# Gender differences in health-related quality-of-life are partly explained by sociodemographic and socioeconomic variation between adult men and women in the US: evidence from four US nationally representative data sets

**Dasha Cherepanov · Mari Palta · Dennis G. Fryback · Stephanie A. Robert**

## Abstract

*Purpose* The purpose of this study was to describe gender differences in self-reported health-related quality-of-life (HRQoL) and to examine whether differences are explained by sociodemographic and socioeconomic status (SES) differentials between men and women.

*Methods* Data were from four US nationally representative surveys: US Valuation of the EuroQol EQ-5D Health States Survey (USVEQ), Medical Expenditure Panel Survey (MEPS), National Health Measurement Study (NHMS) and Joint Canada/US Survey of Health (JCUSH). Gender differences were estimated with and without adjustment for sociodemographic and SES indicators using regression within and across data sets with SF-6D, EQ-5D, HUI2, HUI3 and QWB-SA scores as outcomes.

*Results* Women have lower HRQoL scores than men on all indexes prior to adjustment. Adjusting for age, race, marital status, education and income reduced but did not remove the gender differences, except with HUI3. Adjusting for marital status or income had the largest impact on estimated gender differences.

*Conclusions* There are clear gender differences in HRQoL in the United States. These differences are partly explained by sociodemographic and SES differentials.

**Keywords** Sex differences · Men's health · Women's health · Quality of life · Outcome assessment · Health status · Gender differences · SF-6D · EQ-5D · HUI2 · HUI3 · QWB-SA · Population study

## Introduction

There are well-documented gender differences in health and health behaviors as reflected in various measures of morbidity, mortality and health care utilization [1–3]. In 2002, age-adjusted death rates for 8 out of 10 leading causes of death in the United States were greater for men than for women, and men were more likely to suffer from chronic conditions that are more severe [1, 2, 4]. Men have more life-threatening chronic diseases, including coronary heart disease, cancer, cerebrovascular disease, emphysema, cirrhosis of the liver, kidney disease and atherosclerosis, while women have higher rates of chronic disabling disorders [3]. Women are more likely to suffer from autoimmune and rheumatologic disorders and from other non-life-threatening diseases, such as anemia, thyroid conditions, gallbladder conditions, migraines, arthritis and eczema [3]. Women are also more likely to experience depressive and anxiety disorders while men more commonly experience antisocial behavior, substance abuse and suicide [1–3]. Differences in morbidity and mortality between men and women create a complex relationship between gender and health.

Preference scored, self-reported health-related quality-of-life (HRQoL) indexes are widely used to summarize the

D. Cherepanov (✉)
Department of Health Services, University of California
Los Angeles School of Public Health, P.O. Box 90095-1772,
Los Angeles, CA, USA
e-mail: dcherepanov@ucla.edu

D. Cherepanov
The RAND Corporation, 1776 Main Street. M5S,
Santa Monica, CA 90407-2138, USA

M. Palta · D. G. Fryback · S. A. Robert
Department of Population Health Sciences, University of
Wisconsin School of Medicine and Public Health, 707 WARF
Building, 610 North Walnut Street, Madison, WI 53726, USA

overall health of individuals or groups and to estimate quality-adjusted life-years for use in cost-effectiveness analyses [5, 6]. Although men may seem less healthy than women based on their lower life expectancy at all ages and a greater likelihood of suffering from life-threatening diseases, several studies have found women to have lower HRQoL scores than men [5, 7–9]. However, the evidence to date is not clear, because population-based analyses of HRQoL more often control for potential confounding by gender than study gender as an independent variable. No study has systematically examined the relationship of gender and HRQoL across surveys and measures.

Research on HRQoL and more simple measures of self-reported health status demonstrate that self-rated health differs across gender and other sociodemographic and socioeconomic status (SES) characteristics, such as race, marital status, education and income [6, 7, 10–13]. Liu and Umberson (2008) found self-reported health of those widowed, divorced and separated to be poor relative to those who are married, especially among women [12]. The percentage of women who are married decreases with age and is lower than that of men after age 45 [15]. Moreover, HRQoL varies significantly by SES [13], and it is well known that women have lower individual income than men [14]. Hence, gender differences in HRQoL may be partly due to differences in sociodemographic and SES factors.

Several publicly available, nationally representative surveys of HRQoL present an opportunity to examine the consistency of gender difference in HRQoL. Each survey utilized at least one of five commonly used HRQoL indexes, the Short Form 6 dimension (SF-6D) [16, 17], EuroQol 5 dimension (EQ-5D) [18], the Health Utilities Index Mark 2 (HUI2) [20] and Mark 3 (HUI3) [21] and the Quality of Well-Being Scale Self-Administered form (QWB-SA) [22, 23], and each index, except for the QWB-SA, was used by at least two different surveys. Each survey also collected information on respondent sociodemographic and SES characteristics.

In this study, parallel analyses within each dataset are performed, as well as analyses that pool the surveys. The following hypotheses are tested: (1) women report lower HRQoL than men on the five HRQoL indexes, and (2) differences in sociodemographic and SES characteristics between men and women explain gender differences in HRQoL.

## Methods

### Measures

The five health indexes, SF-6D, EQ-5D, HUI2, HUI3 and QWB-SA, are preference-based generic HRQoL measures

anchored on a cardinal scale by 0 (dead) and 1 (full health). The EQ-5D, HUI2 and HUI3 allow health states with utilities less than zero (worse than dead). Although the five indexes are assumed to be measuring the same concept of health on the same theoretical scale, they in fact differ in construction, in coverage of health domains and in their numerical ranges.

