

Medical Decision Making

<http://mdm.sagepub.com>

Report of Nationally Representative Values for the Noninstitutionalized US Adult Population for 7 Health-Related Quality-of-Life Scores

Janel Hanmer, William F. Lawrence, John P. Anderson, Robert M. Kaplan and Dennis G. Fryback

Med Decis Making 2006; 26; 391

DOI: 10.1177/0272989X06290497

The online version of this article can be found at:
<http://mdm.sagepub.com/cgi/content/abstract/26/4/391>

Published by:



<http://www.sagepublications.com>

On behalf of:



<http://www.smdm.org>
Society for Medical Decision Making

Additional services and information for *Medical Decision Making* can be found at:

Email Alerts: <http://mdm.sagepub.com/cgi/alerts>

Subscriptions: <http://mdm.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations <http://mdm.sagepub.com/cgi/content/refs/26/4/391>

Report of Nationally Representative Values for the Noninstitutionalized US Adult Population for 7 Health-Related Quality-of-Life Scores

Janel Hanmer, William F. Lawrence, MD, MS, John P. Anderson, PhD,
Robert M. Kaplan, PhD, Dennis G. Fryback, PhD

Background. Despite widespread use of generic health-related quality-of-life (HRQoL) scores, few have publicly published nationally representative US values. **Purpose.** To create current nationally representative values for 7 of the most common HRQoL scores, stratified by age and sex. **Methods.** The authors used data from the 2001 Medical Expenditures Panel Survey (MEPS) and the 2001 National Health Interview Survey (NHIS), nationally representative surveys of the US noninstitutionalized civilian population. The MEPS was used to calculate 6 HRQoL scores: categorical self-rated health, EuroQoL-5D with US scoring, EuroQoL-5D with UK scoring, EuroQoL Visual Analog Scale, mental and physical component summaries from the SF-12, and the SF-6D. The authors estimated Quality of Well-being

scale scores from the NHIS. **Results.** They included 22,523 subjects from MEPS 2001 and 32,472 subjects from NHIS 2001. Most age and sex categories had instrument completion rates above 85%. Females reported lower scores than males across all ages and instruments. In general, those in older age groups reported lower scores than younger age groups, with the exception of the mental component summary from the SF-12. **Conclusion.** This is one of the first sets of publicly available, nationally representative US values for any standardized HRQoL measure. These values are important for use in both generalized comparisons of health status and in cost-effectiveness analyses. **Key words:** health-related quality of life; EQ-5D; SF-12; SF-6D; QWB; national norms. (*Med Decis Making* 2006;26:391-400)

There is a wide interest in measuring health-related quality of life (HRQoL) to both describe the health state of individuals or groups and measure the health change these individuals or groups experience over time. To meet this interest, researchers have developed a variety of off-the-shelf HRQoL measures for general and disease-specific health. Some of the general measures have scoring

algorithms that yield utility values, with dead anchored at 0 and full health anchored at 1.0. Of these general measures, some also allow scores less than 0 to indicate states worse than death. These utility values are interval scaled and preference based, so they are appropriate for construction of quality-adjusted life-years to inform decision making and in cost-effectiveness analyses.¹ Scores that are not based on utility values can be used for more general comparisons of health status.

Despite widespread use of HRQoL measures, only 3 publicly published articles provide nationally representative US values for general health measures that researchers and policy makers can use to compare individuals or groups because these measures have previously not been included in nationally representative surveys. Nationally representative values have been published for other countries, including Australia (SF-36),² Ireland (SF-36),³ New Zealand (SF-36),⁴ Norway (SF-36),⁵ Singapore (SF-36),⁶ Spain (SF-36 and HUI-3),^{7,8} and the United Kingdom (SF-36).⁹

Received 23 June 2005 from Department of Population Health Sciences, University of Wisconsin-Madison (JH, DGF); Agency for Healthcare Research and Quality, Rockville, Maryland (WFL); University of California-San Diego (JPA); and University of California-Los Angeles (RMK). The views expressed in this article are those of the authors, and no official endorsement by the Agency for Healthcare Research and Quality (AHRQ) or the US Department of Health and Human Services is intended or should be inferred. This work was partially supported by a P01 grant (AG206079-01) from the National Institute on Aging, an AHRQ training grant (HS000083), and "Centers for Disease Control Project—Quality of Well-Being Scale," from Project MOVE, Physical Activity and Health Branch, Division of Nutrition and Physical Activity, funded by the Robert Wood Johnson Foundation.

Options for comparison within the United States depend on the measure used. For the SF-36 version 1, McHorney and colleagues published values for both telephone and mail administration from a nationally representative sample in 1992 by age group for the 8 SF subscales.¹⁰ For the Quality of Well-being (QWB) Scale, Anderson published estimates using National Health Interview Survey data between 1979 and 1996.¹¹ For the EuroQol-5D (EQ-5D) with US scoring and Health Utilities Index (HUI) Mark 2 and 3, Luo and colleagues published values from a nationally representative sample who self-completed the questionnaire in a home visit survey in 2002.¹² For the SF-36 and SF-12 version 2 family of measures, there are privately published proprietary values that the vendor reports to be from a 1998 nationally representative sample.¹³ For the SF-36 version 1 and QWB Scale, there is a peer-reviewed journal article that lists values from a community sample that is commonly used in lieu of nationally representative values.¹⁴

