

Predicting EQ-5D-US and SF-6D societal health state values from the Osteoporosis Assessment Questionnaire

C. M. McDonough · M. R. Grove · A. D. Elledge ·
A. N. A. Tosteson

Received: 31 August 2010 / Accepted: 14 February 2011 / Published online: 12 April 2011
© International Osteoporosis Foundation and National Osteoporosis Foundation 2011

Abstract

Summary Linear regression was applied to data from 275 persons with osteoporosis-related fracture to estimate EQ-5D-US and SF-6D health state values from the Osteoporosis Assessment Questionnaire. The models explained 56% and 58% of the variance in scores, respectively, and root mean square error values (0.096 and 0.085) indicated adequate prediction for use when actual values are unavailable.

Introduction This study was conducted to provide models that predict EQ-5D-US and SF-6D societal health state values from the Osteoporosis Assessment Questionnaire (OPAQ).

Methods OPAQ, EQ-5D, and SF-6D data from individuals at two centers with prior osteoporosis-related fracture were used. Fractures were classified by type as hip/hip-like,

spine/spine-like, or wrist/wrist-like. Spearman rank correlations between preference-based system (EQ-5D and SF-6D) dimensions and OPAQ subscales were estimated. Linear regression was used to estimate preference-based system health state values based on OPAQ subscales. We assessed models including age, sex, and fracture type and chose the model with the best performance based on the root mean square error (RMSE) estimate.

Results Among the 275 participants (198 women), with mean age of 68 years (range 50–94), the distribution of fracture types included 10% hip/5% hip-like, 18% spine/11% spine-like, and 24% wrist/18% wrist-like. The final regression model for EQ-5D-US included three OPAQ attributes (physical function, emotional status, and symptoms), predicted 56% of the variance in EQ-5D-US scores, and had a RMSE of 0.096. The final model for SF-6D, which included all four OPAQ dimensions, predicted 58% of the variance in SF-6D scores and had a RMSE of 0.085. **Conclusions** Two models were developed to estimate EQ-5D-US and SF-6D health state values from OPAQ and demonstrated adequate prediction for use when actual values are not available.

C. M. McDonough (✉) · M. R. Grove · A. N. A. Tosteson
The Dartmouth Institute for Health Policy and Clinical Practice,
Dartmouth College,
Lebanon, NH 03756, USA
e-mail: Christine.M.McDonough@Dartmouth.edu

A. N. A. Tosteson
Department of Medicine, Dartmouth-Hitchcock Medical Center,
Lebanon, NH, USA

A. N. A. Tosteson
Department of Community and Family Medicine,
Dartmouth Medical School,
Hanover, NH, USA

C. M. McDonough
Health & Disability Research Institute,
Boston University School of Public Health,
Boston, MA, USA

A. D. Elledge
New York, NY, USA

Keywords Cost-utility analysis · Economic evaluation · Health state values · Osteoporosis · Societal preferences

Introduction

To characterize the value of competing osteoporosis therapies, cost-utility analyses that use societal health state values to measure effectiveness are often undertaken [1]. The incremental cost-effectiveness ratio, typically defined as the ratio of added cost to changes in health using the Quality-Adjusted Life Year (QALY) when increasingly costly alternatives are compared, is the focus of the analysis.

Estimating QALYs requires a measure of health using a single score, anchored at 0 for dead and 1 for perfect health, and weighted for the relative desirability of the health state. Standards for economic evaluations recommend societal values (also called preferences or utilities) [2]. Unfortunately, measures of health state value are rarely available from clinical trial participants. Instead, clinical trials more commonly include disease-targeted health status instruments, such as the Osteoporosis Assessment Questionnaire (OPAQ) [3]. The OPAQ yields four domain scores (physical function, psychological status, symptoms, and social interaction) that cannot be used directly for cost-utility analyses. There are two main approaches used to obtain “societal health state values”(HSV): (1) direct measurement of value for health states from a representative sample of the population using methods such as standard gamble [4], time tradeoff [5], and visual analogue scale ratings [6] and (2) indirect measurement using preference-based measurement systems such as the Quality of Well Being Scale [7], the EuroQOL EQ-5D [8], the SF-36-based SF-6D [9], and the Health Utilities Index [10]. These systems include a generic self-report health state classification system and an algorithm to apply a set of societal preference weights for the range of health states described by the system.

Preference-based systems offer the opportunity to measure changes in terms of societal HSVs without the resource requirements of direct measurement. However, in reality, many clinical trials in osteoporosis do not utilize preference-based systems. Choosing outcome measurement instruments for clinical trials involves serious tradeoffs. Researchers must balance the need to maximize measurement precision for the detection of differences against practicality of administration and limits in available resources. With these concerns in mind, unless cost-utility analysis is a specific research aim, investigators often favor disease-specific measures over the generic preference-based instruments because the former focus on the key aspects targeted by treatment. Therefore, HSVs required for economic evaluation are rarely available from trial participants.

