Five-year stability in associations of health-related quality of life measures in community-dwelling older adults: the Rancho Bernardo Study

Claudia Der-Martirosian · Donna Kritz-Silverstein · Elizabeth Barrett-Connor

Abstract

Objective This study examines the five-year stability of the association of SF-12 and SF-6D scores with scores on the longer SF-36 and its domains in community-dwelling older men and women.

Methods Participants were 653 men and 917 women aged 50 and older who completed mailed surveys of HRQOL (1995, 2000). SF-36 physical (PCS) and mental (MCS) component scores, domain scores; SF-12 PCS and MCS scores; and SF-6D scores were computed.

Results Average age in 1995 was 68.2 ± 10.7 for men and 69.8 ± 11.3 for women. In 1995 and 2000, men had significantly higher scores on all measures (P's < 0.001).

Sex-specific Pearson correlations of SF-12 PCS and MCS scores with SF-36 PCS and MCS scores ranged from 0.91 to 0.97 (P's < 0.000). Health utility scores (SF-6D) were also associated with SF-36 PCS and MCS scores, but correlations were lower, ranging from 0.61 to 0.79 (P's < 0.000). Age-stratified comparisons of 5-year change scores for SF-36 PCS showed significant declines in physical health for both men and women within all four age-groups (50–59, 60–69, 70–79, 80+) with steepest declines in the oldest age-group (age 80+). Over time, mental health scores showed significant modest improvement in the younger age-groups but declines in men >70 and women >80 years. The SF-6D scores improved somewhat for the youngest men and women but mean change scores declined for the other age-groups.

Conclusion Similar conclusions would result using either instrument SF-12 or SF-36. However, SF-6D and SF-36 assess, at least in part, different underlying aspects of HRQOL. Both age and sex impact HRQOL.

Keywords Health-related quality of life · Physical and mental functioning · Aging

C. Der-Martirosian · D. Kritz-Silverstein · E. Barrett-Connor
Department of Family and Preventive Medicine, University of California, San Diego, 9500 Gilman Drive, MC 0607, La Jolla, CA 92093-0631, USA

E. Barrett-Connor
e-mail: ebarrettconnor@ucsd.edu

C. Der-Martirosian (✉)
Department of Family and Preventive Medicine, University of California, San Diego, 9500 Gilman Drive, MC 0631-C, La Jolla, CA 92093-0631, USA
e-mail: cldermartirosian@ucsd.edu

Introduction

With longer life expectancy, it has become increasingly important to evaluate quality of life among the elderly. As described by Khanna and Tsevat [1], research in health-related quality of life (HRQOL) originated from two fundamentally distinct approaches: health status and health value/preference/utility assessment. Health status measures attempt to assess a person's functional ability in one or more domains, whereas health utility measures are designed to assess the value or desirability of a state of

health for each individual. There are various measures of health status and health utility; this study focuses on the most widely used measures, SF-36, SF-12, and a relatively new measure SF-6D.

The Short-Form 36 Health Survey (SF-36) is a 36-item measure that has been widely used to assess self-rated HRQOL [2, 3] and to monitor trends and variations in population health status for the past few decades [4]. The SF-36 yields subscale scores in eight domains representing various aspects of physical and mental functioning which is further aggregated into two subscale scores representing physical (PCS) and mental health components (MCS). The SF-36 has been used extensively in observational studies and clinical trials in individuals having a wide range of illnesses, ages and other characteristics [5, 6]. The Short-Form 12 Health Survey (SF-12), a shorter version of SF-36, is another valid and reliable HRQOL measure that also provides physical and mental subscale scores and has been used widely [7]. The SF-12 is advantageous because it can be completed in less than 2 min—about 10 min faster than SF-36 [8]. The Short-Form 6D (SF-6D) consists of a sub-sample of eleven items from the SF-36 and is a relatively new approach to measuring HRQOL. Unlike the SF-36 and SF-12, which provide assessments of *health status* with regard to physical and mental function [1, 9], the SF-6D assigns population-based preferences to the subject's responses that represents the respondent's *health value* or *health preference* [1, 10–12].

Numerous studies show high correlations between SF-36 and SF-12 scores within samples of older individuals and conclude that the shorter SF-12 is a reliable indicator of quality of life [7, 13–15]. For example, Muller et al. [14] reported that in elderly patients with heart disease, SF-12 scores correlated 0.96 with PCS and 0.96 with MCS scores from the SF-36. Similarly, high correlations between SF-12 and SF-36 subscale scores were reported among elderly stroke patients [8]. However, most previous studies were cross-sectional, predominantly focused on specific health conditions, and included few individuals aged 80 and older, when the largest declines in quality of life would be expected. Few studies have reported the associations of HRQOL based on the SF-36 or SF-12 with the SF-6D measure of health state utilities in older adults. Furthermore, although the test–retest reliability of these instruments has been well studied [16–18], few studies reported the stability of these associations over a period of time longer than 12 months [16].