In this report, we present age- and sex-stratified nationally representative values for several commonly used preference-based and non-preference-based HRQoL scores. We use data from the 2001 wave of the Medical Expenditures Panel Survey (MEPS), a nationally representative sample of the US noninstitutionalized civilian population. The 2001 wave of MEPS included 3 off-the-shelf measures of HRQoL: the EQ-5D, the EuroQol Visual Analog Scale (EQ-VAS), and the SF-12 version 1. From these 3 measures, we obtained 6 HRQoL scores: categorical self-rated health, EQ-5D with US scoring,¹⁵ EQ-5D with UK scoring,¹⁶ EQ-VAS, mental and physical component summaries from the SF-12,^{17,18} and the SF-6D.¹⁹ We estimated QWB^{11,20} scores from the 2001 National Health Interview Survey (NHIS), another nationally

representative survey of the US noninstitutionalized population that is used as the sampling frame for MEPS.

SUBJECTS AND METHODS

Subjects

Data for this study come from 2 sources, the 2001 MEPS and the 2001 NHIS surveys. The MEPS is a nationally representative survey of health care utilization and expenditures for the US noninstitutionalized civilian population. The MEPS is a 2-year panel survey, with an overlapping cohort design, taken from the National Health Interview Survey cohort. Each year, a new cohort is initiated and followed longitudinally through a series of 5 in-person interviews at 6-month intervals. The MEPS conducts interviews with 1 or more persons per household, who report on health care utilization, expenditures, insurance coverage, and medical conditions for each household member. Cross-sectional analyses combine information from 2 MEPS cohorts.

In 2000, the MEPS initiated a self-administered questionnaire (SAQ) to obtain information that potentially would be unreliable if reported by a proxy. The SAQ was distributed to all adults aged 18 years old or older in eligible households participating in the MEPS. In 2001, the questionnaire included both the SF-12 version 1 and the EQ-5D instruments.

The NHIS is a nationally representative survey of the US noninstitutionalized civilian population. Sampling and interviewing are continuous throughout each year with oversampling of both black persons and Hispanic persons. The NHIS is a cross-sectional household survey that gathers information on all household members and detailed information about 1 household adult and child. All information is collected by an interviewer. The detailed information includes health status information used in this analysis.

In these analyses, we used all respondents aged 20 or older from the surveys. This included 22,523 subjects from the MEPS 2001 and 32,472 subjects from the NHIS 2001.

Categorical Self-Rated Health—Self-Administered

Categorical self-rated health was indicated by the first question of the SF-12: "In general, would you say your health is: Excellent, Very good, Good, Fair,

The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. The impetus for calculating these national values was the case study work of the Institute of Medicine (IOM) Committee to Evaluate Measures of Health Benefits for Environmental, Health, and Safety Regulation in collaboration with federal regulatory and health agencies. The collaborating agencies include AHRQ, Centers for Disease Control and Prevention, Food and Drug Administration, Environmental Protection Agency, and the National Center for Statistics and Analysis (NHTSA). Revision accepted for publication 20 March 2006.

Address correspondence and reprint requests to Janel Hanmer, Department of Population Health Sciences, University of Wisconsin-Madison, 644 WARF, 610 Walnut St., Madison, WI 53726; telephone: (608) 265-3298; fax: (608) 263-2820; e-mail: jehanmer@wisc.edu.

DOI: 10.1177/0272989X06290497

or Poor?” We report the full distribution of responses to this question.

EQ-5D with UK Scoring—Self-Administered

The EuroQol EQ-5D has 5 multiple-choice questions that form a descriptive system with 5 dimensions concerning the respondent's health today: mobility, self-care, usual activities, pain, and anxiety/depression.²¹ Each question has 3 possible responses: no problems, some problems, or extreme problems/unable to. The pattern of responses for an individual can be converted to a single summary score by applying weights from a population-based valuation set, yielding a utility-based score. We used a time tradeoff valuation set derived from a sample representative of the United Kingdom for these UK scores.¹⁶ These weights allow for “states worse than death,” which are given negative values.

EQ-5D with US Scoring—Self-Administered

The EQ-5D with US scoring is based on the same questionnaire as the UK scoring but uses time tradeoff weights recently collected from a nationally representative sample in the United States.¹⁵ These weights also allow for “states worse than death,” which are given negative values.

Mental Component Summary and Physical Component Summary—Self-Administered

The 12 multiple-choice items of the SF-12 relate to 8 dimensions: physical functioning, physical role limitations, emotional role limitations, pain, general health, vitality, social functioning, and mental health. The SF-12 is an abridged version of the SF-36, which was constructed to reflect the mental and physical component summary scores of the parent scale.^{17,18} The MCS and PCS were developed from a reduction of the 8 dimensions to two dimensions by factor analysis. The factor scores were normalized so that both the mental component summary (MCS) and physical component summary (PCS) have averages of 50 and standard deviations of 10 with respect to the proprietary US national data set held by QualityMetric, Inc.¹⁸ We include imputed scores, calculated using a proprietary algorithm of QualityMetric, Inc.

