Volume 9 • Number 4 • 2006
VALUE IN HEALTH

Comparing the EQ-5D and the SF-6D Descriptive Systems to Assess Their Ceiling Effects in the US General Population

Murtuza Bharmal, PhD[1], Joseph Thomas III, PhD[2]

[1]Quintiles Transnational Corp., Falls Church, VA, USA; [2]Purdue University, School of Pharmacy, West Lafayette, IN, USA

ABSTRACT

Objectives: The EuroQol (EQ-5D) and SF-6D (derived from the SF-12) were compared to assess any ceiling effect in the EQ-5D and the SF-6D descriptive systems. In addition, the Physical Component Summary (PCS-12), the Mental Component Summary (MCS-12) and the EuroQol Visual Analog Scale (EQ-VAS) were compared on their discriminative ability to detect differences among individuals with different morbidities and sociodemographic characteristics.

Methods: Data from the 2000 Medical Expenditure Panel Survey were used for the analysis. A total of 11,248 individuals that were 18 years or older and had data on all the study variables were included in the analysis.

Results: A total of 5104 individuals (47%) reported no limitations on all of the EQ-5D dimensions and only 683 (5.8%) were classified in full health based on the SF-6D descriptive system. Approximately 49% of the respondents that reported no limitations on the EQ-5D reported feeling "tense or downhearted and low," "a little," (level 2) or "some" (level 3) of the time on SF-6D. PCS-12 scores and EQ-VAS scores among individuals reporting no limitations on the EQ-5D descriptive system were significantly lower for respondents reporting coronary heart disease, angina, diabetes, myocardial infarction, high blood pressure or joint pain compared with respondents that reported no medical condition. Effect sizes for medical conditions using the PCS-12 were larger than the effect sizes using the EQ-VAS.

Conclusions: Unlike the EQ-5D descriptive system, the SF-6D descriptive system derived from the SF-12 does not seem to have a ceiling effect. Nevertheless, the SF-6D does not discriminate between individuals with different morbidities who report full health on the EQ-5D, as does the PCS-12 and the EQ-VAS.

Keywords: ceiling, EQ-5D, EuroQol, SF-12, SF-6D.

Introduction

The 12-item Short Form Health Survey (SF-12) and the EuroQol (EQ-5D) are easy to administer generic health-related quality of life measures that are widely used in population surveys, clinical practice, and clinical trials [1–4]. The EQ-5D index is a preference-based index measure, where an individual provides an assessment of each component of his/her health status according to a structured health-status classification system and a single preference-based score is derived for each individual based on societal preferences [5]. In contrast, the SF-12 and the EuroQol Visual Analog Scale (EQ-VAS) that is included in the EQ-5D as a separate measure to assess health-related quality of life (HRQOL) do not assign an external value to individuals' assessments of their health. The EQ-VAS contains one item that assesses individuals' self-rated health status, whereas, the SF-12 generates summary scores that are psychometrically derived [5].

Address correspondence to: Murtuza Bharmal, Quintiles Transnational Corp., 3130 Fairview Park Drive, Suite # 800, Falls Church, VA 22042, USA. E-mail: murtuzabharmal@alumni.purdue.edu

10.1111/j.1524-4733.2006.00108.x

The advantage with preference-based index measures is the ability to obtain a single index score for an individual that can be used to calculate quality-adjusted life-years. Brazier et al. provided an algorithm to derive a preference-based single index score, SF-6D index, from the 36-item Short Form Health Survey (SF-36) [6]. Nevertheless, because the respondent burden with the 12-item SF-12 is substantially lower compared with the respondent burden with the 36-item SF-36, there is growing use of the SF-12 instrument among researchers. As a result, recently, Brazier and Roberts provided an algorithm to derive SF-6D from the SF-12 [7]. The SF-6D derived from the SF-12 varies from the SF-6D derived from the SF-36 in some aspects. The SF-6D derived from the SF-36 uses 11 items from the SF-36, whereas the SF-6D derived from the SF-12 uses seven items from the SF-12. The physical functioning domain in the SF-6D derived from the SF-12 contains three levels compared with six levels in the SF-6D derived from the SF-36. There is also one less level in the pain domain of SF-6D derived from SF-12 as compared with the SF-6D derived from SF-36 (five levels vs. six levels, respectively).

Previous studies have compared the EQ-5D and the SF-6D derived from the SF-36 [8–11]. Brazier et al.

found an interclass correlation of 0.51 between the SF-6D and the EQ-5D [11]. They also found some evidence of ceiling effects in the EQ-5D descriptive system. In their analysis, they found 214 observations where patients reported full health on the EQ-5D, of which only 12 (6%) were classified in full health on the SF-6D. Similar comparison between the SF-6D derived from the SF-12 and the EQ-5D are not available.

The purpose of this analysis was to compare the EQ-5D and the SF-6D derived from the SF-12 to examine any ceiling effects in the EQ-5D and the SF-6D descriptive systems in the US general population. In addition, the study explored the discriminative ability of SF-12 component summary scores and the EQ-VAS to discriminate between individuals with various medical conditions and different sociodemographic characteristics among individuals reporting full health on the EQ-5D or the SF-6D descriptive systems. Our study hypotheses were that the EQ-5D descriptive system would have a ceiling effect in the US general population [11–14] and the SF-12 component summary scores would have better discriminative ability than the one-item EQ-VAS in differentiating between individuals with various medical conditions and different sociodemographic characteristics among individuals reporting full health on the EQ-5D or the SF-6D descriptive systems [15].