The purpose of the present study is to examine (1) the association of SF-12 and SF-6D scores with scores on the SF-36 and its domains in a large population-based sample of older men and women and (2) the stability of these associations in members of this cohort who returned the same HRQOL survey 5 years later.

Methods

Sample

Study participants were community-dwelling men and women aged 50 and older from the Rancho Bernardo cohort. In 1972 to 1974, a census-based cohort of 6,339 individuals representing 82% of the adult population of Rancho Bernardo (a community in Southern California) was enrolled in a study of heart disease risk factors. All were Caucasian, relatively well educated, and middle class. Most were married and living in single-family homes [19].

These individuals are followed by periodic clinic visits and annual mailed surveys. In October 1995 and again in October 2000, a survey including the SF-36 was sent to all surviving members of the Rancho Bernardo cohort. A total of 2,867 participants responded to the 1995 survey and 2,297 responded to the 2000 survey. After excluding the 544 aged less than 50 years at the time of the 1995 survey and the 753 who did not respond to both mailed surveys, there remained 1,570 participants (653 men and 917 women) who are the focus of this study. Responding to the mailed questionnaire was considered implicit informed consent. The Human Research Protections Program at the University of California San Diego approved this study.

HRQOL measures

The SF-36 consists of 36 multiple-choice items originally derived from a large series of health status instruments including 245 items employed in the Medical Outcomes study [15, 20]. The first item of the SF-36 asks participants to rate their health with response choices of excellent, very good, good, fair, and poor. The remaining 35 forced choice items ask participants to rate various aspects of their health. Specific items and scoring for the SF-36 are detailed elsewhere [2, 21]. Briefly, responses to items are grouped into scales assessing eight domains: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health. Scores for the physical functioning, role physical, bodily pain, and general health domains are further aggregated into a physical component summary (PCS) score, and scores for vitality, social functioning, role emotional, and mental health domains are aggregated into a mental component summary (MCS) score. For all SF-36 derived summary and domain scores, higher scores are indicative of better health status. In the general US population, the physical and mental component summaries each have mean scores of 50 and standard deviations of 10, whereas mean scores for the eight domains range from 61 to 84 and standard deviations range from 18 to 34 [21].

The validated SF-12 consists of 12 items from the SF-36. As described elsewhere [7], its PCS and MCS summary scores are constructed from all 12 items where each question item is weighted according to each (physical or mental/emotional) dimension. No domain scores are available from SF-12. Higher scores on the SF-12 PCS and MCS are indicative of better health status.

The SF-6D is a health utility measure of HRQOL, which assesses the value or desirability of a state of health against an external metric [1], is most often used in cost-effectiveness research. The SF-6D is based on six SF-36 domains, specifically, physical functioning, role limitations, social functioning, pain, mental health and vitality; scores are reported as a percent, where 100% is optimal health value [10]. As described elsewhere [11, 12], the SF-6D can be constructed in two ways either using 11 items from the SF-36 or 7 items from the SF-12. In either case, higher scores indicate greater health value or preference.

Statistical analysis

Only individuals who responded to both the 1995 and the 2000 SF-36 questions and were 50 years of age or older in 1995 were included in this analysis. Responses were scored based on existing algorithms to obtain SF-36 physical component (PCS) and mental component (MCS) scores and domain scores as well as SF-12 PCS and MCS scores for each of those years [22]. Additionally, participants were assigned one SF-6D score based on the 11 of the SF-36 items (SF6D_SF36) and a second SF-6D score based on seven items from the SF-12 instrument (SF6D_SF12). SF-36 PCS, MCS, and SF-6D change scores were also computed by subtracting the 1995 scores from the 2000 scores.

Descriptive statistics were calculated for all SF-36 and SF-12 PCS and MCS scores and SF-6D scores. Comparisons of men and women on each of these scores were performed with independent t-tests. Because of observed sex differences, all further analyses were sex specific. Pearson correlations were used to examine the associations of the eight SF-36 domain scores as well as the SF-36 PCS and MCS scores with SF-12 PCS and MCS scores and SF-6D scores; r^2 was used to indicate the amount of shared common variance between SF-12 and SF6D with SF-36. Scatter plots were used to evaluate the agreement between SF-36 and SF-12 as well as SF-36 with SF-6D in 1995. Correlations were used to examine the associations of each HRQOL score obtained in 1995 and with those obtained in 2000 after stratifying by decades of age (50–59, 60–69, 70–79, 80+). Analysis of variance was used to examine the differences in mean SF6D_SF36 by self-reported health status groups (excellent, very good, good, fair/poor) in 1995 and 2000. Analysis of variance was also used to examine differences in mean SF-36 PCS, MCS, and SF-6D

change scores between 1995 and 2000 by age-group (50–59, 60–69, 70–79, 80+).