The SF-6D index was developed as a summary utility scale based on items from the SF-36v2™ and SF-12v2™ (a subset of SF-36v2™) questionnaires [16, 17]. It refers to health in the "past 4 weeks" and covers six health domains: physical function, role limitation, social function, pain, mental health and vitality. The SF-6D index produces single summary utility scores ranging from 0.35 to 1.0. In this study, the SF-12v2™ version of SF-6D is used.

The EuroQoL EQ-5D refers to health "today" and incorporates five domains of health, mobility, self-care, usual activities, pain/discomfort and anxiety/depression. This index produces summary utility scores that range from −0.11 to 1 [18]. US weights are used in this study [19].

The HUI2 and HUI3 indexes refer to health "in the past week." HUI2 has six domains (sensation, mobility, emotion, cognition, self-care, pain), and HUI3 has eight domains (vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain) for which data are collected using the proprietary Health Utilities Index questionnaire [20, 21]. Summary utility scores range from −0.03 to 1.0 for HUI2 and from −0.36 to 1.0 for HUI3.

The QWB-SA index refers to the past three days and covers four domains: mobility/self-care, physical activity, self-care/usual activity and acute/chronic symptoms [22, 23]. The QWB-SA produces a summary utility score ranging from 0.09 to 1.0.

### Data and variables

These analyses use four publicly available data sets that were collected during similar time frames and contain HRQoL measures of interest as described in Table 1. All four data sets contain survey weights for producing estimates generalizable to the non-institutionalized adult population in the United States. All participants gave written informed consent at the inception of each of the four surveys. The use of these publically available data sets for this study was approved by the University of Wisconsin-Madison Health Sciences Institutional Review Board.

The US Valuation of the EuroQoL EQ-5D Health States Survey (USVEQ) produced a nationally representative sample from civilian non-institutionalized residents of ages 18 and older within the 50 US states and the District of Columbia [19]. Data were collected between June and October of 2002. It was administered via paper-and-pencil in face-to-face interviews (PAPI). Survey weights were

**Table 1** Descriptions of four US representative surveys

|  | USVEQ | MEPS | NHMS | JCUSH |
|---|---|---|---|---|
| Year of administration | June–October 2002 | Full year of 2003 | June 2005–August 2006 | November 2002–March 2003 |
| Form of administration | PAPI | SAQ | CATI | CATI |
| Age of respondents | 18+ years old | 18–90 years old | 35–89 years old | 18+ years old |
| No. of US respondents (proxy, non-proxy) | 4,048 | 22,684 | 3,844 | 5,183 |
| No. of US respondents[a] (non-proxy, ages 35–89, Blacks + Whites only) | 2,471 (includes 719 non-Hispanic Blacks, 1,752 Hispanics and Whites) | 13,195 (includes 38 Hispanic Blacks, 2,582 Hispanic Whites) | 3,648 (includes 23 Hispanic Blacks, 47 Hispanic Whites) | 3,186 (includes 3 Hispanic Blacks, 59 Hispanic Whites) |
| Response rate (proxy, non-proxy) | 59% | 65% | 56% | 50%[b] |
| HRQoL measures |  |  |  |  |
| SF6D-SF36v2 |  |  | X |  |
| SF6D-SF12v2 |  | X | X* |  |
| QWB-SA |  |  | X |  |
| EQ-5D | X | X | X |  |
| HUI2 | X |  | X |  |
| HUI3 | X |  | X | X |

*USVEQ* US Valuation of the EuroQol EQ-5D health states survey; *MEPS* Medical expenditure panel survey; *NHMS* National health measurement study; *JCUSH* Joint Canada/US survey of health (JCUSH); *PAPI* Paper-and-pencil in face-to-face interview; *SAQ* Self-administered questionnaire; *CATI* Computer-assisted telephone interview

\* The SF-12v2$^{TM}$ is a subset of the SF-36v2$^{TM}$, and thus the SF6D-SF12v2 can be scored in NHMS

[a] The number of US respondents represents the analytical samples used from each of the four surveys

[b] This is the United States response rate (excluding Canada)

post-stratified by age (18+), gender and race (Hispanic, non-Hispanic black and other).

The Medical Expenditure Panel Survey (MEPS) has been administered annually to the US non-institutionalized population ages 18–90 years old (for confidentiality purposes, MEPS coded all ages >85 as '85') since 1996 to obtain information on health care utilization and expenditures via mailed self-administered questionnaire (SAQ) [25]. Survey weights were post-stratified by six variables (race/ethnicity: Hispanic, black non-Hispanic, and other; sex; age; poverty status; census region; metropolitan statistical area). The 2003 MEPS data are used in the current study to best match the time frame of the other nationally representative data sets.

The National Health Measurement Study (NHMS) is a nationally representative computer-assisted telephone interview (CATI) survey of non-institutionalized adults between the ages of 35 and 89 residing in the continental US [6]. The survey was conducted between June 2005 and August 2006. The survey weights were post-stratified by gender, race (black, white and other) and age (35–44, 45–66, 65+).

The Joint Canada/US Survey of Health (JCUSH) is a cross-sectional random-digit-dialed telephone survey conducted in both Canada and the United States, administered via a CATI [24]. Data collection took place between November 2002 and March 2003. The study included

people ages 18 and over in both countries and excluded people who were institutionalized or living in United States or Canadian territories. JCUSH survey weights were post-stratified by age (18–44, 44–64, 65+) and gender.

The current analyses target non-proxy US respondents whose ages were between 35 and 89 years and who reported their race/ethnicity as either white/Caucasian or black/African American (Table 1). The 'other' race/ethnicity category was excluded from our analyses since these data contained too few individuals from each of these many racial and ethnic subgroups of the US population. NHMS, JCUSH and MEPS specified their race categories as white/Caucasian and black/African American while race categories in USVEQ were black/African American (non-Hispanic), white (non-Hispanic) and Hispanic/Latino. For comparability with the other surveys, we coded Hispanic/Latino from USVEQ as white/Caucasian. The majority (at least 67%) of Hispanic respondents in the other three studies self-reported themselves as white.