Methods

Data Source

The data used for analysis were taken from the household component of the 2000 Medical Expenditure Panel Survey (MEPS) conducted by the Agency for Healthcare Research and Quality [16]. The household component survey was a year long panel survey of 25,096 persons representative of the civilian, noninstitutionalized US population [16]. The survey used a stratified multistage area probability design with over sampling of African Americans and Hispanic persons.

The MEPS is a panel survey, with an overlapping cohort design. A new cohort is initiated each year and provides information for a 2-year reporting period. The MEPS involves five in-person interviews with one or more persons per household, who report on healthcare utilization, expenditures, insurance coverage, and medical conditions for each household member. Cross-sectional analyses combine information from two cohorts. The EQ-5D and SF-12 scores for this study were collected in 2000, representing round two of data collection for the 2000 cohort and round four of data collection for the 1999 cohort.

The survey also gathered data on sociodemographic characteristics and medical conditions of each household resident by face-to-face interviews with a member of each household. Respondents were asked whether each household member had ever been diagnosed, by a doctor or other health professional, as having certain chronic clinical conditions including diabetes, asthma, high blood pressure, coronary heart disease, angina, myocardial infarction, stroke, emphysema, or joint pain. In 2000, the MEPS for the first time also gathered data using a self-administered questionnaire, which was distributed to all individuals 18 years of age or older in the sample. The questionnaire included the SF-12 and the EQ-5D instruments.

Health-Related Quality of Life Measures

The EQ-5D descriptive system is a preference-based HRQOL measure with one question for each of the five dimensions that include mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [17]. Each question has three levels of response (no limitation, some limitation, severe or complete limitation). An index score (EQ-5D index) also was calculated from the EQ-5D descriptive system using weights recently provided by Shaw, Johnson, and Coons in the US general population [18]. In addition to the multidimensional descriptive system, the EQ-5D also includes a separate 20 cm EQ-VAS to measure self-assessed health status. Respondents are asked to rate their health on a 101-point scale, the end-points of which are labeled "best imaginable health state" and "worst imaginable health state" anchored at 100 and 0, respectively [19].

The SF-12 is a shorter version of the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36) and consists of a subset of 12 items from the SF-36. The SF-12 was constructed to reduce the respondent burden of longer instruments used for measuring HRQOL [20]. The 12 questions of the SF-12 measure physical functioning, role limitations due to physical problems, role limitations due to emotional problems, pain, general health, vitality, social functioning, and mental health. The SF-12 also provides two summary scores, namely the Physical Component Summary (PCS-12) and the Mental Component Summary (MCS-12) [21]. The PCS-12 and the MCS-12 are scored such that a higher score represents better function, and each is standardized such that in the general population, the mean score is 50 and the standard deviation is 10 [21].

The SF-6D was derived from the SF-12 using the algorithm provided by Brazier and Roberts [7]. The SF-6D uses SF-12 data to classify individuals into a six-dimensional health state classification system where the dimensions are physical functioning, role limitations, social functioning, pain, mental health, and vitality. Each dimension has three to five levels and thus, 7500 possible health states can be defined using the SF-6D descriptive system. Brazier and Roberts also provide valuations derived from a UK general population to generate the SF-6D index score from the SF-6D descriptive system [7].

Study Sample

For this analyses, individuals of 18 years or older, with responses on all items in the SF-12, EQ-5D descriptive system, EQ-VAS, medical conditions, and sociodemographic characteristics, were included in the analysis. Additional subgroup analyses were conducted on respondents that reported no limitations on each of the five dimensions of the EQ-5D descriptive system and respondents that were classified in full health on each of the six dimensions of the SF-6D descriptive system.

Data Analyses

Analyses were conducted using STATA version 8 and SAS for Windows Version 8.2 [22,23]. All the analyses incorporated sampling weights to account for the complex survey design used in the MEPS. Although a select subgroup was used in this study, the sample estimates are representative of the US general population using the subgroup analysis, as recommended in MEPS documentation [24]. In the MEPS, any subgroup with greater than 100 observations can be used to support national estimates [25]. Mean PCS-12 scores, MCS-12 scores, and EQ-VAS scores were calculated for the sample. A checklist of nine medical conditions: angina, asthma, diabetes, congestive heart disease, emphysema, hypertension, joint pain, myocardial infarction, and stroke were included in the MEPS. For analysis, dichotomous variables were constructed indicating presence or absence of each of the nine reported medical conditions. Another variable to indicate the mean number of chronic medical conditions also was constructed by summating responses on all nine medical conditions.

To assess ceiling effects in the EQ-5D and the SF-6D descriptive systems, respondents that reported no limitations on all the five dimensions of EQ-5D were identified and their responses on the SF-6D descriptive system were assessed. Similarly, respondents that were classified in full health on all the six dimensions of SF-6D were identified and their responses on EQ-5D descriptive system were assessed.