To have a better understanding of the differences in terms of health quality of life between responders and non-responders, the 1995 HRQOL mean scores were compared between responders and non-responders to the 2000 mailing. All analyses were two-tailed with statistical significance considered at $P < 0.05$. No attempt was made to adjust for multiple comparisons; exact P-values are shown instead. Statistics were performed with SAS, version 9.12 (SAS Institute, Inc., Cary, NC).

Results

In 1995, the age ranged from 50 to 99 with an average of 68.2 ± 10.7 in men and 69.8 ± 11.3 in women; median age was 71 for men and 74 for women. Table 1 shows descriptive statistics for HRQOL scores in 1995 and 2000 for all participants and separately for men and women. In both years, comparisons with independent t-tests showed that there were significant sex differences ($P < 0.05$) on all measures, with men having higher self-rated quality of life than women. The differences between men and women were fairly stable over time.

Table 2 shows sex-specific Pearson correlations of the SF-12 physical and mental subscale scores and SF-6D scores with SF-36 physical and mental subscale scores in both 1995 and 2000. As shown, there were high correlations between SF-12 and SF-36 subscales in men and women at both time points. Coefficients ranged from 0.96 to 0.97 for PCS and 0.94 to 0.97 for MCS scores (P's < 0.001), indicating over 88% of shared common variance between SF-12 and SF-36 scores. Scores on SF-6D were also significantly associated with SF-36 subscale scores, but correlations were lower, ranging from 0.61 to 0.69 (P's < 0.001) in 1995 and 0.71 to 0.79 (P's < 0. 001) in 2000 for PCS and from 0.62 to 0.69 (P's < 0.001) in 1995 and 0.61 to 0.67 (P's < 0.000) in 2000 for MCS. Thus, the shared common variance between SF-6D and SF-36 ranged from 37% to 63% for PCS and from 37% to 41% for MCS.

Table 3 presents sex-specific Pearson correlations of SF-12 PCS, MCS and SF-6D scores in 1995 and 2000 with each of the eight SF-36 domains. As shown, although all correlations were statistically significant, for both men and women, physical functioning, role physical, and bodily pain domains had the highest correlations with SF-12 physical summary score (PCS), (r's = 0.80–0.86, P's < 0.001), whereas role emotional and mental health domains had the highest correlations with SF-12 mental summary score (MCS) (r's = 0.79–0.85, P's < 0.001). Social functioning and vitality SF-36 domain scores had similar

Table 1 Comparisons[a] of sex-specific health-related quality of life (HRQOL) measures in men and women aged 50 and older

HRQOL measures	Total mean (SD) (n = 1570)	Men mean (SD) (n = 653)	Women mean (SD) (n = 917)	t-value	P-value
Year 1995					
SF36_PCS[a]	48.4 (9.6)	49.7 (8.6)	47.5 (10.1)	−4.62	<0.001
SF36_MCS[a]	54.1 (8.4)	54.9 (7.6)	53.5 (9.0)	−3.40	<0.001
SF12_PCS	48.5 (9.6)	50.0 (8.8)	47.5 (10.1)	−5.03	<0.001
SF12_MCS	53.8 (8.0)	54.6 (7.3)	53.2 (8.5)	−3.61	<0.001
SF6D_sf12	82.6 (12.2)	84.9 (10.8)	80.9 (12.8)	−6.46	<0.001
SF6D_sf36	78.3 (11.2)	80.4 (10.3)	76.8 (11.6)	−6.31	<0.001
Year 2000					
SF36_PCS	45.4 (11.2)	46.8 (10.6)	44.4(11.5)	−4.06	<0.001
SF36_MCS	54.5 (8.2)	55.0 (7.8)	54.0 (8.5)	−2.32	<0.021
SF12_PCS	45.6 (11.1)	47.1 (10.7)	44.4 (11.3)	−4.60	<0.001
SF12_MCS	53.9 (8.0)	54.4 (7.7)	53.5 (8.2)	−2.17	<0.030
SF6D_sf12	80.4 (13.9)	82.3 (13.4)	80.0 (14.1)	−4.54	<0.001
SF6D_sf36	76.2 (12.6)	77.9 (12.5)	74.9 (12.5)	−4.49	<0.001

[a] Results of independent sample t-tests; PCS physical component score, MCS mental component score

Table 2 Sex-specific Pearson correlations for SF-12 and SF6D scores with SF-36

	Men (n = 653)		Women (n = 917)	
	SF36_PCS	SF36_M CS	SF36_PCS	SF36_MCS
Year 1995				
SF12_PCS	0.96**	0.08*	0.96**	0.08*
SF12_MCS	0.05	0.96**	0.03	0.97**
SF6D_sf12	0.61**	0.69**	0.62**	0.67**
SF6D_sf36	0.69**	0.63**	0.66**	0.65**
Year 2000				
SF12_PCS	0.97**	0.22**	0.96**	0.14**
SF12_MCS	0.22**	0.95**	0.13**	0.95**
SF6D_sf12	0.76**	0.67**	0.71**	0.62**
SF6D_sf36	0.79**	0.64**	0.73**	0.61**