SF-6D—Self-Administered

The SF-6D scoring algorithm uses 7 of the questions from the SF-12. These questions were used to

construct health scenarios that were evaluated using the standard gamble technique in a representative sample of the UK population. Regression analysis was then used to model the preferences assigned to each health state. With the resulting scoring algorithm, a utility-based score can be assigned to each health state.¹⁹

Visual Analog Scale from the EQ-5D—Self-Administered

The full EuroQol instrument includes the 5 multiple-choice items mentioned above and a visual analog scale. This 20-cm vertical scale runs from the “worst imaginable health state” at 0 to the “best imaginable health state” at 100. A subject places a mark to indicate “how good or bad [his or her] health is today,” which is converted to an integer between 0 and 100.²¹

Quality of Well-being Scale (Estimated)—Interviewer Administered

The QWB Scale categorizes a respondent in mobility, physical activity, social activity, and symptom/problem scales. Preference weights for each function level were derived from 867 raters, and a scoring algorithm was developed to yield scores between 0 and 1.²²

The QWB Scale was not administered in the 2001 MEPS, but a QWB estimation procedure has been developed from 1979 to 1996 NHIS data. Details of the imputation methodology are given elsewhere.^{12,20} QWB-estimates (QWBX1) were estimated using 2001 NHIS data, and the estimate algorithm was modified for the reorganized NHIS. Specifically, the 1997 and later NHIS had questions on functional limitations that more closely match with the QWB Social Activity and Physical Activity subscales. The analysis of these data was weighted to take into account the NHIS sampling design. The 2001 QWB data nonresponse was generally less than 1%. This is because the NHIS, like the interviewer-administered QWB Scale, employs patterns of questions and follow-up probes that allow pursuit of health classification information to near-definitive conclusions in nearly all cases.

Analyses

Data were analyzed using STATA (version 8.2, StataCorp, College Station, TX) to allow adjustment for the complex sampling design of the MEPS or the NHIS. The reported results incorporate the sampling and post-stratification weights, yielding nationally representative estimates for noninstitutionalized adults answering

Table 1 Unweighted Age- and Sex-Stratified Counts of Respondents for Each Health-Related Quality-of-Life Measure

Males, <i>n</i> (%)							
Age Group	Total MEPS	Completed EQ-5D	Completed SF-12	Completed EQ-VAS	Completed Categorical	Total NHIS	QWB Computed
20–29	2087	1663 (80)	1664 (80)	1585 (76)	1662 (80)	2587	2574 (99)
30–39	2221	1912 (86)	1922 (86)	1856 (83)	1918 (86)	3102	3086 (99)
40–49	2276	2002 (88)	2021 (89)	1964 (86)	2021 (89)	3003	2986 (99)
50–59	1772	1545 (87)	1563 (88)	1521 (86)	1565 (88)	2321	2312 (100)
60–69	1070	962 (90)	976 (91)	959 (90)	975 (91)	1434	1423 (99)
70–79	781	704 (90)	715 (92)	702 (90)	716 (92)	1127	1120 (99)
80–89	305	263 (86)	266 (87)	260 (85)	267 (88)	515	515 (100)
Total	10,512	9051(86)	9127 (87)	8847 (84)	9124 (87)	14,089	14,016 (99)
Females, <i>n</i> (%)							
Age Group	Total MEPS	Completed EQ-5D	Completed SF-12	Completed EQ-VAS	Completed Categorical	Total NHIS	QWB Computed
20–29	2225	1881 (85)	1907 (86)	1800 (81)	1907 (86)	3202	3182 (99)
30–39	2456	2177 (89)	2203 (90)	2102 (86)	2209 (94)	3810	3794 (100)
40–49	2550	2277 (89)	2307 (90)	2205 (86)	2300 (90)	3708	3697 (100)
50–59	1910	1712 (90)	1743 (91)	1709 (89)	1744 (91)	2771	2756 (99)
60–69	1261	1145 (91)	1167 (93)	1145 (91)	1171 (93)	2025	2010 (99)
70–79	1005	887 (88)	899 (89)	881 (88)	914 (91)	1794	1782 (99)
80–89	604	499 (83)	500 (83)	491 (81)	519 (86)	1073	1073 (100)
Total	12,011	10,578 (88)	10,726 (89)	10,333 (86)	10,764 (88)	18,383	18,294 (99)

Note: MEPS, Medical Expenditures Panel Survey; EQ-5D, EuroQol-5D; EQ-VAS, EuroQol Visual Analog Scale; NHIS, National Health Interview Survey; QWB, Quality of Well-being Scale.

questionnaires. We use all respondents aged 20 or older. For categorical self-rated health, we report the full distribution of responses, stratified by sex and age by decade. For each continuous scale, we report the estimated mean value and 95% confidence interval around this estimate, stratified by sex and age by decade. We also report the quartile point estimates for each continuous scale, stratified by sex and age by decade.

RESULTS

The total number of respondents in each sample, as well as the number of respondents for each HRQoL measure, is presented by age and sex in Table 1. These counts represent total respondents, not the effective sample size after weighting. In general, instrument completion rates were very high; most age and sex categories had completion rates above 85%. Completion rates do vary with age and sex. For females, the highest rates of completion occur in the 60- to 69-year-old age bracket, with those aged 80 to 89 having the lowest completion rates. For males, the highest completion rates occur in the 70- to 79-year-old age bracket,

with those aged 20 to 29 having the lowest completion rates. The group with the lowest completion rates overall was 20- to 29-year-old males, of whom only 75.9% completed the EQ-VAS. The EQ-VAS was the least completed measure, with 3 groups having response rates under 85%: 20- to 29-year-old females at 80.9%, 30- to 39-year-old males at 83.6%, and 80- to 89-year-old females at 81.3%. The group with the highest completion rates was 60- to 69-year-old females, who had completion rates over 90% for each measure. In general, nonresponders had lower income, were less likely to have finished high school, were less likely to be married, were less likely to be white, and were more likely to be Hispanic than responders (data not shown).