Our analyses adjusted for sociodemographic and SES variables. The variables in the different data sets were coded as similarly as possible. Table 2 describes the coding and shows the distributions of the resulting independent variables by dataset and gender.

The datasets differed only slightly in coding of the income variable. All studies but MEPS measured total household income, whereas MEPS measured income as

family income in terms of percent of the poverty line based on number of family members. According to Department of Health and Human Services Poverty Guidelines, income as percent of the Federal Poverty Line (FPL), is classified as poor: <100%, near poor: 100–125%, low: 125–200%, middle: 200–400% and high: 400% or greater. FPL varies according to number of members in a family [26]. The number of family members and midpoints of FPL percentage ranges were used to convert the MEPS income classifications into dollar values. These were grouped into four income categories comparable to those of the other three studies.

The categorization of age was the same in all four data sets. Marital status and education questions had variations in phrasing and answer options. Categories were created to be relatively similar across surveys.

## Analyses

We fit survey-weighted least squares (WLS) regression models separately for each data set with HRQoL scores as outcome variables. Gender, age and race (model 1) were included in all models; and marital status (model 2), education (model 3) and income (model 4), were added one at

**Table 2** Weighted proportions (shown in percents) of sample characteristics by dataset and gender

| Data (year) | USVEQ (2002) | | | MEPS (2003) | | | NHMS (2005–2006) | | | JCUSH (2002–2003) | | | Census 2000 Proportions |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender | Women | Men | Total | Women | Men | Total | Women | Men | Total | Women | Men | Total | |
| Gender (%) | 52.3 | 47.7 | | 52.5 | 47.5 | | 53.5 | 46.5 | | 54.0 | 46.0 | | |
| Age group | | | | | | | | | | | | | |
| 35–44 | 33.3 | 31.8 | 32.6 | 28.0 | 30.5 | 29.2 | 32.6 | 28.7 | 30.8 | 29.5 | 31.5 | 30.4 | – |
| 45–54 | 25.6 | 28.5 | 27.0 | 26.5 | 27.9 | 27.1 | 24.8 | 22.4 | 23.7 | 26.6 | 29.0 | 27.7 | – |
| 55–64 | 17.2 | 19.1 | 18.1 | 19.4 | 20.1 | 19.7 | 20.0 | 21.1 | 20.5 | 18.6 | 18.8 | 18.7 | – |
| 65–74 | 12.3 | 12.4 | 12.4 | 13.4 | 12.7 | 13.0 | 13.2 | 15.8 | 14.4 | 15.2 | 11.6 | 13.5 | – |
| 74–89 | 11.5 | 8.1 | 9.9 | 12.7 | 8.9 | 10.9 | 9.4 | 12.0 | 10.6 | 10.2 | 9.0 | 9.6 | – |
| Total N (unweighted) | 2,471 | | | 13,195 | | | 3,648 | | | 3,186 | | | |
| Race[a] | | | | | | | | | | | | | |
| Black | 11.8 | 9.7 | 10.8 | 11.9 | 10.5 | 11.2 | 12.5 | 10.3 | 11.4 | 14.5 | 12.3 | 13.5 | |
| White | 88.2 | 90.3 | 89.2 | 88.1 | 89.5 | 88.8 | 87.5 | 89.7 | 88.6 | 85.5 | 87.7 | 86.5 | |
| Marital status | | | | | | | | | | | | | |
| Never married/single | 5.5 | 8.8 | 7.1 | 9.0 | 11.6 | 10.3 | 7.3 | 6.5 | 6.9 | 8.0 | 8.6 | 8.3 | 9.6* |
| Widowed/divorced/separated | 35.4 | 16.8 | 26.5 | 32.2 | 18.5 | 25.7 | 23.9 | 12.8 | 18.8 | 29.2 | 13.4 | 21.9 | 26.1* |
| Married/living with partner | 59.1 | 74.4 | 66.4 | 58.8 | 69.9 | 64.1 | 68.7 | 80.7 | 74.3 | 62.8 | 78.0 | 69.8 | 64.3* |
| Education | | | | | | | | | | | | | |
| Less than high school (HS) | 18.4 | 18.7 | 18.6 | 18.3 | 18.9 | 18.6 | 8.8 | 6.8 | 7.9 | 9.8 | 7.4 | 8.7 | 17.8** |
| HS graduate/equivalent | 34.8 | 32.0 | 33.5 | 35.8 | 32.1 | 34.0 | 28.9 | 28.7 | 28.8 | 38.7 | 32.2 | 35.7 | 29.6** |
| Some post-HS education | 24.9 | 25.4 | 25.2 | 22.5 | 20.0 | 21.3 | 24.7 | 19.4 | 22.2 | 15.3 | 14.6 | 15.0 | 28.0** |
| College degree or higher | 21.8 | 23.9 | 22.8 | 23.4 | 29.0 | 26.1 | 37.6 | 45.1 | 41.1 | 36.2 | 45.9 | 40.7 | 24.6** |
| Household income[b] | | | | | | | | | | | | | |
| Lowest income level | 28.6 | 19.5 | 24.1 | 9.2 | 7.0 | 8.2 | 12.7 | 9.2 | 11.0 | 16.9 | 10.1 | 13.7 | 21.3*** |
| 2nd lowest income level | 28.9 | 27.3 | 28.1 | 14.1 | 10.1 | 12.2 | 16.1 | 14.0 | 15.1 | 22.3 | 16.4 | 19.5 | 18.2*** |
| 3rd lowest income level | 28.6 | 32.4 | 30.5 | 55.0 | 58.5 | 56.7 | 34.3 | 39.1 | 36.6 | 35.1 | 40.4 | 37.6 | 35.4*** |
| Highest income level | 13.9 | 20.8 | 17.3 | 21.7 | 24.4 | 23.0 | 36.9 | 37.7 | 37.3 | 25.6 | 33.1 | 29.2 | 25.1*** |