* $P < .05$; ** $P < .001$

associations with both SF-12 physical and mental summary scores (see Table 3). Thus, for both survey years, the general health domain score had a stronger correlation with the physical than the mental summary score (e.g., for men in 1995, $r = 0.66$ for SF-12 PCS vs. $r = 0.37$ for SF-12 MCS). SF-36 domain scores and SF6D scores were also significantly correlated in both sexes in both years with correlations that did not vary much over the SF-36 domains (see Table 3).

Table 4 shows sex-specific correlations between HRQOL measures in 1995 and 2000 after stratification by age (50–59, 60–69, 70–79 and 80+). As shown, scores on the SF-36 subscales, SF-12 subscales, and SF-6D in 1995 were significantly correlated with scores on those measures in 2000 (all P's < 0.001). The associations between HRQOL measures (in 1995) are depicted graphically using scatter plots

(see Figs. 1a, b, and 2a, b). Scatter plots were similar for men and women in 1995 and 2000; only in 1995 sex-specific data (for men) are presented. In Fig. 1a and b, the tight clustering along the $y = x$ line indicates the close agreement between SF-36 and SF-12. Figure 2a and b show less clustering along the $y = x$ line indicative of a lower agreement between SF-36 and SF-6D_SF36. Results were similar using HRQOL scores from 2000 (data not shown).

Figure 3a and b show the association between responses to the self-rated health question and the SF-6D utility scores. As shown, as *self-rated* health declines, there is a monotonic decrease in mean SF6D_SF36 utility scores in both sexes in both time periods (1995 and 2000).

Figure 4a and b show mean change scores between 1995 and 2000 for SF-36 PCS, MCS, and SF6D_SF36. In both men and women, all age-groups (50–59, 60–69, 70–79, 80+) showed a decline in PCS scores over 5 years with the greatest decline observed for the oldest age-group (age 80+). Women in the three younger age-groups and men in the two younger age-groups had a modest improvement in MCS score, whereas older men and women showed a decline. In both men and women, SF-6D scores improved over time among those aged 50–59 but declined for the other three age-groups ($P < 0.001$).

Study participants who did not respond to 2000 mailer had lower SF-36 PCS, MCS, and SF6D_SF36 mean scores (P's < 0.001) and were younger when compared to respondents who participated in both 1995 and 2000 mailings.

Discussion

In this large sample of older community-dwelling men and women, there are strong and similar associations between

Table 3 Sex-specific Pearson correlations between SF-36 HRQOL domain scores and SF-12 physical (PCS) and mental (MCS) scores and SF6D scores

SF36 domains	Year 1995		Year 2000	
	SF12_PCS	SF12_MCS	SF12_PCS	SF12_MCS
Men (n = 653)				
General health	0.66	0.37	0.71	0.46
Physical functioning	0.80	0.14	0.86	0.28
Role physical	0.82	0.26	0.84	0.37
Bodily pain	0.70	0.23	0.73	0.36
Vitality	0.62	0.56	0.72	0.57
Social functioning	0.54	0.63	0.62	0.69
Role emotional	0.15	0.79	0.32	0.79
Mental health	0.22	0.83	0.39	0.83
Women (n = 917)				
General health	0.61	0.38	0.69	0.40
Physical functioning	0.86	0.13	0.86	0.24
Role physical	0.82	0.32	0.86	0.32
Bodily pain	0.74	0.23	0.75	0.23
Vitality	0.61	0.57	0.71	0.56
Social functioning	0.54	0.61	0.60	0.62
Role emotional	0.24	0.81	0.30	0.80
Mental health	0.22	0.85	0.28	0.80
SF36 domains	Year 1995		Year 2000	
	SF6D_SF12	SF6D_SF36	SF6D_SF12	SF6D_SF36
Men (n = 653)				
General health	0.58	0.63	0.69	0.70
Physical functioning	0.57	0.62	0.72	0.74
Role physical	0.69	0.68	0.75	0.75
Bodily pain	0.60	0.70	0.71	0.76
Vitality	0.70	0.74	0.77	0.78
Social functioning	0.78	0.74	0.80	0.77
Role emotional	0.63	0.56	0.65	0.63
Mental health	0.68	0.67	0.70	0.70
Women (n = 917)				
General health	0.59	0.62	0.62	0.66
Physical functioning	0.62	0.63	0.71	0.71
Role physical	0.74	0.74	0.75	0.75
Bodily pain	0.63	0.71	0.65	0.71
Vitality	0.71	0.74	0.74	0.77
Social functioning	0.76	0.74	0.78	0.78
Role emotional	0.71	0.67	0.64	0.61
Mental health	0.65	0.65	0.61	0.64

Note: All correlations are significant at $P < 0.0001$

PCS and MCS scores obtained with the SF-36 and with the SF-12 in all age-groups and both sexes. Furthermore, the SF-12 PCS and MCS sub-scale scores were highly correlated with each of the individual SF-36 domain scores. These results indicate that there is little loss of information with use of the SF-12 instead of the SF-36; similar conclusions would be made using either instrument, in accord with other studies [7, 13–15].