Individuals that indicated full health on the EQ-5D or the SF-6D descriptive systems were further classified as being in better or worse health based on whether their PCS-12 score, MCS-12 score or EQ-VAS score was above or below the mean score on each of the scales, respectively. In addition, individuals that indicated full health on the EQ-5D descriptive system also were classified as being in better or worse health based on whether their SF-6D index score was above or below the mean SF-6D index score. Similarly, individuals with full health on the SF-6D descriptive system also were classified as being in better or worse health based on whether their EQ-5D index score was above or below the mean EQ-5D index score. Individuals classified as being in better health based on the above scales were compared with those classified as being in worse health on sociodemographic characteristics including age, sex, marital status, employment status, income, number of years of education, and number of chronic medical condition using chi-square tests of associations or *t*-tests. Based on similar comparisons in a previous study, we hypothesized that individuals that were male, younger, more educated, with a higher income, employed, married, and with lower number of chronic medical conditions would more likely be classified in better health as opposed to worse health by a more discriminative health measure [15].

Mean PCS-12, MCS-12, EQ-VAS, and SF-6D index scores were calculated for individuals that reported full health on the EQ-5D descriptive system and who reported having no medical condition. Similarly, mean PCS-12, MCS-12, EQ-VAS, and EQ-5D index scores were calculated for individuals that reported full health on the SF-6D descriptive system and reported having no medical condition. *T*-tests were used to compare the mean scale scores among respondents who reported having specific medical conditions with those who had no medical conditions. Effect sizes were calculated for groups with statistical differences by subtracting the mean for individuals with the medical condition from the mean for individuals with no medical condition and dividing by the standard deviation of the mean for the sample. An a priori alpha level of 0.05 was used for statistical significance. No adjustments were made for multiple comparisons because effect sizes were used to assess if the differences observed between groups were greater than the minimum clinically important effect size of 0.2 [26].

Results

Sample Characteristics

A total of 11,248 individuals were 18 years or older and had data on all the variables used in the study. A total of 48 respondents (0.4%) were proxy respondents in the data used in this study and were included in the current analyses.

Of the 11,248 individuals, 5104 (47%) reported no limitations on all the EQ-5D dimensions and 683 (5.8%) were classified in full health on all the SF-6D dimensions. Table 1 provides the demographic characteristics for the entire sample, respondents that reported no limitations on all the EQ-5D dimensions and respondents that were classified in full health on all the SF-6D dimensions.

Approximately 33% of all respondents were between the ages of 18 and 34 years (Table 1). The majority of these individuals were white (86.2%), married (57.5%), had a high school diploma (47.8%), and employed (75.7%). The mean PCS-12 score among all the respondents was 50.10 (SE = 0.13). The mean MCS-12 score among all the respondents was 51.56 (SE = 0.12). The mean EQ-VAS score among all the

Table 1 Study sample characteristics

Category	Initial sample N = 11,248 Percent*	Full Health On EQ-5D N = 5104 Percent*	Full Health On SF-6D N = 683 Percent*
Age (years)			
18–24	12.64	17.28	21.95
25–34	20.39	25.16	18.44
35–44	22.60	24.19	23.92
45–54	19.66	17.51	15.54
55–64	11.33	8.93	11.46
65–74	7.61	4.69	5.34
75 or above	5.77	2.24	3.35
Sex			
Male	48.38	52.87	64.65
Female	51.62	47.13	35.35
Income (dollars)			
Less than 10,000	22.95	19.65	22.43
10,000–30,000	36.84	34.05	37.40
30,000–50,000	22.45	25.09	19.54
50,000–70,000	10.58	12.32	12.00
More than 70,000	7.18	8.89	8.62
Education			
No degree	14.28	11.16	18.60
GED	4.24	3.32	2.92
High school diploma	47.76	45.96	47.55
Bachelor's degree	18.25	22.13	17.66
Master's degree	6.81	7.82	6.38
Doctorate degree	1.81	2.43	2.89
Other degree	6.84	7.16	4.00
Race			
Native American or Eskimo	0.86	8.85	2.26
Asian or Pacific Islander	3.21	3.99	4.69
African American	9.75	10.04	15.82
White	86.18	85.07	77.22
Hispanic ethnicity	9.50	10.61	19.02
Marital status			
Married	57.52	56.71	50.40
Widowed	5.28	2.62	2.87
Divorced	11.07	9.06	10.68
Separated	1.72	1.73	2.14
Never married	24.41	30.54	33.91
Employment status			
Employed at present	75.71	84.28	80.91
Not employed at present	24.29	15.72	19.09

*Percentages are weighted and account for the complex survey design of MEPS.
GED, general educational development.

respondents was 80.05 (SE = 0.22). Among all respondents, the mean EQ-5D index score was 0.8739 (SE = 0.0022) and the mean SF-6D index score was 0.8126 (SE = 0.0017).

Among individuals that reported no limitations on the EQ-5D descriptive system, approximately 42% were between the ages of 18 and 34 years (Table 1). The majority of these individuals were white (85.1%), married (56.7%), had a high school diploma (45.9%), and employed (84.3%).

Among individuals that were classified in full health on the SF-6D, approximately 40% were between the ages of 18 and 34 years (Table 1). The majority of these individuals were white (77.2%), married (50.4%), had a high school diploma (47.6%), and employed (80.9%).

Distribution of Responses on the EQ-5D and the SF-6D Descriptive Systems

Table 2 provides the distribution of responses on the EQ-5D descriptive system. A majority of the respondents (more than 55%) indicated no limitations (level 1) on at least one of the five EQ-5D dimensions. Very few respondents (less than 4%) indicated severe limitations (level 3) on any of the five dimensions of the EQ-5D. Table 2 also shows the distribution of responses on the SF-6D descriptive system. A large proportion of respondents also were classified in the top category (level 1) on the SF-6D for the dimensions of physical functioning (81%), role limitation (74.5%), social functioning (69.4%), and pain (57.4%). Very few respondents (less than 4%) were classified in the bottom category (level 5) on any of the six dimensions of the SF-6D.