In regression analyses, marital status, education, and household income were adjusted to Census 2000 proportions (www.factfinder.census.gov) that include

\* All races, ages 35+ (never married; separated/widowed/divorced; married, spouse present)

\*\* Whites and blacks only, ages 25+ (less than 9th grade/9–10th grade, no diploma; high school graduate (includes equivalency); some college, no degree/Associate degree; Bachelor's degree/graduate or professional degree)

\*\*\* Whites and blacks only, ages 35+ (<$20,000; $20,000–$34,999; $35,000–$74,999; ≥$75,000); All Census 2000 proportions exclude Alaska, Hawaii and Puerto Rico

[a] In regression analyses, race was adjusted to average proportions across datasets (Black 12%; White 88%); [b] Income levels by dataset: USVEQ (<$20,000, $20,000–$40,000, $40,001–$75,000, ≥$75,001), MEPS family income (<$20,000, $20,000–$34,999, $35,000–$74,999, ≥$75,000), NHMS (<$20,000, $20,000–$34,999, $35,000–$74,999, ≥$75,000), JCUSH (<$20,000, $20,000–$40,000, $40,001–$80,000, ≥$80,001)

a time and simultaneously (model 5). All covariates were modeled as indicator variables. Gender differences used women as the reference category by coding men as "1" and women as "0".

Estimates of gender difference were standardized via model-based adjustment to the marginal Census 2000 proportions (www.factfinder.census.gov) of variables not already used for post-stratification as shown in Table 2. This was achieved by including interactions of gender with these sociodemographic and SES variables centered at their percentage representation in the Census.

Additional analyses were conducted by pooling data sets: MEPS and NHMS (for SF-6D); USVEQ, MEPS and NHMS (for EQ-5D); JCUSH, NHMS and USVEQ (for HUI3); and USVEQ and NHMS (for HUI2). In the WLS analyses of the pooled data, each dataset was specified as a sampling stratum, and indicator variables for the particular data sets were included in the models.

All analyses were conducted using SAS/STAT® System for Windows (version 9.1) applying procedures PROC SURVEYFREQ, SURVEYMEANS and SURVEYREG, which incorporated survey weights to produce US nationally representative estimates (Copyright 2002–2003 SAS Institute Inc., Cary, NC, USA).

## Results

### Respondent characteristics

Table 2 shows the weighted sample characteristics underlying each dataset separately for women and men. Based on all four surveys, the distributions of marital status and income were different for men and women, yet education distributions were similar. Notably, a greater proportion of women than men reported being widowed, divorced or separated while more men reported being married or living with a partner. All four surveys showed a higher proportion of women reporting the two lowest levels of income while a greater proportion of men reported having the two highest levels of income.

### Results from regression analyses

Gender differences from multivariable WLS regressions from the four studies and pooled analyses are presented in Table 3 and are shown graphically in Fig. 1. The table summarizes gender parameter estimates from five WLS regression models with women as the reference category. All estimates, except for fully adjusted HUI3, were positive, indicating men on average have better HRQoL than women. Gender differences in HRQoL were statistically significant ($P \leq .05$) for most HRQoL measures when only age and

race (model 1) were in the model with the exception of EQ-5D and HUI3 in USVEQ ($P \leq .08$). Adjustment for marital status (model 2) or income (model 4) tended to result in the greatest reductions in the gender difference. The magnitude of gender differences for SF-6D in MEPS was larger than in NHMS. Adjustment for income greatly reduced the gender difference for SF-6D in NHMS, and statistical significance of gender was lost in the models that included income (models 4 and 5). Gender differences in EQ-5D from MEPS and NHMS were nearly identical across models until income was added into the model. In both NHMS and USVEQ, gender differences in EQ-5D decreased when adjusted for marital status or income. The magnitudes of gender differences in HUI2 and HUI3 were smaller in NHMS than in USVEQ when adjusted for education and all three sociodemographic and SES variables at once (marital status, education, income). The gender differences were most prominent for HUI2 in USVEQ, SF-6D in MEPS, and QWB-SA in NHMS. EQ-5D showed the smallest gender differences. HUI3 was the only measure that had a negative but statistically non-significant gender estimate after all adjustments in both JCUSH and NHMS. Figure 1 illustrates the results seen in Table 3 for each HRQoL measure across the surveys where it was obtained.

Pooled analyses are shown in Table 3 and are plotted in Fig. 1. These analyses confirmed that adjusting for marital status or income had the greatest impact on the gender difference in HRQoL. Positive gender parameter estimates were statistically significant in all models for SF-6D and EQ-5D, although these results were mostly driven by the MEPS. The HUI2 and HUI3 had statistically significant gender estimates up to model 3 (which included education). In model 2 (including marital status), gender differences in HUI3 had a $P \leq .08$. Addition of income or all covariates at once in the model decreased precision of the gender estimates, and generally made the statistically significant difference disappear. HUI3 was the only measure that showed no gender differences after all adjustments in the pooled analysis. Only SF-6D and QWB-SA retained gender difference estimate of at least 0.02.

## Discussion

This study confirms that women self-report worse health than men on five commonly used HRQoL indexes. This difference in HRQoL appears to be explained in large part by differences in sociodemographic and SES characteristics between men and women in the US population. Large differences in HRQoL by SES have been previously documented [13]. The lower average income of women appears to account for much of their disadvantage in HRQoL.