There are situations in which researchers wish to conduct cost-utility analysis without measuring health states directly. They may choose to use age- and sex-specific HSVs from a published source, such as time tradeoff-based values of current health from a population sample [11]. However, populations used to estimate published HSVs may be different from the population of interest in important ways. To address this challenge, methods have been developed to estimate HSVs from data provided by more widely used outcome measures. Several methods have been explored to estimate HSVs from the SF-36 [12–14], the SF-12 [15–21], condition-specific tools [22–25], and national survey instruments and therefore allow the calculation of QALYs [2, 26]. The objectives of this study were to provide a model to

predict societal HSVs from the preference-based systems EQ-5D-US and SF-6D using data from the OPAQ and to characterize the impact of osteoporotic fracture in terms of mean OPAQ dimension scores and EQ-5D-US and SF-6D HSVs by fracture type.

Methods

Study population

This is a secondary analysis of pooled data from two studies of persons at least 50 years old with a non-traumatic fracture within the prior 5 years recruited at Dartmouth-Hitchcock Medical Center in Lebanon, NH and the Mayo Clinic in Rochester, MN [27, 28]. The first study included women with non-traumatic fractures of the hip or spine within the prior 5 years, and the second included men with non-traumatic fractures of the hip and spine and women with fractures in other locations.

Fracture categorizations used in this study were based on published expert panel classifications [29]. In summary, three main fracture locations of hip, vertebra, and wrist were identified, and fracture classification criteria were created for other fractures in order to characterize their effects relative to these more common fractures. A panel of physician experts familiar with aspects of osteoporosis-related fractures defined the following “like” categories: (1) *like a hip* = shaft/distal femur, proximal tibia/fibula; (2) *like a spine* = cervical spine, ribs, sternum, clavicle, scapula, pelvis; (3) *like a wrist* = humerus, shaft/proximal radius, ankle, distal tibia/fibula; and (4) *like a finger/toe* = hand/fingers, foot/toes, and patella.

Participation was requested from every patient visiting one of the participating medical centers’ osteoporosis clinics. Each participant completed a series of questionnaires at the time of visit to the medical center or via mail following a visit to a participating medical center. Patients were excluded if they had experienced multiple fractures or demonstrated impaired cognitive function. Participants with missing essential variables for analyses due to incomplete questionnaires were removed from the analysis.

Measures

The measures used in this study, including the OPAQ, the Medial Outcomes Study Short Form (SF-36), the SF-6D, and the EuroQoL EQ-5D, are described below.

OPAQ

The Osteoporosis Assessment Questionnaire is a disease-targeted instrument designed to measure health status in

fracture patients [3, 30, 31]. There are several versions including the original 73-question OPAQ 1.0, the 60-question OPAQ 2.0, and the 34-question OPAQ-SV. In an early validity study, the OPAQ 1.0 demonstrated evidence of validity and mean Cronbach's alpha of 0.86 ± 0.01 [3]. In this study, we used the self-completed OPAQ 2.0, which consists of 60 questions categorized within 14 domains, and further grouped within the following four dimensions: physical function, emotional status, symptoms, and social interaction. Item scores are summed to create the 14 domain scores, which are reported from 0 to 10. Domain scores are summed and rescaled to yield dimension scores from 0 to 100, with higher scores indicating better status.

SF-36-derived SF-6D

The SF-36 was developed in the Medical Outcomes Study for the measurement of health status, and its measurement properties have been extensively studied [32]. The SF-36 has been used in the Multicentre Osteoporosis Study (CaMos). It measures eight dimensions on up to six levels. Our research used version 1 of the SF-36, with the reference time frame for health status of "the past 4 weeks." The SF-6D provides a method for deriving a preference-weighted health state classification system from the SF-36 instrument using items from six of the eight dimensions with response categories up to six levels [9, 33]. We used version 2 of the SF-6D, which utilizes an additive statistical model to provide health state values ranging from 0.29 to 1.00. Societal HSVs were modeled from a representative sample of 611 adults from the UK using the standard gamble.

EQ-5D

The widely used and extensively studied EuroQoL EQ-5D includes five attributes rated on three levels to define 245 health states (with "dead" and "unconscious" added) [8]. Preferences are available from multiple populations with different elicitation methods using the same EQ-5D descriptive system. We used preference weights from a US population sample [34]. We will refer to the health state classification system as EQ-5D and the HSVs as EQ-5D-US. The scoring algorithm uses the "D1" random effects model for EQ-5D-US, which provides scores between -0.11 to 1.00 based on a representative sample of 4,048 adults from the USA [34]. The reference time frame for health status used in our study was "today."