The full distribution of categorical self-rated health responses, stratified by age and sex, is presented in Table 2. These values take the sampling weights into account. From this table, we can see that females are less likely to report "excellent" health than males, except in the oldest age groups. Among both males and females, older age groups are less likely to report "excellent" or "very good" health and more likely to

Table 2 Age- and Sex-Stratified Distribution of Responses for Categorical Self-Rated Health

Age Group	Males					Females				
	E	V	G	F	P	E	V	G	F	P
20–29	.276	.427	.240	.054	.004	.218	.441	.263	.071	.006
30–39	.231	.437	.267	.057	.009	.193	.425	.291	.073	.018
40–49	.193	.390	.307	.091	.019	.145	.407	.308	.113	.028
50–59	.169	.358	.319	.118	.036	.140	.360	.339	.122	.039
60–69	.112	.331	.339	.164	.054	.090	.315	.375	.176	.044
70–79	.060	.255	.391	.224	.071	.065	.237	.397	.248	.055
80–89	.048	.218	.407	.262	.066	.054	.181	.375	.296	.094

Note: E, excellent; V, very good; G, good; F, fair; P, poor. Sampling weights are taken into account.

report “good,” “fair,” or “poor” health than the younger age groups.

The mean estimates and 95% confidence intervals, stratified by age and sex, and incorporating sampling weights for each continuous HRQoL score are presented in Table 3. These values take the sampling weights into account. Respondents report lower scores at higher ages except for the MCS score, which seems to slightly increase for those ages 60 to 79. Females report lower scores than males in all age groups and on all scores. The confidence intervals for these estimates are small because of the large number of respondents in the MEPS survey, although confidence intervals for means widen at older ages, where there were fewer respondents.

The mean estimates of the EQ-5D with UK scoring, EQ-5D with US scoring, EQ-VAS, QWB Scale, and SF-6D are plotted in Figure 1. EQ-VAS scores were rescaled between 0 and 1 for comparison. We did not include the MCS and PCS in this figure because they are not single summary scores of health.

Figure 1 shows that females report lower scores than males and that older respondents report lower scores within each measure. Although the trend with age is the same for each scoring system, the slope of change is not the same, so the rank order of scoring systems within an age range depends on the age range. Although the EQ-5D with US scoring, QWB Scale, and EQ-VAS appear to have similar slopes across the age groups, it appears that the SF-6D changes less with age and the EQ-5D with UK scoring changes more with age.

The point estimates for the quartiles of each continuous HRQoL measure, stratified by age and sex, are presented in Table 4. Quartiles were computed accounting for sampling weights. From this table, we can see the range of values for each summary score as well as the presence of ceiling effects. The most

prominent ceiling effects occur in the EQ-5D based instruments, with both the UK and US scoring, where more than 50% of the youngest age groups are given scores of perfect health (1.0). In general, the UK scoring gives lower values to each quartile than the US scoring.

DISCUSSION

In this report, we have presented nationally representative US values for a variety of HRQoL summary scores. These values were computed from large, recent, nationally representative surveys of the US noninstitutionalized civilian population. Previously, there has not been a publicly published set of nationally representative US values for any HRQoL score using an off-the-shelf instrument except for the SF-36 version 1, an estimated QWB, the EQ-5D with US scoring, and the HUI 2/3. (However, averages were reported for the “HALex” measure derived post hoc from the National Health Interview Survey.²³) To compare our findings to the privately published SF-12 values, we recomputed means with comparable age groups (data not shown) and found our estimates to be within ± 2 of the published PCS means and within ± 3 of the published MCS means.

We also compared our EQ-5D estimates to estimates from the data used in the recent paper by Luo and others¹² (data not shown) and found that the 95% confidence intervals from the estimated means overlapped in all sex and age groups. Use of either set of estimates would be appropriate for US averages, where we report averages of sex by age group; Luo and others report averages of race/ethnicity by other demographic variables. This MEPS sample may be more desirable because of its larger sample size, although details about the mode of administration may also determine the most appropriate data

Table 3 Age- and Sex-Stratified Mean and 95% Confidence Interval (CI) for Each Continuous Health-Related Quality-of-Life Summary Score