Assessing Ceiling Effects in the EQ-5D and the SF-6D Descriptive Systems

A total of 5104 individuals (47%) reported no limitations on all the EQ-5D dimensions. Table 3 provides distribution of responses on the SF-6D dimensions among individuals who reported no limitations on all the EQ-5D dimensions. In this group, a majority of individuals were classified in level 1 for the SF-6D dimension of physical functioning (97%), role limitation (95.2%), social functioning (89.7%), and pain (88.1%). Nevertheless, 84.2% of the respondents were

Table 2 Proportion of responses (%) on each level of the EQ-5D and the SF-6D descriptive systems in the study sample (N = 11,248)

	Sample distribution on scale dimensions					
EQ-5D scale levels (N = 11,248)	Mobility	Self-care	Usual activities	Pain/ discomfort	Anxiety/ depression	
1	83.9	97.2	83.8	55.6	73.6	
2	15.8	2.6	15.0	41.0	24.3	
3	0.2	0.2	1.2	3.3	2.2	
SF-6D scale levels (N = 11,248)	Physical functioning	Role limitation	Social functioning	Pain	Mental health	Vitality
1	81.0	74.5	69.4	57.4	36.0	9.3
2	12.9	9.7	15.4	23.6	37.1	47.5
3	6.0	8.1	10.2	10.6	19.6	31.6
4	—	7.6	3.3	6.0	5.5	8.4
5	—	—	1.7	2.4	1.8	3.2

Table 3 Ceiling effects of EQ-5D. Distribution of responses (%) on the SF-6D dimensions for those with EQ-5D = 11111 (N = 5104)

Level	Physical functioning	Role limitation	Social functioning	Pain	Mental health	Vitality
1	97.0	95.2	89.7	88.1	50.8	15.8
2	2.3	1.9	8.1	10.4	37.2	61.2
3	0.7	2.6	1.3	1.0	8.8	20.6
4	—	0.3	0.5	0.3	1.4	1.5
5	—	—	0.5	0.2	1.9	1.0

classified in level 2 or higher in the vitality dimension and 49.2% of the respondents were classified in level 2 or higher in the mental health dimension.

A total of 683 individuals (5.8%) were classified in full health on the SF-6D descriptive system. Further analysis of this group revealed that a large majority, 628 (92.6%) had reported no limitations on all the EQ-5D dimensions. Of the remaining 55 individuals that reported limitations on the EQ-5D, 47 respondents (6.2%) indicated some limitations, level 2, in the pain dimension of EQ-5D, with less than eight respondents reporting some or severe limitations in all other dimensions.

Comparison of HRQOL Measures on Sociodemographic Characteristics for Individuals Reporting Full Health on EQ-5D Descriptive System

Table 4 compares individuals reporting no limitations on the EQ-5D but classified as being in better or worse health based on PCS-12, MCS-12, EQ-VAS, or SF-6D index scores. Significant differences were observed in most sociodemographic characteristics between individuals classified in better physical health and those classified as in worse physical health based on PCS-12 score. As expected, individuals classified as in better physical health were more likely to be men, employed, younger, have a higher income, more educated and have fewer chronic medical conditions than individuals classified in worse physical health.

Significant differences also were observed in some of the sociodemographic characteristics between individuals classified in better mental health and worse mental health based on MCS-12 score. As expected, individuals classified as in better mental health were more likely to be men and more likely to be married than individuals classified in worse mental health. Nevertheless, opposed to findings by Johnson and Coons [15], individuals classified as in better mental health were less likely to be employed than individuals classified in worse mental health.

No significant differences were observed on any of the sociodemographic characteristics among individuals classified in better or worse health based on the EQ-VAS score. Nevertheless, as expected, individuals classified in better health based on EQ-VAS had

Table 4 Individuals reporting no limitation on the EQ-5D dimensions: a comparison of those in better health (SF-12/EQ-VAS/SF-6D ≥ mean) with those in worse health (SF-12/EQ-VAS/SF-6D < mean)

	PCS-12		MCS-12		EQ-VAS		SF-6D index	
Category	<50.10 (n = 378)	≥50.10 (n = 4726)	<51.56 (n = 1014)	≥51.56 (n = 4090)	<80.05 (n = 1314)	≥80.05 (n = 3790)	<0.8126 (n = 775)	≥0.8126 (n = 4329)
Mean age (years)	44.47 (1.17)*	39.04 (0.30)*	36.89 (0.47)*	40.04 (0.37)*	40.57 (0.63)	39.03 (0.32)	37.79 (0.62‡	39.71 (0.34)‡
Total income (dollars)	27,323 (1727)†	33,901 (547)†	32,718 (985)	33,618 (567)	32,532 (996)	33,749 (579)	30,716 (1201)‡	33,918 (575)‡
Number of years education (mean)	12.84 (0.22)†	13.54 (0.06)†	13.44 (0.10)	13.51 (0.07)	13.39 (0.12)	13.53 (0.06)	13.37 (0.14)	13.52 (0.06)
Chronic medical conditions (mean)	0.75 (0.06)*	0.37 (0.01)*	0.38 (0.03)	0.40 (0.01)	0.54 (0.03)*	0.35 (0.01)*	0.41 (0.03)	0.40 (0.01)
% Female	54.49 (2.96)†	46.58 (0.79)†	50.03 (1.67)	46.43 (0.85)	46.39 (1.47)	47.38 (0.87)	52.43 (1.71)†	46.21 (0.89)†
% Married	61.53 (3.26)	56.34 (0.99)	49.61 (1.97)*	58.44 (1.14)*	58.24 (1.84)	56.19 (1.07)	47.27 (2.16)*	58.36 (1.01)*
% Employed	75.18 (2.58)*	84.96 (0.73)*	87.59 (1.28)‡	83.47 (0.79)‡	83.44 (1.04)	84.56 (0.82)	85.04 (1.61)	84.14 (1.04)