**Table 3** HRQoL measures regressed (WLS) on gender (controlling for age, race, marital status, education and income) with women as the reference category

| Gender parameter estimates | SF-6D | | EQ-5D | | HUI2 | | HUI3 | | QWB-SA | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Estimate | SE | Estimate | SE | Estimate | SE | Estimate | SE | Estimate | SE |
| *USVEQ dataset* | | | | | | | | | | |
| Model: (adjusting for) | | | | | | | | | | |
| 1: age, race | – | – | **0.016**$^{\ddagger}$ | 0.009 | **0.024**\*\* | 0.008 | **0.023**$^{\ddagger}$ | 0.013 | – | – |
| 2: age, race, marital status | – | – | 0.009 | 0.009 | 0.014 | 0.008 | 0.010 | 0.013 | – | – |
| 3: age, race, education | – | – | **0.016**\* | 0.009 | **0.024**\*\* | 0.008 | **0.023**$^{\ddagger}$ | 0.013 | – | – |
| 4: age, race, income | – | – | 0.008 | 0.009 | 0.013 | 0.008 | 0.012 | 0.013 | – | – |
| 5: all covariates | – | – | 0.008 | 0.009 | 0.012 | 0.009 | 0.011 | 0.013 | – | – |
| *MEPS dataset* | | | | | | | | | | |
| Model: (adjusting for) | | | | | | | | | | |
| 1: age, race | **0.029**\*\*\* | 0.002 | **0.017**\*\*\* | 0.003 | – | – | – | – | – | – |
| 2: age, race, marital status | **0.025**\*\*\* | 0.003 | **0.013**\*\*\* | 0.003 | – | – | – | – | – | – |
| 3: age, race, education | **0.028**\*\*\* | 0.003 | **0.016**\*\*\* | 0.003 | – | – | – | – | – | – |
| 4: age, race, income | **0.027**\*\*\* | 0.003 | **0.017**\*\*\* | 0.004 | – | – | – | – | – | – |
| 5: all covariates | **0.022**\*\*\* | 0.003 | **0.013**\*\*\* | 0.004 | – | – | – | – | – | – |
| *NHMS dataset* | | | | | | | | | | |
| Model: (adjusting for) | | | | | | | | | | |
| 1: age, race | **0.023**\*\*\* | 0.007 | **0.019**\*\* | 0.007 | **0.024**\*\* | 0.008 | **0.023**\* | 0.012 | **0.030**\*\*\* | 0.008 |
| 2: age, race, marital status | **0.015**\* | 0.007 | **0.014**$^{\ddagger}$ | 0.008 | 0.015 | 0.009 | 0.009 | 0.013 | **0.024**\*\* | 0.008 |
| 3: age, race, education | **0.020**\* | 0.008 | 0.013 | 0.009 | 0.018 | 0.011 | 0.014 | 0.016 | **0.027**\*\* | 0.009 |
| 4: age, race, income | 0.012 | 0.008 | 0.010 | 0.010 | 0.013 | 0.012 | 0.011 | 0.017 | **0.023**\*\* | 0.009 |
| 5: all covariates | 0.007 | 0.009 | 0.005 | 0.011 | 0.006 | 0.013 | −0.002 | 0.018 | **0.021**\* | 0.010 |
| *JCUSH dataset* | | | | | | | | | | |
| Model: (adjusting for) | | | | | | | | | | |
| 1: age, race | – | – | – | – | – | – | **0.024**\*\* | 0.008 | – | – |
| 2: age, race, marital status | – | – | – | – | – | – | 0.013 | 0.009 | – | – |
| 3: age, race, education | – | – | – | – | – | – | 0.008 | 0.012 | – | – |
| 4: age, race, income | – | – | – | – | – | – | 0.004 | 0.011 | – | – |
| 5: all covariates | – | – | – | – | – | – | −0.008 | 0.014 | – | – |
| *Pooled datasets* | | | | | | | | | | |
| Model: (adjusting for) | | | | | | | | | | |
| 1: age, race | **0.028**\*\*\* | 0.002 | **0.018**\*\*\* | 0.003 | **0.024**\*\*\* | 0.006 | **0.024**\*\*\* | 0.007 | – | – |
| 2: age, race, marital status | **0.023**\*\*\* | 0.003 | **0.014**\*\*\* | 0.003 | **0.017**\*\* | 0.007 | **0.012**$^{\ddagger}$ | 0.007 | – | – |
| 3: age, race, education | **0.027**\*\*\* | 0.003 | **0.016**\*\*\* | 0.003 | **0.023**\*\* | 0.007 | **0.017**\* | 0.008 | – | – |
| 4: age, race, income | **0.024**\*\*\* | 0.003 | **0.013**\*\*\* | 0.003 | **0.013**$^{\ddagger}$ | 0.007 | 0.007 | 0.008 | – | – |
| 5: all covariates | **0.020**\*\*\* | 0.003 | **0.010**\*\*\* | 0.004 | 0.010 | 0.008 | 0.000 | 0.009 | – | – |

The table summarizes regression parameter estimates for gender

\*\*\* $P \leq .001$; \*\* $P \leq .01$; \* $P \leq .05$; $^{\ddagger}$ $P \leq .08$

The gender-associated differences seen in HRQoL depend somewhat on the particular HRQoL index used. The estimates range from 0.02 to 0.03 (higher HRQoL for men) without adjustment and from −0.01 to 0.03 with adjustment for marital status, education and income. Only the SF-6D and QWB-SA displayed magnitudes of gender differences that approached 0.03, considered substantively important for preference-based HRQoL indexes [10, 27, 28, 34].