Measure	Age Group	Male			Female		
		Lower 95% CI	Mean	Upper 95% CI	Lower 95% CI	Mean	Upper 95% CI
EQ-5D UK	20-29	0.902	0.910	0.919	0.882	0.892	0.902
	30-39	0.888	0.897	0.906	0.854	0.864	0.874
	40-49	0.844	0.854	0.864	0.808	0.820	0.832
	50-59	0.804	0.816	0.829	0.772	0.785	0.797
	60-69	0.768	0.786	0.804	0.731	0.747	0.763
	70-79	0.716	0.736	0.756	0.670	0.689	0.708
	80-89	0.675	0.711	0.747	0.589	0.622	0.656
EQ-5D US	20-29	0.922	0.928	0.934	0.905	0.913	0.920
	30-39	0.912	0.918	0.925	0.886	0.893	0.900
	40-49	0.880	0.887	0.894	0.855	0.863	0.871
	50-59	0.853	0.861	0.870	0.829	0.837	0.846
	60-69	0.827	0.840	0.852	0.800	0.811	0.822
	70-79	0.788	0.802	0.816	0.758	0.771	0.784
	80-89	0.757	0.782	0.807	0.701	0.724	0.747
MCS (SF-12)	20-29	51.7	52.1	52.6	49.0	49.5	50.1
	30-39	51.5	52.0	52.4	49.2	49.7	50.2
	40-49	50.9	51.4	51.9	49.2	49.7	50.1
	50-59	51.7	52.1	52.6	49.9	50.4	50.9
	60-69	52.1	52.7	53.3	51.2	51.8	52.3
	70-79	51.9	52.7	53.6	51.0	51.8	52.5
	80-89	50.2	51.5	52.7	49.4	50.4	51.3
PCS (SF-12)	20-29	53.5	53.8	54.2	52.7	53.0	53.3
	30-39	52.6	53.0	53.3	51.2	51.6	52.1
	40-49	50.5	50.9	51.3	49.0	49.5	50.0
	50-59	48.1	48.6	49.2	46.3	46.8	47.4
	60-69	44.7	45.6	46.4	43.2	44.0	44.7
	70-79	40.1	41.1	42.1	39.2	40.0	40.9
	80-89	37.2	38.7	40.2	34.9	36.0	37.1
QWB Scale ^a	20-29	0.826	0.826	0.826	0.819	0.819	0.819
	30-39	0.831	0.831	0.831	0.820	0.820	0.820
	40-49	0.803	0.803	0.803	0.797	0.797	0.797
	50-59	0.768	0.768	0.768	0.763	0.763	0.763
	60-69	0.737	0.737	0.737	0.738	0.738	0.738
	70-79	0.724	0.724	0.724	0.718	0.718	0.718
	80-89	0.662	0.662	0.662	0.645	0.645	0.645
SF-6D	20-29	0.851	0.857	0.862	0.822	0.827	0.833
	30-39	0.843	0.849	0.855	0.812	0.818	0.824
	40-49	0.825	0.831	0.836	0.798	0.804	0.810
	50-59	0.811	0.819	0.827	0.781	0.788	0.795
	60-69	0.793	0.803	0.813	0.775	0.784	0.794
	70-79	0.758	0.770	0.783	0.736	0.748	0.759
	80-89	0.723	0.742	0.761	0.684	0.700	0.716
EQ-VAS	20-29	86.4	87.2	88.0	83.7	84.5	85.2
	30-39	84.2	84.9	85.6	81.0	81.8	82.5
	40-49	81.2	81.9	82.7	79.4	80.3	81.1
	50-59	78.4	79.5	80.6	77.6	78.5	79.4
	60-69	75.4	76.9	78.3	75.1	76.3	77.5
	70-79	71.6	72.8	74.4	71.2	72.6	74.1
	80-89	67.7	70.2	72.7	64.1	66.1	68.1

Note: EQ-5D UK, EuroQol-5D with UK scoring; EQ-5D US, EuroQol-5D with US scoring; MCS, mental component summary from the SF-12; PCS, physical component summary from the SF-12; QWB, Quality of Well-being Scale; EQ-VAS, Visual Analog Scale from the EuroQol instrument.

a. The upper and lower bounds on the confidence interval (CI) for the QWB estimates differ only in the fourth significant figure.