*Estimate statistically different between the groups with *P*-value of <0.001.
†Estimate statistically different between the groups with *P*-value of <0.01.
‡Estimate statistically different between the groups with *P*-value of <0.05.

significantly lower mean number of chronic medical conditions compared with individuals classified in worse health (0.35 vs. 0.54).

There were significant differences on some sociodemographic characteristics among individuals classified in better or worse health based on the SF-6D index score. As expected, individuals classified in better health based on the SF-6D index score were more likely to be men, have a higher income and married compared with individuals classified in worse health. Nevertheless, contrary to expectation, individuals classified in better health based on SF-6D index score were significantly older compared with individuals classified in worse health (Table 4).

Comparison of HRQOL Measures on Sociodemographic Characteristics for Individuals Classified in Full Health on SF-6D Descriptive System

Among individuals classified in full health on the SF-6D descriptive system, only three individuals could be classified as in better physical health (PCS-12 score higher than the mean PCS-12 score), as a result, no comparisons were made between better physical health and worse physical health. Similarly, only five individuals out of the 683 individuals (in full health on SF-6D) could be classified as in better mental health based on the MCS-12, as a result, no comparisons were made between better mental health and worse mental health in those groups.

Based on EQ-VAS score being higher than the mean EQ-VAS score, 92 out of 683 individuals (in full health on SF-6D) were classified in better health. Further analysis of this group found no significant differences on most of the sociodemographic characteristics among individuals classified in better or worse health based on the EQ-VAS score. Nevertheless, individuals classified in better health based on EQ-VAS compared with individuals classified in worse health had significantly higher number of years of education (13.1 vs. 11.9 years), a higher proportion of employed individuals (82.6% vs. 67.9%), and individuals with higher incomes ($32,516 vs. $20,687).

Based on the EQ-5D index score being higher than the mean EQ-5D index score, 55 out of 683 individuals (in full health on SF-6D) were classified in better health. Further analysis of this group found no significant differences on most of the sociodemographic characteristics among individuals classified in better or worse health based on the EQ-5D index score. Nevertheless, individuals classified in better health based on EQ-5D index score compared with individuals classified in worse health were significantly younger (39.5 vs. 47.9 years).

Comparison of HRQOL Measures on Medical Conditions for Individuals Reporting Full Health on EQ-5D Descriptive System

Among individuals that indicated no limitations on the EQ-5D were individuals that reported having a diagnosis of coronary heart disease, angina, stroke, diabetes, myocardial infarction, high blood pressure, joint pain, asthma, or emphysema (14 respondents). A total of 3523 individuals (69.4%) among those that reported full health on EQ-5D descriptive system reported having no medical condition. Table 5 provides a comparison of PCS-12 scores, MCS-12 scores, EQ-VAS scores, and SF-6D index scores between individuals that reported having a specific medical condition and those that reported having no medical condition. Data are presented only for the medical

Table 5 Comparison of mean (SE) scores on SF-12, SF-6D index, and EQ-VAS among individuals with chronic conditions and those with no medical conditions for individuals reporting no limitations on the EQ-5D

Condition	n	PCS-12	MCS-12	SF-6D	EQ-VAS
No medical condition	3523	55.28 (0.06)	54.95 (0.06)	0.8852 (0.0015)	88.44 (0.27)
Coronary heart disease	51	51.54 (1.06)*	56.85 (0.86)†	0.8819 (0.0102)	80.62 (2.28)*
Effect size‡		0.92	0.34		0.45
Angina	30	51.71 (0.84)*	55.55 (1.01)	0.8726 (0.0122)	82.43 (1.92)§
Effect size		0.88			0.34
Stroke	31	52.07 (0.91)*	54.92 (1.23)	0.8707 (0.0108)	84.29 (2.34)
Effect size		0.79			
Diabetes	165	52.28 (0.42)*	55.59 (0.42)	0.8854 (0.0073)	79.86 (1.48)*
Effect size		0.74			0.49
Myocardial infarction	55	52.34 (0.43)*	55.68 (0.87)	0.8807 (0.0095)	79.59 (1.74)*
Effect size		0.72			0.51
High blood pressure	658	53.48 (0.19)*	55.62 (0.23)†	0.8853 (0.0031)	84.34 (0.49)*
Effect size		0.44	0.12		0.23
Joint pain	730	54.25 (0.15)*	55.09 (0.19)	0.8837 (0.0032)	85.72 (0.52)*
Effect size		0.25			0.16
Asthma	333	54.56 (0.27)*	54.31 (0.37)§	0.8755 (0.0040)	87.00 (0.63)
Effect size		0.18	0.12		

*Group means are statistically different in expected direction with *P*-value <0.001.