The gender differences on all HRQoL indexes followed parallel trajectories of change with adjustment for sociodemographic and SES variables. Taking age and race differences into account generally did not explain the gender difference in HRQoL and produced estimates that were quite similar across measures (except somewhat higher for SF-6D and lower for EQ-5D). Once further adjustments for marital status, education and income were taken into account, gender differences became small on most HRQoL measures.

**Fig. 1** The figure shows estimated gender differences in HRQoL measures by WLS regression model with women as the reference category
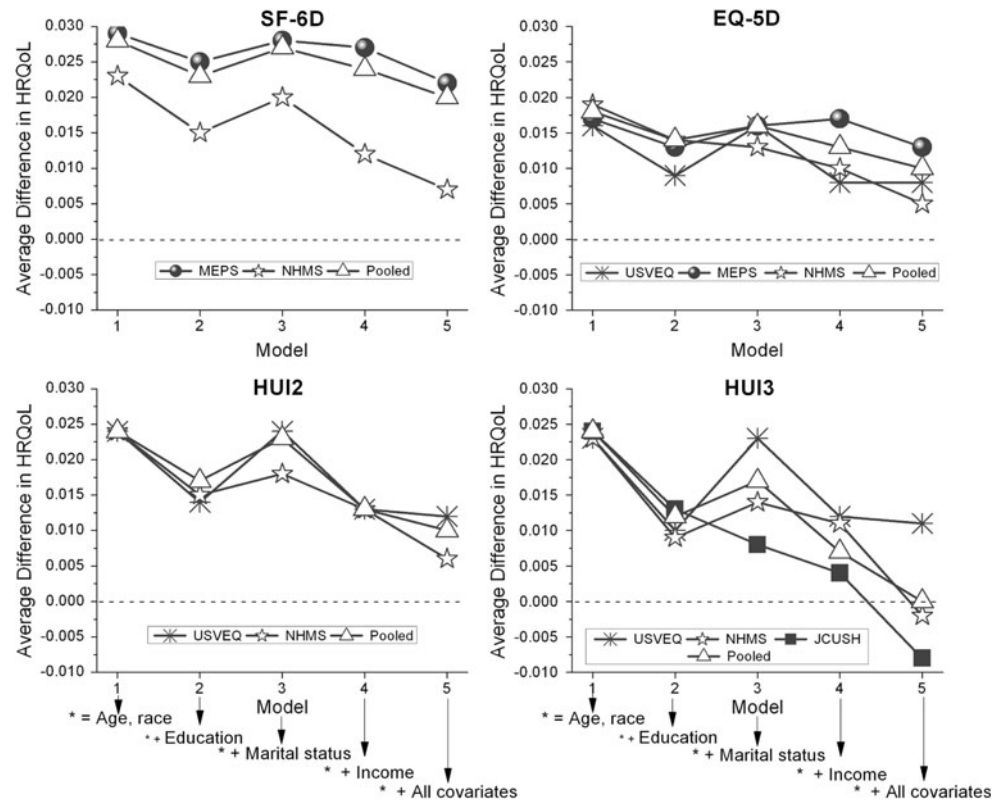


Taking income differentials into account resulted in estimates that were again quite consistent across measures with a few exceptions. EQ-5D retained a statistically significant gender difference in the pooled analysis due to the large sample size of MEPS, although the magnitude of the difference was similar to the non-significant differences in other measures. In contrast, after controlling for income or marital status or all covariates simultaneously, the gender difference in SF-6D remained moderate and statistically significant in MEPS and pooled analyses, as did the difference in QWB-SA from NHMS.

Several circumstances may have contributed to the greater and persistent gender differences in SF-6D and QWB-SA. It is well known that the five HRQoL measures differ in construction (e.g., HRQoL domains, time frames, elicitation methods and scoring equations) and distributional properties (e.g., ceiling/floor effects, numerical range), although they purport to represent the same evaluation of a given level of health [6]. The observed gender differences in HRQoL reported here may then be in part an artifact of the particular HRQoL index used. For example, SF-6D and QWB-SA have minimal ceiling effects compared to the other three indexes [6], which may lead to these two measures identifying gender differences between relatively healthy men and women. The QWB-SA is also the only measure that incorporates in its summary score a list of 58 acute/chronic symptoms and health conditions, and such

symptoms may differ by gender. The fact that the remaining gender difference in SF-6D after adjustment for marital status and income (in the pooled analysis) arose almost entirely from MEPS and not from the NHMS raises the possibility that differences in adjustment variable categorization may have led to differences in results, such as differences in measurement of the income variable in MEPS.

MEPS and NHMS also differ in mode of administration. Hanmer et al. (2007) [7] showed that self-administered surveys (e.g., MEPS) yield lower HRQoL scores than telephone surveys (e.g., NHMS) on the EQ-5D, Visual Analog Scale, HUI3 and general self-rated health question. Hays et al. (2009) [29] found HRQoL scores are more positive for phone administration following mail administration with differences between modes: 0.06 (SF-6D), 0.03 (QWB-SA), 0.08 (EQ-5D), 0.04 (HUI2) and 0.10 (HUI3). Further analyses (not shown) revealed that distributions on domains of SF-6D in MEPS and NHMS are similar except for the mental health domain. Greater proportions of both men and women reported lower levels of mental health in MEPS than in NHMS; however, the proportions differed more for women than men between the two datasets. This may indicate that differences in SF-6D results between MEPS and NHMS may be related to a greater mode effect among women on SF-6D.