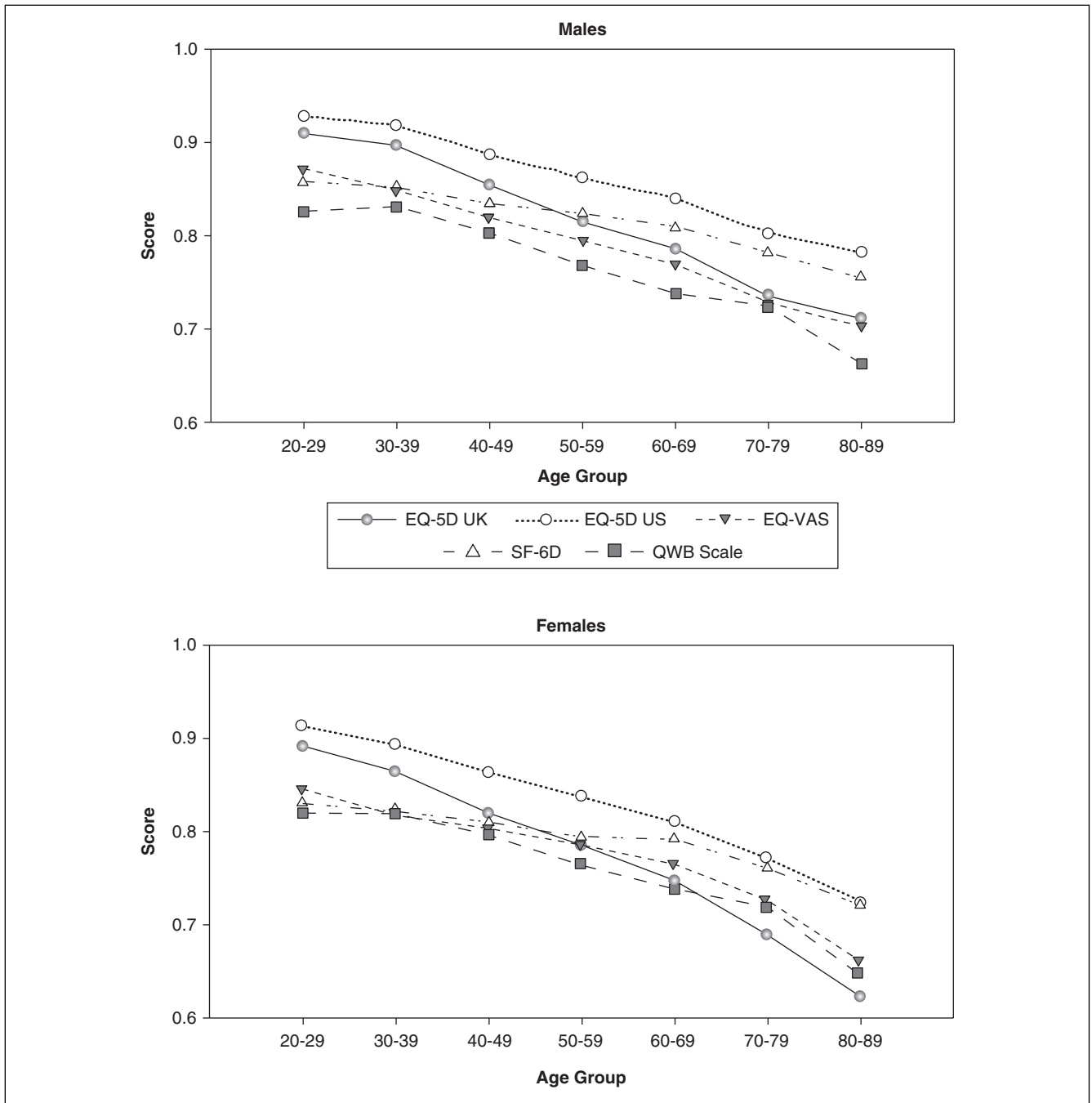


Figure 1 Age- and sex-stratified means for the EQ-5D UK, EQ-5D US, QWB Scale, SF-6D, and EQ-VAS. "EQ-5D UK" is the EuroQol-5D with UK scoring, "EQ-5D US" is the EuroQol-5D with US scoring, QWB is the Quality of Well-being Scale, and "EQ-VAS" is the Visual Analog Scale from the EuroQol instrument.

set to use for future comparisons; the EQ-5D was self-completed on paper for both surveys, the MEPS was a mailed questionnaire, and the questionnaire used by Luo and others was completed while an interviewer was present.

Table 3 shows that, as expected, females report lower mean scores than males and older respondents generally report lower scores within each measure. There are 2 exceptions to the age trend. The first is an increase in estimated QWB scores from the

Table 4 Age- and Sex-Stratified Quartile for Each Continuous Health-Related Quality-of-Life Summary Score