†Group means are statistically different in nonexpected direction with *P*-value <0.05.

‡Effect size were calculated by subtracting the mean for individuals with the condition from the mean for individuals without any medical condition and dividing by the standard deviation (SD) of the mean for the whole sample (N = 5104). SD of PCS-12 was 4.08, SD of MCS-12 was 5.53 and SD of EQ-VAS was 17.49.

§Group means are statistically different in expected direction with *P*-value <0.05.

conditions with 30 or more observations to maintain stability of the standard errors of estimates.

Individuals who reported having coronary heart disease, angina, stroke, diabetes, myocardial infarction, high blood pressure, joint pain or asthma had significantly lower mean PCS-12 scores than individuals who reported having no medical condition. Effect sizes using the PCS-12 ranged from 0.18 for asthma to 0.92 for coronary heart disease, most of them were above the minimum clinically important effect size of 0.2 [26].

Mean scores on the MCS-12 were significantly lower for individuals with asthma compared with individuals with no medical condition. Nevertheless, mean scores on the MCS-12 were not significantly different between individuals that reported most other medical conditions and those that did not have any medical condition. Furthermore, contrary to expectations, individuals who reported having coronary heart disease or high blood pressure had significantly higher mean scores on MCS-12 compared with individuals having no medical condition (Table 5).

Mean scores on the SF-6D index score were not significantly different between individuals that reported having a medical condition and those that did not have any medical condition (Table 5). Individuals who reported having coronary heart disease, diabetes, myocardial infarction, high blood pressure or joint pain also had significantly lower mean EQ-VAS scores, than individuals who reported having no medical condition. Effect sizes using the EQ-VAS ranged from 0.16 for pain to 0.51 for myocardial infarction (Table 5).

Comparison of HRQOL Measures on Medical Conditions for Individuals Classified in Full Health on SF-6D Descriptive System

Among individuals that were classified in full health on the SF-6D were individuals that reported having a diagnosis of high blood pressure, asthma, joint pain, diabetes (16 respondents), myocardial infarction (five respondents), stroke (four respondents), coronary heart disease (three respondents) or angina (one respondent). A total of 522 respondents (76.43%) in this group reported having no medical condition. Table 6 provides a comparison of PCS-12 scores, MCS-12 scores, EQ-VAS scores, and EQ-5D index scores between individuals that reported having a specific medical condition and those that reported having no medical condition. Data are presented only for the medical conditions with 30 or more observations.

Individuals who reported having high blood pressure had significantly lower mean PCS-12 scores than individuals who reported having no medical condition. Effect size using the PCS-12 was 0.52 for high blood pressure. Mean scores on the MCS-12, EQ-5D index or the EQ-VAS were not significantly different between individuals that reported any of the three medical conditions and those that reported having no medical condition (Table 6).

Discussion

This analysis compared two widely used generic health-related quality of life measures to assess their ceiling effects in the US general population. To our knowledge, this is the first such comparison of the SF-6D derived from the SF-12 and the EQ-5D. The study provides evidence suggesting that the EQ-5D descriptive system compared with the SF-6D descriptive system derived from the SF-12 has a substantial ceiling effect in the US general population. Out of the 11,248 individuals in the study sample, a large proportion, 5104 (47%) reported full health on the EQ-5D dimensions. Nevertheless, only 683 respondents, 5.8% of the study sample were classified in full health on the SF-6D. Such a distribution of responses on the EQ-5D descriptive system has been observed in the United States, UK, Canadian, and Catalan general population surveys [12–15].

A great majority of those reporting full health on the EQ-5D descriptive system were not classified in full health on the SF-6D descriptive system. Based on the SF-6D responses, individuals reporting full health on the EQ-5D may still have problems on the mental health function and vitality dimensions. Of particular interest is that of those who reported no limitations on the EQ-5D, 49.2% of the respondents reported feeling "tense or downhearted and low" a little (level 2) or some (level 3) of the time.

Table 6 Comparison of mean (SE) scores on SF-12, EQ-5D, and EQ-VAS among individuals with chronic conditions and those with no medical conditions for individuals reporting no limitations on the SF-6D

Condition	n	PCS-12	MCS-12	EQ-5D	EQ-VAS
No medical condition	522	56.16 (0.07)	59.97 (0.08)	0.9888 (0.0025)	92.87 (0.56)
High blood pressure	60	55.30 (0.18)*	60.20 (0.13)	0.9845 (0.0065)	91.52 (1.32)
Effect size†		0.52			
Asthma	31	55.65 (0.29)	60.18 (0.24)	0.9847 (0.0088)	93.10 (1.49)
Joint pain	83	55.90 (0.18)	60.03 (0.17)	0.9792 (0.0063)	91.64 (1.11)

*Group means are statistically different in expected direction with *P*-value <0.001.

†Effect size was calculated by subtracting the mean for individuals with the condition from the mean for individuals without any medical condition and dividing by the standard deviation (SD) of the mean for the whole sample (N = 683). SD of PCS-12 was 1.66.