Alternatively, persistent gender differences in SF-6D and QWB-SA may reflect unique health variation captured

by these indexes but not by EQ-5D, HUI2 or HUI3, or gender biases in item responses. Fleishman and Lawrence (2003) found, based on the MEPS data, that some questions on SF-12v2$^{TM}$ are prone to differential item functioning (DIF) between the genders, i.e., different responses to the questions by men and women having the same underlying health [30]. In particular, the two questions that are used to estimate the mental heath component of SF-6D ("felt downhearted", "had energy") were found to have DIF by gender [17, 30, 31]. Fleishman and Lawrence (2003) showed that adjusting for DIF generally reduced gender differences in mental health [30].

HUI3 in NHMS and JCUSH was the only HRQoL index that showed a reverse direction of the gender difference, though not statistically significant, with simultaneous adjustment for age, race, marital status, education and income. HUI3 was also more affected by adjustments for SES than were other measures. The relatively greater sensitivity to SES adjustments on HUI3 is consistent with findings of Robert et al. (2009) [13] who showed that income-associated disparities appear wider for the HUI3 than for SF-6D or EQ-5D.

There may be several reasons for minor differences in results between the studies. Random variation, differences in sampling and mode effects (JCUSH and NHMS were phone surveys; MEPS and USVEQ were self-administered) may have led to such differences [7]. Our analysis aims to represent the sociodemographic and SES mix of the US population and may be smaller or larger in certain subgroups or in populations with a different sociodemographic and SES mix. Future research on such interactions is warranted. Furthermore, the analysis adjusted the results to marginal percentages in age, race and socioeconomic subgroups, but not to percentages in subgroups formed by 2 or 3-way cross-tabulation of these variables, which may also have been affected by response biases.

Small gender differences in HRQoL that persist even after adjustments for available sociodemographic and SES characteristics indicate either that this variation has been incompletely measured, or that other factors are contributing to gender differences in HRQoL. We may not be measuring accurately the full distribution of marital status, education and income, especially as we have collapsed these variables in our attempt to achieve consistency across data sets. In addition, it may be that dynamics of marital status and income over the life course contribute significantly to gender differences in HRQoL—dynamic effects that our cross-sectional analyses cannot detect. Clearly, age, race, marital status, education and income are likely not the only sociodemographic and SES variables that contribute to HRQoL differences between men and women.

Our analysis focuses on whether there are gender differences in HRQoL and whether these differences are related to sociodemographic and SES differentials between men and women. Alternatively, the analysis could have examined whether other measures of self-perceived health (e.g., general self-rated health question, symptoms/conditions and morbidity indexes) explain gender differences in HRQoL. Such analyses would aim to answer different questions and may lead to different conclusion regarding gender differences. Several considerations would affect the interpretation of analyses attempting to adjust for such measures, including measurement error or bias in the adjustment variables themselves. HRQoL measures have been developed to capture the multiple dimensions of the health experience beyond the presence of disease. However, examining gender differences in HRQoL among men and women with the same illness or disease may be a fruitful direction of future research.

This study has several limitations. These analyses are based on cross-sectional data so it was not possible to assess gender differences in changes in HRQoL over time. The inability to consistently control for household or family size in measuring household income may be a weakness of this study. Additionally, this study is based on a non-institutionalized sample and at least two institutionalized segments of the US population, those who are hospitalized or in nursing homes, are likely to have the worst HRQoL than people in our study and are predominantly women [32, 33]. Due to sample age and race restrictions in the analyses, only community-residing 35–89-year-old black and white US adults are represented in the results.

The primary strength of the study is the simultaneous use of four recently conducted, large nationally representative surveys among US adults and several commonly used preference-based measures of HRQoL. To further understand the scope of gender differences in HRQoL, it is important to conduct future research on gender and HRQoL in other subpopulations of the United States, including residents ages <35, other racial subgroups of the population and people living in institutions. Incorporating longitudinal data would allow an assessment of how aging affects changes in HRQoL in men and women. It is also important to assess whether gender differences in HRQoL are larger in some subgroups of the population.

## Conclusion

This study of gender differences in five commonly used measures of HRQoL from four nationally representative studies indicated that US women on average have slightly lower self-reported HRQoL than men. The HRQoL gap is largely explained by sociodemographic and SES differentials between men and women. All five measures of

HRQoL showed similar gender differences prior to sociodemographic and SES adjustments and relatively similar patterns with adjustment. Income and marital status contributed the most to explaining gender differences in HRQoL across measures. Our findings underscore the impact of socioeconomic disparities on the well-being of women and point to the importance of considering such disparities prior to attributing poor HRQoL in women to other gender-related factors.

## References

1. Courtenay, W. H. (2000). Behavioral factors associated with disease, injury, and death among men: Evidence and implications for prevention. *Journal of Men's Studies, 9*(1), 81–142.

2. Courtenay, W. H., McCreary, D. R., & Merighi, J. R. (2002). Gender and ethnic differences in health beliefs and behaviors. *Journal of Health Psychology, 7*(3), 219–231.

3. Bird, D. C., & Rieker, P. P. (2008). *Gender and health: The effects of constrained choices and social policies*. New York: Cambridge University Press.

4. Kochanek, K. D., Murphy, S. L., Anderson, R. N., Scott, C. (2004). Deaths: Final data for 2002. National vital statistics reports. Hyattsville, Maryland: National Center for Health Statistics, 53(5).

5. Hanmer, J., Lawrence, W. F., Anderson, J. P., et al. (2006). Report of nationally representative values for the non-institutionalized US adult population for 7 health-related quality-of-life scores. *Medical Decision Making, 26*, 391–400.

6. Fryback, D. G., Dunham, N. C., Palta, M., et al. (2007). US norms for 6 generic health-related quality of life indexes from the national health measurement study. *Medical Care, 45*, 1162–1170.