Although the SF-6D was constructed to assess a different underlying aspect of health-related quality of life, there are statistically significant correlations with SF-36, indicating a modest overlap between these measures. However, the lower proportion of shared common variance between the SF-36 and the SF-6D when compared to the proportion of variance shared with SF-12 suggests that use of the SF-6D to assess health status would lead to a lower

Table 4 Sex- and age-specific associations for HRQOL measures in 1995 and 2000

	N	SF-36 PCS	SF-36 MCS	SF-12 PCS	SF-12 MCS	SF6D_SF36	SF6D_SF12
Men age-groups[a]							
50–59	166	0.64	0.44	0.58	0.41	0.52	0.56
60–69	189	0.48	0.52	0.47	0.52	0.53	0.56
70–79	188	0.57	0.44	0.52	0.40	0.48	0.52
80+	110	0.56	0.34	0.57	0.33	0.40	0.50
Women age-groups[a]							
50–59	207	0.57	0.54	0.52	0.48	0.45	0.51
60–69	224	0.48	0.60	0.50	0.58	0.47	0.54
70–79	276	0.54	0.39	0.53	0.67	0.46	0.51
80+	210	0.62	0.35	0.61	0.40	0.51	0.52

Results of Pearson correlations. All P-values significant at $P < 0.001$
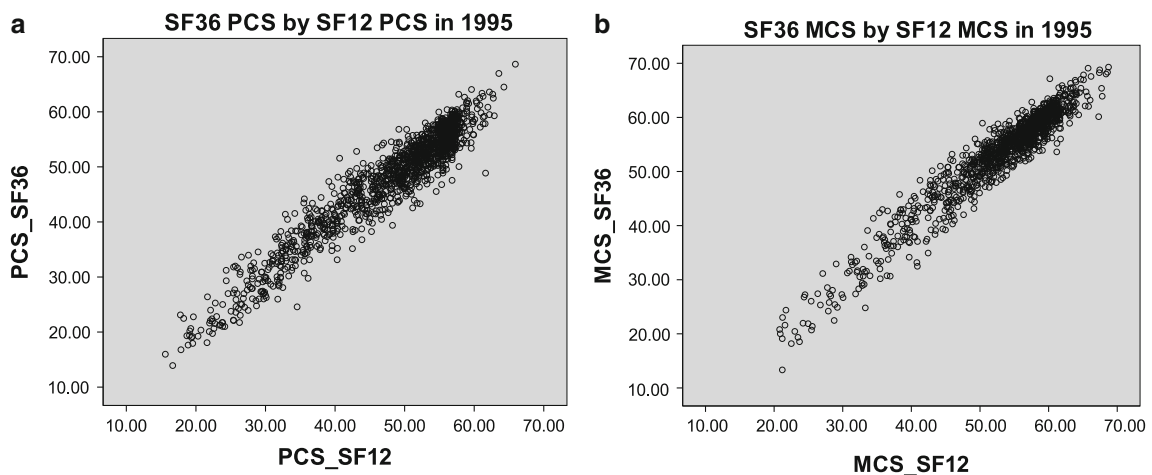
[a] Based on 1995 age

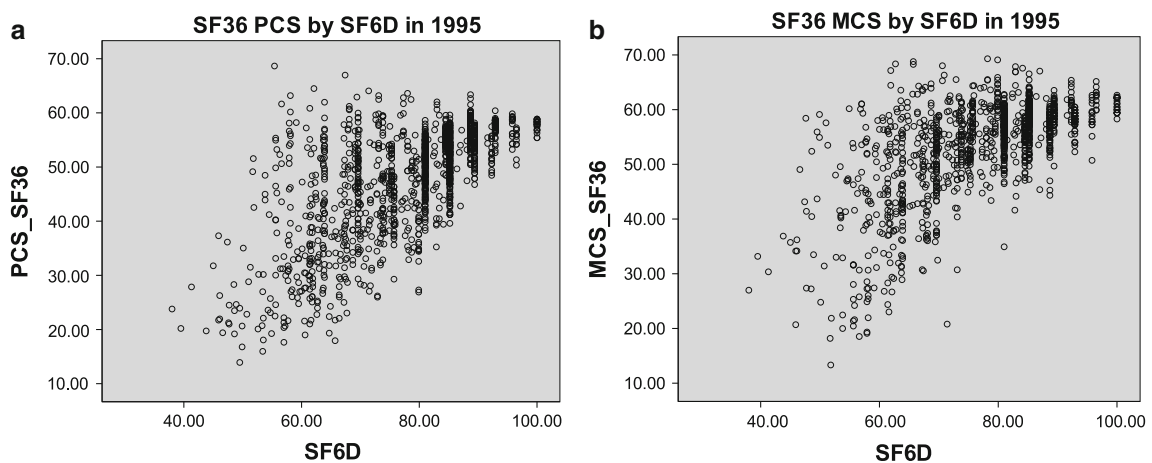**Fig. 1** **a** SF-36 PCS by SF-12 PCS in 1995. **b** SF-36 MCS by SF-12 MCS in 1995