Johnson and Coons compared the EQ-5D and SF-12 in an adult US sample [15]. In their study, 206 of 413 respondents (49.88%) indicated no limitations on all the EQ-5D dimensions. Among individuals reporting no limitations on EQ-5D, they found that respondents reporting poorer physical health (below median PCS-12 score) were older and were less likely to be married or employed than those in the better health group. The worse mental health group (below median PCS-12 score) had a higher proportion of women and persons not married or in a partnership. They also found that respondents with worse health on both components had more chronic medical conditions on average and had lower mean EQ-VAS scores than subjects in the better health groups. The results of our study were mostly consistent with the findings by Johnson and Coons [15]. Nevertheless, the Johnson and Coons study did not find statistically significant differences in the evaluation of the effect of specific chronic medical conditions on the SF-12 component scores for these subjects, primarily resulting from the small sample sizes. In addition, the study did not compare individuals in better or worse health based on the EQ-VAS scores. We found statistically significant differences in the evaluation of the effect of specific medical conditions on the PCS-12 scores and the EQ-VAS scores for individuals that reported no limitations on all EQ-5D dimensions. In addition, our study evaluated the effect sizes for specific medical conditions using the PCS-12, the MCS-12, and the EQ-VAS to compare their discriminative ability.

Among individuals who reported full health on the EQ-5D, mean scores on PCS-12 were significantly lower for those who reported having coronary heart disease, angina, stroke, diabetes, myocardial infarction, high blood pressure, joint pain or asthma than individuals who reported having no medical condition. Effect sizes using the PCS-12 ranged from 0.18 for asthma to 0.92 for coronary heart disease. Among individuals who reported full health on the EQ-5D, mean scores on EQ-VAS were significantly lower for those who reported having coronary heart disease, angina, diabetes, myocardial infarction, high blood pressure or joint pain than individuals who reported having no medical condition. Effect sizes using the EQ-VAS ranged from 0.16 for joint pain to 0.51 for myocardial infarction. These findings indicate that individuals with significant morbidity are misclassified as in full health on the EQ-5D descriptive system, when in fact they have substantially lower health-related quality of life compared with individuals without any medical condition. The findings also indicate that the effect sizes using PCS-12 were higher than those using EQ-VAS providing evidence of the better discriminative ability of the PCS-12 compared with the EQ-VAS.

Among individuals that were classified in full health on the SF-6D descriptive system, fewer individuals reported having a medical condition. Further in most cases, statistically significant differences were not found using the PCS-12, the MCS-12, or the EQ-VAS among individuals who were classified in full health on the SF-6D and reported having a specific medical condition compared with those that reported having no medical condition. Although, the SF-6D descriptive system derived from the SF-12 did not seem to have a ceiling effect as did the EQ-5D descriptive system, the SF-6D index did not discriminate among individuals who reported full health on the EQ-5D and reported having a specific medical condition compared with those that reported having no medical condition. It is likely that this lack of significant difference in SF-6D index score among groups may be due to the use of preference weights for constructing the SF-6D index scores that were derived in the UK general population [7], rather than the US general population. It is recommended that future studies are conducted to derive preference weights for the SF-6D descriptive system for the US general population.

Some of the findings in our study were not in the expected direction. For example, among individuals who reported no limitations on EQ-5D, individuals who reported having coronary heart disease or high blood pressure, had a higher MCS-12 score than individuals without medical conditions. Furthermore, no clinically important differences were observed in the MCS-12 between individuals with any of the other chronic medical conditions and those without a medical condition. Some researchers have suggested that the RAND-12 scoring approach [27] offers several theoretical advantages over the SF-12 scoring algorithm in terms of performance of the summary scores in discriminating between known groups [28,29]. Nevertheless, we also conducted all the analyses described in this study using the RAND-12 scoring algorithm, but the trends across all results remained. The MHC-12 (Mental Health Composite of the RAND-12) also did not find significant differences between individuals with chronic medical conditions and those without a medical condition. Thus, this result does not appear to be due to the differences in scoring algorithms used to construct the SF-12 and RAND-12 component summary scores [28–31], but is more likely due to the fact that most of the chronic medical conditions included in the study would have a primary impact on individuals' physical functioning rather than mental functioning.

Other studies also have reported ceiling effects in EQ-5D descriptive system. Brazier et al. compared the EQ-5D and an Anglicized version of the SF-36 in a postal survey of patients from two primary care practices in the UK [32]. They concluded that the SF-36 provides a more sensitive description of health across the dimensions compared with the EQ-5D and may be more useful for measuring health in populations with less severe morbidity. The authors observed a large

ceiling effect in the EQ-5D suggesting that EQ-5D would be more suitable for measuring the health of populations with more morbidity, where the distribution of responses is less skewed. Nevertheless, studies conducted in morbid populations, also have found some evidence of ceiling effect in the EQ-5D [33–35]. In a study of patients with Hodgkin's disease [33], 19 of 49 patients reported the maximum level of health for all the five items of the EQ-5D, whereas in a study of patients with advanced HIV disease [34], 30% reported no limitations in all dimension of EQ-5D.

There are several possible reasons for the ceiling effect observed in the EQ-5D descriptive system. The EQ-5D dimensions are limited to three response categories per item and this may consequently lead it to having a ceiling effect [12]. It may be that three response choices per dimension do not allow it to capture variations among individuals. In contrast, the SF-6D derived from the SF-12 has levels ranging from three to five which makes it more sensitive to differences among individuals. The exclusion of some important health-related quality of life dimensions may be responsible for the observed ceiling effect [13,14,35]. For example, the vitality dimension which is an additional dimension in the SF-6D, is not in the EQ-5D. In this study, 31.6% of the respondents that reported full health on the EQ-5D reported having "a lot of energy" some (level 3) of the time.