Measure	Age Group	Males					Females				
		Min	25%	50%	75%	Max	Min	25%	50%	75%	Max
EQ-5D UK	20-29	-0.239	0.848	1	1	1	-0.594	0.796	1	1	1
	30-39	-0.484	0.796	1	1	1	-0.536	0.796	1	1	1
	40-49	-0.484	0.796	1	1	1	-0.484	0.725	0.848	1	1
	50-59	-0.239	0.727	0.796	1	1	-0.484	0.725	0.796	1	1
	60-69	-0.594	0.691	0.796	1	1	-0.594	0.691	0.796	1	1
	70-79	-0.594	0.691	0.796	0.883	1	-0.594	0.620	0.725	0.848	1
	80-89	-0.184	0.620	0.760	0.850	1	-0.594	0.516	0.691	0.796	1
	EQ-5D US	20-29	0.122	0.844	1	1	1	-0.109	0.827	1	1
30-39		-0.040	0.827	1	1	1	-0.100	0.827	1	1	1
40-49		-0.038	0.827	1	1	1	-0.040	0.800	0.844	1	1
50-59		0.063	0.810	0.827	1	1	-0.038	0.800	0.827	1	1
60-69		-0.109	0.778	0.827	1	1	-0.109	0.778	0.827	1	1
70-79		-0.109	0.778	0.827	0.860	1	-0.109	0.707	0.800	0.843	1
80-89		0.118	0.708	0.816	0.853	1	-0.109	0.597	0.778	0.827	1
MCS (SF-12)		20-29	13.1	49.1	55.1	57.8	69.2	9.5	45.1	53.0	56.3
	30-39	12.9	49.0	54.6	57.8	70.6	8.6	45.6	53.0	56.7	70.5
	40-49	13.9	47.6	54.3	57.8	68.8	14.7	44.8	53.0	56.8	68.5
	50-59	12.9	49.0	55.1	57.9	67.0	15.5	46.1	53.6	57.4	67.7
	60-69	17.0	49.1	55.7	58.7	68.5	16.7	47.2	55.0	58.5	69.0
	70-79	13.7	48.8	55.8	59.7	71.4	19.1	45.6	54.9	58.9	71.3
	80-89	19.0	45.4	54.6	58.7	67.4	12.8	42.1	52.4	58.9	69.0
	PCS (SF-12)	20-29	21.1	52.9	55.5	56.8	66.4	17.6	51.2	55.3	56.8
30-39		14.2	51.8	55.5	56.7	65.4	12.3	49.7	54.8	56.6	68.0
40-49		11.9	48.3	53.8	56.3	67.1	14.3	45.7	53.2	56.1	66.3
50-59		16.4	45.2	52.4	55.8	65.3	13.2	39.2	51.1	55.5	65.6
60-69		14.5	37.1	49.8	54.8	68.8	14.6	35.1	47.1	53.8	64.5
70-79		12.5	31.9	42.5	51.7	64.5	11.2	30.2	40.2	51.7	63.1
80-89		16.7	29.7	37.9	48.9	60.2	12.3	26.2	34.0	46.4	58.9
QWB Scale		20-29	0.460	0.743	0.814	1	1	0.451	0.741	0.814	1
	30-39	0.422	0.743	0.814	1	1	0.439	0.741	0.814	1	1
	40-49	0.394	0.701	0.796	0.856	1	0.394	0.701	0.770	0.856	1
	50-59	0.378	0.701	0.770	0.830	1	0.378	0.701	0.770	0.856	1
	60-69	0.378	0.660	0.743	0.830	1	0.394	0.660	0.756	0.830	1
	70-79	0.378	0.640	0.718	0.830	1	0.378	0.606	0.735	0.830	1
	80-89	0.378	0.563	0.667	0.743	1	0.378	0.518	0.640	0.743	1
	SF-6D	20-29	0.345	0.800	0.863	0.922	1	0.373	0.793	0.863	0.863
30-39		0.345	0.800	0.863	0.922	1	0.345	0.755	0.863	0.880	1
40-49		0.345	0.797	0.863	0.922	1	0.345	0.737	0.863	0.863	1
50-59		0.345	0.782	0.863	0.922	1	0.345	0.723	0.821	0.863	1
60-69		0.345	0.734	0.863	0.922	1	0.345	0.695	0.817	0.863	1
70-79		0.345	0.660	0.800	0.880	1	0.345	0.618	0.758	0.863	1
80-89		0.366	0.618	0.754	0.863	1	0.357	0.561	0.695	0.859	1
EQ-VAS		20-29	0	80	90	95	100	0	80	89	95
	30-39	0	80	90	95	100	0	75	85	91	100
	40-49	0	75	85	92	100	0	75	85	90	100
	50-59	0	75	85	90	100	0	70	85	90	100
	60-69	0	70	80	90	100	0	66	80	90	100
	70-79	0	60	80	89	100	0	60	78	88	100
	80-89	0	60	75	85	100	0	50	70	80	100

Note: EQ-5D UK, EuroQol-5D with UK scoring; EQ-5D US, EuroQol-5D with US scoring; MCS, mental component summary from the SF-12; PCS, physical component summary from the SF-12; QWB, Quality of Well-being Scale; EQ-VAS, Visual Analog Scale from the EuroQol instrument.

20- to 29-year-old age group to the 30- to 39-year-old age group. This increase is very small (.005 for males and .001 for females) and may be noise from a loss of discrimination in the QWB estimation procedure; before 1997, the NHIS reported 300 diseases/conditions in Recode B, plus another 133 in condition lists. By contrast, the 2001 NHIS reported 61 conditions. The second exception is that MCS scores peak in the 60- to 69-year-old and 70- to 79-year-old age groups. This increase appears to be real as a similar increase is reported in the privately published SF-12 values, with a peak in the 65- to 75-year-old age group.¹³

The values for preference-based scores can be used as a basis for judging incremental effects in cost-effectiveness analyses, whereas the values for the other scores can be used for descriptive comparisons of health status. For analysts intending to use values reported here for cost-effectiveness analyses, it is important to note that none of the measures has a mean of perfect health (1.0) in any stratum; an analysis should not assume that preventing or curing a condition will return a person to perfect health.²⁴ It is also important to note that the values for any age and sex stratum vary by summary score, and mixing absolute values across scoring systems could lead to different conclusions than using the same scoring system consistently for all valuation within 1 analysis. For instance, comparing the mean EQ-5D US score for a 60- to 69-year-old to the mean SF-6D score for a 40- to 49-year-old would lead—erroneously—to the conclusion that those aged 60 to 69 report better health than those aged 40 to 49. If these age brackets were compared using mean scores within either scoring system, the opposite—and correct—conclusion would be obtained.

The confidence intervals around the mean estimates by age and sex are small in this report because of the large number of respondents and generally high completion rates for each measure. In the MEPS, the measures in order from highest to lowest completion rates were the SF-12, EQ-5D, and EQ-VAS, which is the same order that these measures were placed in the questionnaire. Other recent, large population surveys administered by mail in the United States²⁵ and Canada,²⁶ which included the EQ-5D, EQ-VAS, and the SF-12, found lower response rates for the SF-12. The SF-12 appeared after the EQ-5D and EQ-VAS in the US questionnaire, suggesting that the differing completion rates may be dependent on placement within the questionnaire and not characteristics of the indexes. Unfortunately, the variability in completion rates in different measures means that the set of respondents

within each measure is different. Although 82% of respondents completed all measures in the MEPS, the reported values may be subject to a measure-by-measure self-selection bias.