Fig. 2 **a** SF-36 PCS by SF6D_SF36 in 1995. **b** SF-36 MCS by SF6D_SF36 in 1995. For figure **a** and **b**, SF6D is based on SF6D_SF36

a Mean SF6D scores by health status in 1995

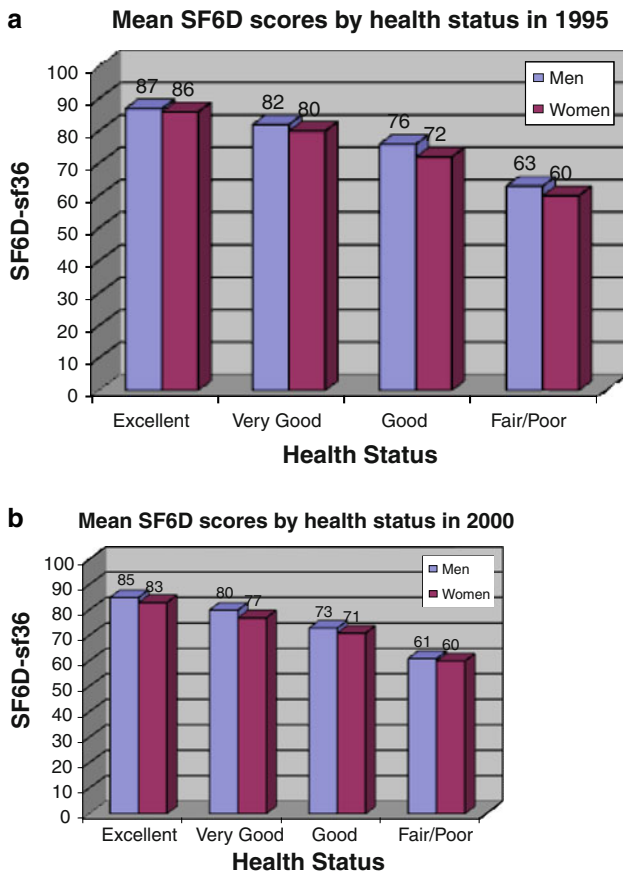**b** Mean SF6D scores by health status in 2000

Fig. 3 **a** Mean SF6D_SF36 scores by health status in 1995. **b** Mean SF6D_SF36 scores by health status in 2000. *Notes*: All SF6D comparisons by self-reported health status significant at $P < 0.001$ using ANOVA. Health Status Question: "In general would you say your health is?" Due to small sample size, "fair and poor" response categories were combined into one response category
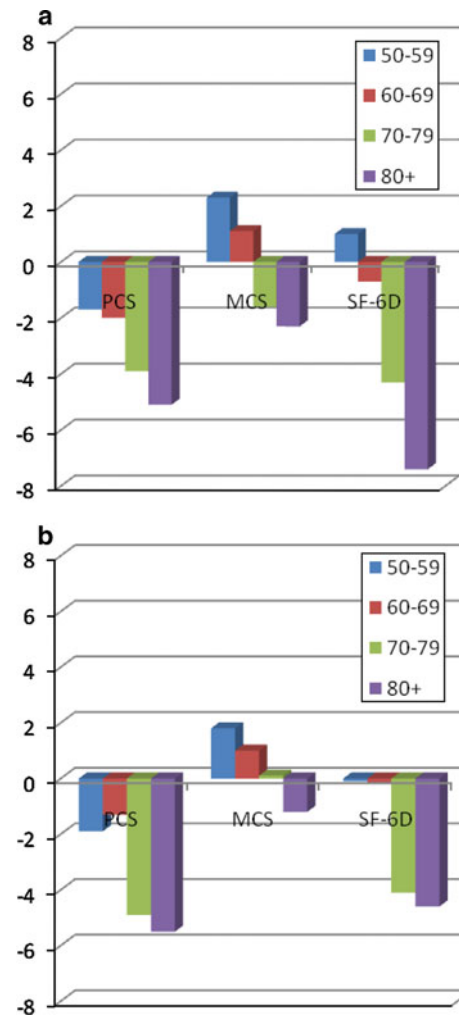