7. Hanmer, J., Hays, R. D., Fryback, D. G., et al. (2007). Mode of administration is important in US national estimates of health-related quality of life. *Medical Care, 45*, 1171–1179.

8. Jia, H., Lubetkin, E. I., Moriarty, D. G., & Zack, M. M. (2007). A comparison of healthy days and EuroQol EQ-5D measures in two US adult samples. *Applied Research in Quality of Life, 2*(3), 209–221.

9. Luo, N., Johnson, J. A., Shaw, J. W., et al. (2005). Self-reported health status of general adult US population as assessed by the EQ-5D and health utilities index. *Medical Care, 43*(11), 1078–1086.

10. Sullivan, P. W., Lawrence, W. F., & Ghushchyan, V. (2005). A national catalog of preference-based scores for chronic conditions in the United States. *Medical Care, 43*(7), 736–749.

11. Shinberg, D. S. (1999). For richer, for poorer, in sickness and in health: Socioeconomic status and health among married couples. *Annals of the New York Academy of Sciences, 896*, 341–343.

12. Liu, H., & Umberson, D. J. (2008). The times they are a changin': Marital status and health differentials from 1972 to 2003. *Journal of Health and Social Behavior, 49*, 239–253.

13. Robert, S. A., Cherepanov, D., Palta, M., Dunham, N. C., Feeny, D., & Fryback, D. G. (2009). Socioeconomic status and age variations in health-related quality of life: Results from the national health measurement study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 64B*(3), 378–389.

14. Heckert, M. T., Droste, E. H., Adams, J. P., Griffin, M. C., Roberts, L. L., Mueller, M. A., et al. (2002). Gender differences in anticipated salary: Role of salary estimates for others, job characteristics, career paths, and job inputs. *Sex roles, 47*, 139–151.

15. Kreider, R. M., Simmons, T. (2000). Marital Status: 2000. US Department of Commerce. Economics and Statistics Administration. US Census Bureau, 2003. Census 2000 Brief. Publication No. C2KBR-30.

16. 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.

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

18. Brooks, R., Rabin, R., & de Charro, F. (2003). *The measurement and valuation of health status using EQ-5D: A European perspective*. Dordrecht. Netherlands: Kluwer Academic Publishers.

19. Shaw, J. W., Johnson, J. A., & Coons, S. J. (2005). US valuation of the EQ-5D health states: development and testing of the DI valuation model. *Medical Care, 43*(3), 203–220.

20. Torrance, W. G., Feeny, D. H., Furlong, W. J., Barr, R. D., Zhang, Y., & Wang, Q. (1996). Multiattribute utility function for a comprehensive health status classification system: Health utilities index mark 2. *Medical Care, 34*(7), 702–722.

21. Feeny, D., Furlong, W., Torrance, G. W., et al. (2002). Multiattribute and singleattribute utility functions for the health utilities index mark 3 system. *Medical Care, 40*, 113–128.

22. Kaplan, R. M., & Anderson, J. P. (1988). A general health policy model: Update and applications. *Health Services Research, 23*, 203–235.

23. Andresen, E. M., Rothenberg, B. M., & Kaplan, R. M. (1998). Performance of a self-administered mailed version of the quality of well-being (QWB-SA) questionnaire among older adults. *Medical Care, 36*, 1349–1360.

24. Sanmartin, C., Berthelot, J. M., Ng, E., et al. (2006). Comparing health and health care use in Canada and the United States. *Health Affairs (Millwood), 25*, 1133–1142.

25. Cohen, J. W., Monheit, A. C., Beauregard, K. M., et al. (1996–1997). The Medical Expenditure Panel Survey: a national health information resource. Rockville, MD: Center for Cost and Financing Studies, Agency for Health Care Policy and Research. Inquiry, 33(4), 373–389.

26. Rhoades, J. A. (2005). The Long-Term Uninsured in America, 2002–2003: Estimates for the US Population under Age 65. Rockville, MD: Agency for Healthcare Research and Quality, 2005. Statistical Brief #104.

27. Lubetkin, E. I., Jia, H., Franks, P., & Gold, M. R. (2005). Relationship among sociodemographic factors, clinical conditions, and health-related quality of life: Examining the EQ-5D in the U.S. general population. *Quality of Life Research, 14*(10), 2187–2196.

28. Wyrwich, K. W., Bullinger, M., Aaronson, N., Hays, R. D., Patrick, D. L., & Symonds, T. (2005). The clinical significance consensus meeting group. Estimating clinically significant differences in quality of life outcomes. *Quality of Life Research, 14*(2), 285–295.

29. Hays, R. D., Kim, S., Spritzer, K. L., Kaplan, R. M., Tally, S., Feeny, D., et al. (2009). Effects of mode and order of administration on generic health-related quality of life scores. *Value in Health, 12*(6), 1035–1039.

30. Fleishman, J., & Lawrence, W. (2003). Demographic variation in SF-12 scores: True differences or differential item functioning? *Medical Care, 41*(7), 75–86.

31. Ware, J., et al. (2002). How to Score Version 2 of the SF12 Health Survey, Quality Metric Inc.

32. Merrill, C. T., Elixhauser, A. (2005). Hospitalization in the United States, 2002. Rockville, MD: Agency for Healthcare Research and Quality, 2005. HCUP Fact Book No. 6. AHRQ Publication No. 05-0056. ISBN 1-58763-217-9.

33. National Center for Health Statistics. (2009). Health, United States, 2008, with Special Feature on the Health of Young Adults. Hyattsville, MD: US Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics, 2009. Library of Congress Catalog No. 76-641496.

34. Kaplan, R. M. (2005). The minimally clinically important difference in generic utility-based measures. *COPD, 2*(1), 91–97.