A strength of using the MEPS data for this analysis is that the survey included 2 intact HRQoL instruments. From these 2 instruments, we directly determined 6 of the summary scores presented in this report. The seventh summary score, the QWB Scale, was estimated using a procedure developed from questions in the contemporaneous NHIS data.²⁰ It would have been preferable to have access to a nationally representative sample that directly answered the questions within the QWB Scale and other commonly used measures.

Although we have reported nationally representative means and quartiles for 3 of the most commonly used HRQoL measures, a variety of other off-the-shelf measures were not included. The most notable exclusion is the Health Utilities Index, although nationally representative mean estimates of this measure are publicly available.¹² If the MEPS continues to include HRQoL measures, this report should be updated periodically for convenient and timely use by analysts.

In conclusion, we have presented the first publicly available set of nationally representative US values for any standardized HRQoL measure, except those for the SF-36 version 1,¹⁰ the estimated QWB,¹¹ the EQ-5D with US scoring,¹² and the HUI 2/3.¹² We have presented these values for 7 different summary scores; 6 scores come directly from intact measures, and 1 was estimated. We believe nationally representative values are important for both generalized comparisons of health status and cost-effectiveness analyses.

REFERENCES

1. Gold MR. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996.
2. Australian Bureau of Statistics. *National Health Survey Australia, 1995: SF-36 Population Norms* (Catalogue no 4399.0). Canberra: Australian Bureau of Statistics; 1997.
3. Blake C, Codd MB, O'Meara YM. The Short Form 36 (SF-36) Health Survey: normative data for the Irish population. *Irish J Med Sci*. 2000;169:195–200.
4. Scott KM, Tobias MI, Sarfati D, Haslett SJ. SF-36 health survey reliability, validity and norms for New Zealand. *Aust NZ J Publ Heal*. 1999;23:401–6.
5. Loge JH, Kaasa S. Short Form 36 (SF-36) Health Survey: normative data from the general Norwegian population. *Scand J Soc Med*. 1998;26:250–8.
6. Chan SP, Machin D, Soh CH, et al. Measuring health-related quality of life in Singapore: normal values for the English and Chinese SF-36 Health Survey. *Ann Acad Med Singapore*. 2002;31:366–74.

7. Alonso J, Regidor E, Barrio G, Prieto L, Rodríguez C, de la Fuente L. Valores poblacionales de referencia de la versión española del cuestionario de salud SF-36. *Med Clin (Barc)*. 1998;111:410-6.
8. Ruiz M, Rejas J, Soto J, Pardo A, Rebollo I. Adaptation and validation of the Health Utilities Index Mark 3 into Spanish and correction norms for Spanish population. *Med Clin (Barc)*. 2003;120:89-96.
9. Bowling A, Bond M, Jenkinson C, Lamping DL. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the omnibus survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *J Public Health Med*. 1999;21:255-70.
10. McHorney CA, Kosinski M, Ware JE. Comparison of the costs and quality of norms for the SF-36 Health Survey collected by mail versus telephone interview: results from a national survey. *Med Care*. 1994;32:551-67.
11. Anderson JP. Activity limitations reported by the National Health Interview Survey: an anomaly and its effect on QWBX1 estimates of national well-being [research letter]. *Am J Public Health*. 2001;91:1135-6.
12. Luo N, Johnson JA, Shaw JW, Feeny D, Coons SJ. Self-reported health status of the general adult US population as assessed by the EQ-5D and Health Utilities Index. *Med Care*. 2005;43:1078-86.
13. Ware JE, Kosinski M, Keller SD. How to Score the SF-12[®] Physical and Mental Health Summary Scales. Lincoln, RI: Quality-Metric Incorporated; 1998.
14. Fryback DG, Dasbach EJ, Klein R, et al. The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors. *Med Decis Making*. 1993;13:89-102.
15. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005;43:203-20.
16. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35:1095-1108.
17. Gandek B, Ware JE, Aaronson NK, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51:1171-8.
18. Ware JE Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220-33.
19. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care*. 2004;42:851-9.
20. Anderson JP, Kaplan RM, Ake CF. Arthritis impact on US life quality: morbidity and mortality effects from National Health Interview Survey data 1986-1988 and 1994 using QWBX1 estimates of well-being. *Soc Indic Res*. 2004;69:67-91.
21. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33:337-43.
22. Kaplan RM, Bush JW, Berry CC. Health status: types of validity and the Index of Well-Being. *Health Serv Res*. 1976;11:478-507.
23. Erickson P, Wilson R, Shannon I. Years of Healthy Life. Statistical Notes No. 7. Hyattsville, MD: US Department of Health and Human Services, CDC, National Center for Health Statistics; 1995.
24. Fryback DG, Lawrence WF. Dollars may not buy as many QALYs as we think: a problem with defining quality of life adjustments. *Med Decis Making*. 1997;17:276-84.
25. Johnson JA, Coons SJ. Comparison of the EQ-5D and SF-12 in and adult US sample. *Qual Life Res*. 1998;7:155-66.
26. Johnson JA, Pickard AS. Comparison of the EQ-5D and SF-12 health surveys in a general population survey in Alberta, Canada. *Med Care*. 2000;38:115-21.