Fig. 4 **a** Change in HRQOL (based on SF36) measures between 1995 and 2000 by age-group for men. **b** Change in HRQOL (based on SF36) measures between 1995 and 2000 by age-group for women. *Notes*: For figure **a** and **b**, SF-36 measures compared by four age-groups significant for both men and women at $P < .001$ using ANOVA. Age-groups based on 1995 age. The SF-6D measure is based on SF6D-SF36

precision of results and confirms that these instruments assess, at least in part, different underlying aspects of HRQOL. In this study, SF-6D scores decreased monotonically with decreasing self-ratings of health in accord with Petrou et al. [9] who reported a similar pattern in a large sample of women in England. It should be noted that participants responded to SF-36 items, and scores for SF-12 and SF-6D scales were constructed from the SF-36 responses, and different scales were not presented to participants as separate standardized scales.

The associations observed in the present study were stable over time for both men and women; similar patterns of associations were observed using data from 1995 and 2000. Unlike previous studies where the focus of test–retest reliability entailed much shorter-term periods, this study provides evidence of the stability of these associations over 5 years in adults aged 50 and over as well as in those aged 80 and older, when the largest declines in quality of life (both physical and mental) and health utility would likely be found. The observed decline with age was likely

attenuated in the present dataset which required that participants complete the questionnaire on two occasions 5 years apart. Although the mailed questionnaires had high response rates, there were greater numbers of older and likely more unwell seniors who did not complete the second survey. This would have minimized the ability to quantify the loss in HRQOL in old age and is an inevitable limitation of all geriatric studies.

Age and sex-specific results of 5-year change scores for SF-36 PCS show statistically significant *declines* in physical health for both men and women within all four age–groups, although the decline was worse for the oldest age-group (age 80+). Over time, mental health scores showed statistically significant but *modest improvement* in the younger age–groups, although men over 70 and women

over 80 exhibited declines. Health utility scores (SF-6D) improved somewhat for the youngest men and women (aged 50–59); however, the mean change scores declined for the other three age-groups. These results are comparable to other reports [16, 23] of the expected association of age on decline in quality of life. The examination of specific morbid conditions and other variables as explanatory variables for the age-related change in HRQOL will be reported elsewhere.

Although the patterns of association and change over the 5-year period were similar in both sexes, there were statistically significant sex differences on all HRQOL measures at both time points, with men reporting significantly higher self-rated quality of life. These results are in accord with Myint et al. [23, 24] who reported that men have higher SF-36 PCS and MCS mean scores compared to women in a population-based study of individuals aged 41–80 [23, 24]. Hemingway et al. [25] also reported higher SF-36 mean scores for men compared to women for all SF-36 domains among a large sample of younger civil servants in England aged 39–63.

Several limitations are considered. Data were self-reported via two annual mailers, which are vulnerable to misreporting. The Rancho Bernardo cohort is Caucasian, relatively well educated with high socioeconomic status; the application of these findings to other ethnic and socioeconomic groups might be limited. On the other hand, the homogeneity of the RBS population reduces potential associations of less advantaged socioeconomic status and ethnicity. Selection bias may also affect results of this study. Those who responded to the second mailed questionnaire appear to be healthier, both physically and mentally, than those who did not respond. Selection bias was reduced by the use of mailed survey, allowing those who were too ill or too busy as a caregiver to participate. Nevertheless, we acknowledge that non-responders to the second mailing had lower quality of life when compared to responders, which may have led to the very high correlation between the two assessments by removing those with the least favorable HRQOL at baseline. Furthermore, results of this study may be biased in that they reflect the stability of the associations only in survivors. Demonstration of five-year stability in survivors is novel.

This study also has several strengths. It provides the unique opportunity to examine the associations between two widely used measures of quality of life (SF-36 and SF-12) as well as a relatively new measure of health utility (SF-6D) in a large community-dwelling population of older men and women with a wide age range including people (older than those in many other studies). The availability of two data points obtained 5 years apart provides an opportunity to examine the stability of associations after a longer period of time compared to previous reports.

In conclusion, results suggest that SF-12 is a reliable measure of *health status* and can be used among older men and women in lieu of SF-36 with little loss of information. This has important implications for research on quality of life issues faced by the elderly. Although the SF-6D is a reliable measure of *health preference,* it cannot be substituted for either of the health status measures (SF-36 and/or SF-12), without compromising predictive utility. The significant gender differences in physical and mental health status as well as health preference scores suggest that future studies of HRQOL should include sex-specific analyses. Furthermore, future longitudinal analyses should adjust for the confounding effects of age. Additionally, more longitudinal studies on HRQOL are needed to provide additional information on change in HRQOL over time, especially in the oldest old. Future studies on HRQOL should also be mindful that change in quality of life in older age can occur in either negative or positive direction. Studies in the elderly also need to determine which factors help explain improvement in quality of life and whether they reflect healthy aging or adjustment and compromise in the face of increasing limitations.