In conclusion, although the EQ-5D descriptive system was created with simplicity as a central requirement it may suffer from ceiling effects in general populations. The study results also suggest that although the SF-6D descriptive system derived from the SF-12 does not seem to have a ceiling effect, the SF-6D index score, unlike the PCS-12 and the EQ-VAS, was not able to discriminative between groups with varying morbidity, among those who reported full health on the EQ-5D. In addition, the PCS-12 was more discriminating than the EQ-VAS as a self-rating HRQOL measure. Between the PCS-12 and the EQ-VAS, in addition to discriminating between individuals with chronic medical conditions with higher effect sizes, the PCS-12 was also able to discriminate between individuals with different sociodemographic characteristics. These findings may be important when considering which measures to use to assess HRQOL in general populations.

Source of financial support: None.

References

1 Idler EL, Russell LB, Davis D. Survival, functional limitations, and self-rated health in the NHANES I epidemiologic follow-up study, 1992. First national health and nutrition examination survey. Am J Epidemiol 2000;152:874–83.

2 Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav 1997;38:21–37.

3 Bergner M, Rothman ML. Health status measures: an overview and guide for selection. Annu Rev Public Health 1987;8:191–210.

4 Coons SJ, Rao S, Keininger DL, et al. A comparative review of generic quality-of-life instruments. Pharmacoeconomics 2000;17:13–35.

5 Gold MR, Siegel JE, Russel LB, et al., eds. Cost Effectiveness in Health and Medicine. New York: Oxford University Press, 1996.

6 Brazier JE, Roberts J, Deverill M. The estimation of a preference-based single-index measure of health from the SF-36. Health Econ 2002;21:271–92.

7 Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004;42:851–9.

8 Marra C, Esdaile J, Guh D, et al. A comparison of four indirect methods of assessing utility values in rheumatoid arthritis. Med Care 2004;42:1125–31.

9 Longworth L, Bryan S. An empirical comparison of EQ-5D and SF-6D in liver transplant patients. Health Econ 2003;12:1061–7.

10 Conner-Spady B, Suarez-Almazor M. Variation in the estimation of quality-adjusted life years by different preference-based instruments. Med Care 2003;41: 791–801.

11 Brazier J, Roberts J, Tsuchiya A, et al. A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ 2004;13:873–84.

12 Macran S, Weatherly H, Kind P. Measuring population health: a comparison of three generic health status measures. Med Care 2003;41:218–31.

13 Johnson JA, Pickard AS. Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta. Can Med Care 2000;38:115–21.

14 Badia X, Schiaffino A, Alonso J, et al. Using the EuroQol 5-D in the Catalan general population: feasibility and construct validity. Qual Life Res 1998;7:311–22.

15 Johnson JA, Coons SJ. Comparison of the EQ-5D and SF-12 in an adult US sample. Qual Life Res 1998;7:155–66.

16 Agency for Healthcare Research and Quality. 2000 Medical Expenditure Panel Survey. Rockville, MD: Agency for Healthcare Research and Quality, 2003.

17 The EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.

18 Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states development and testing of the D1 valuation method. Med Care 2005;43:203–8.

19 Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095–8.

20 Ware JE, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–6.

21 Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales (3rd ed.). Lincoln, RL: QualityMetric Inc., 1998.

22 SAS Institute, Inc. SAS/STAT Guide for Personal Computers, Version 8.02. Cary, NC: SAS Institute Inc., 2001.
23 StataCorp. Stata Statistical Software Release 8.0. College Station, TX: Stata Corporation, 2004.
24 Agency for Healthcare Research and Quality. Computing Standard Errors for MEPS Estimates. Rockville, MD: Agency for Healthcare Research and Quality, January 2005. Available from: http://www.meps.ahrq.gov/factsheets/FS_StandardErrors.htm [Accessed September 6, 2005].
25 Cohen S. The Medical Expenditure Panel Survey: an overview. Effective Clinical Practice May/June 2002. Available from: http://www.acponline.org/journals/ecp/mayjun02/cohen.htm [Accessed September 6, 2005].
26 Samsa G, Edelman D, Rothman M, et al. Determining clinically important differences in health status measures. Pharmacoeconomics 1999;15:141–55.
27 Hays RD. RAND-36 Health Status Inventory. San Antonio, TX: The Psychological Corporation, 1998.
28 Taft C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? Qual Life Res 2001;10:395–404.
29 Taft C, Karlsson J, Sullivan M. Reply to Drs Ware and Kosinski. Qual Life Res 2001;10:415–20.
30 Wilson D, Parsons J, Tucker G. The SF-36 summary scales: problems and solutions. Sozial-und Praventivmedzin 2000;45:239–46.
31 Simon G, Revicki D, Grothaus L, et al. SF-36 summary scores: are physical and mental health truly distinct? Med Care 1998;36:567–72.
32 Brazier J, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. Qual Life Res 1993; 2:169–80.
33 Norum J, Angelsen V, Wist E, et al. Treatment costs in Hodgkin's disease: a cost-utility analysis. Eur J Cancer 1996;32A:1510–17.
34 Wu AW, Jacobson DL, Frick KD, et al. Validity and responsiveness of the EuroQol as a measure of health-related quality of life in people enrolled in an AIDS clinical trial. Qual Life Res 2002; 11:273–82.
35 Jenkinson C, Stradling J, Petersen S. How should we evaluate health status? A comparison of three methods in patients presenting with obstructive sleep apnoea. Qual Life Res 1998;7:95–100.