Conflict of interest statement None of the authors have a conflict of interest.

Sponsor's role None.

References

1. Khanna, D., & Tsevat, J. (2007). Health related quality of life—an introduction. *American Journal of Managed Care, 13*, S218–S223.
2. Ware, J. E., Jr., & Gandek, B. (1998). Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. *Journal of Clinical Epidemiology, 51*, 903–912.
3. Prescott-Clarke, P., & Pimatesta, P. (1998). *Health survey for England'96. Vol I: Findings and Vol II: Methodology and documentation*. London: HMSO.
4. Kroenke, C. H., Kubzansky, L. D., Adler, N., et al. (2008). Prospective change in health-related quality of life and subsequent mortality among middle-aged and older women. *American Journal of Public Health, 98*(11), 2085–2091.
5. Ferrucci, L., Guralnik, J. M., Studenski, S., et al. (2004). Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. *Journal of the American Geriatrics Society, 52*, 625–634.

6. Brazier, J. E., Harper, R., Jones, H. M., et al. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ, 305*, 160–164.

7. 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. *Medical Care, 34*, 220–233.

8. Pickard, A. S., Johnson, J. A., Penn, A., et al. (1999). Replicability of SF-36 summary scores by the SF-12 in stroke patients. *Stroke, 30*, 1213–1217.

9. Petrou, S., Morrell, J., & Spiby, H. (2009). Assessing the empirical validity of alternative mult-attribute utility measures in the maternity context. Health and Quality of Life Outcomes, 7.

10. Ackerman, R. T., Edelsteiin, S. L., Venkat, K. M., et al. (2009). Changes in health state utilities with changes in body mass in the diabetes prevention program. *Obesity, 114*, 1–6.

11. Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics, 21*, 271–292.

12. Brazier, J., Usherwook, T., Harper, R., et al. (1998). Deriving a preference-based single index from the UK SF-36 health survey. *Journal of Clinical Epidemiology, 51*, 1115–1128.

13. Cote, I., Gregoire, J. P., Moisan, J., et al. (2004). Quality of life in hypertension: the SF-12 compared to the SF-36. *Canadian Journal of Clinical Pharmacology, 11*(2), e232–e238.

14. Muller-Nordhorn, J., Roll, S., & Willich, S. N. (2004). Comparison of the short-form (SF)-12 health status instrument with the SF-36 in patients with coronary heart disease. *Heart, 90*, 523–527.

15. Hurst, N. P., Ruta, D. A., & Kind, P. (1998). Comparison of the MOS Short Form-12 (SF12) health status questionnaire with the SF36 in patients with rheumatoid arthritis. *British Journal of Rheumatology, 37*, 862–869.

16. Thomsen, D. K., Mehlsen, M. Y., Olesen, F., et al. (2003). Is there an association between rumination and self-reported physical health? One year follow-up in a young and elderly sample. *Journal of Behavioral Medicine, 27*(3), 215–231.

17. Everard, K. M., Lach, H. W., Fisher, E. D., et al. (2000). Relationship of activity and social support to the functional health of older adults. *Journal of Gerontology: Social Sciences, 55B*(4), S208–S212.

18. Andersen, E. M., Bowley, N., Rothenberg, B. M., et al. (1996). Test-retest performance of a mailed version of the medical outcomes study 36-iterm short-form health survey. *Medical Care, 34*(2), 1165–1170.

19. Criqui, M. E., Barrett-Connor, E., & Austin, M. (1978). Differences between responders and non-responders in a population-based in a cardiovascular disease study. *American Journal of Epidemiology, 108*, 367–372.

20. McHorney, C. A., Ware, J. E., Jr, Lu, J. F. R., et al. (1994). The MOS 36-items short form health survey (SF-36): III: Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care, 32*, 40–63.

21. Ware, J. E., Kosinski, M., & Keller, S. D. (1994). *SF-36 physical and mental health summary scales: A user's manual*. Boston, MA: The Health Institute, New England Medical Center.

22. Brazier, J., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care, 42*, 851–859.

23. Myint, P. K., Luben, R. N., Surtees, P. G., et al. (2009). Effect of age and sex on relationship between different socioeconomic indices and self-reported functional health in the EPIC-Norfolk population-based study. *AEP, 19*(5), 289–297.

24. Myint, P. H., Luben, R. N., Surtees, P. G., et al. (2007). Self-reported mental health related quality of life and mortality in men and women in the European prospective investigation into cancer (EPIC-Norfolk): A prospective population study. *Psychosomatic Medicine, 69*, 410–414.

25. Hemingway, H., Stafford, M., Stansfeld, S., et al. (1997). Is the SF-36 a valid measure of change in population health? Results from the Whitehall II study. *BMJ, 315*, 1273–1279.