Statistical analysis

Statistical analyses were conducted after pooling the two groups. To examine potential differences in characteristics

or predictive ability of the models and final algorithms, additional analyses were performed with stratified data by fracture type and/or gender to examine potential differences in characteristics or predictive ability of the models.

Demographic data were summarized for the entire participant population. As appropriate, the mean and standard deviation or percentage and count for each characteristic were calculated and described.

For each EQ-5D and SF-36 dimension, OPAQ domain and dimension means and standard deviations were calculated. Tests of trend across EQ-5D dimensions and OPAQ domain and dimension scores were performed to assess the relationship between the measures, with the hypothesis that EQ-5D and SF-6D dimensions would correlate with similar OPAQ dimensions and related domains.

We calculated Spearman correlations for the EQ-5D and SF-6D dimensions and the respective OPAQ domain. We performed linear regression analyses using OPAQ dimension scores to predict EQ-5D-US and SF-6D HSVs. We first conducted univariate regression to assess the relationships between HSVs and demographic and clinical variables. For each preference-based measure, a comprehensive multivariate model was generated to predict HSVs using OPAQ dimensions and variables found statistically significant in univariate regressions. From the comprehensive model, a parsimonious model was created by removing variables and OPAQ scores that were not statistically significant within the comprehensive model. In addition, the parsimonious model was examined stratified by gender and fracture type. Model fit was assessed by comparing the distribution of observed and expected HSVs, visual inspection of plots for observed versus predicted values and of the residuals. Model performance for prediction of HSVs was assessed using the R^2 and root mean square error estimates (RMSE). Analyses for this study were completed with STATA, version 9 (STATA Corporation, College Station, TX, USA).

Results

Among a total of 291 study participants with fracture, 16 were excluded due to missing questionnaire responses and 275 were included in the analysis. The mean age of the participants was 68.1 years (± 10.6), with a range in age from 50 to 94 years, and 72% were female (Table 1). Grouped by fracture type, hip fracture participants had the highest mean age of 74.2 years (SD 11.23), while wrist-like fracture participants had the lowest average age of 64.41 years (SD 9.09).

Participants' mean EQ-5D-US HSV was 0.85 (SD 0.14) and mean SF-6D HSV was 0.76 (SD 0.13). OPAQ's four main dimension means (SD) for the entire group were:

Table 1 Characteristics of the sample ($n=275$)

	All participants
Age	
50–64 years	111 (40) ^a
65 years and over	164 (60) ^a
Female sex	198 (72) ^a
Fracture type	
Wrist/wrist-like	153 (56) ^a
Spine/spine-like	80 (29) ^a
Hip/hip-like	42 (15) ^a
Ethnicity: Hispanic	1 (<1) ^a
Race: White	271 (99) ^a
Education: > high school	171 (63) ^a
EQ-5D-US health state value	0.85 (0.14) ^b
SF-6D health state value	0.76 (0.13) ^b
OPAQ dimension scores	
Physical function	85.9 (17.5) ^b
Emotional status	70.1 (16.1) ^b
Symptoms	65.7 (18.7) ^b
Social interaction	60.3 (16.8) ^b

^a n (%)^b Mean (SD)

physical function 86.1 (17.4), emotional status 70.4 (16.1), symptoms 65.74 (18.6), and social interaction 60.6 (16.9).

Statistically significant correlations were observed for all EQ-5D dimensions with all OPAQ dimensions except social interaction and for all SF-6D dimensions with all of the OPAQ dimensions ($p<0.001$) (Table 2). OPAQ's physical function dimension exhibited the strongest significant correlations with EQ-5D dimensions, ranging from

–0.34 to –0.60, followed by emotional status, and symptoms. A similar pattern with stronger correlation coefficients was noted between OPAQ dimensions and SF-6D, and the correlation between SF-6D dimensions and OPAQ social interaction was significant. All domains that comprise OPAQ's physical function, emotional status, and symptoms dimensions exhibited significant correlations to all EQ-5D dimensions. For the domains that form OPAQ's social interaction dimension, the only significant, but slight, correlations were for the social network domain to EQ-5D usual activities and self-care dimensions.

Tests for trend demonstrated similar consistency between OPAQ dimensions and EQ-5D and SF-6D dimensions. Tests for OPAQ's physical function, emotional status, and symptoms dimensions were all highly significant ($p<0.001$) with each of the five EQ-5D dimensions. OPAQ's social interaction dimension, again, did not show a significant trend with EQ-5D. OPAQ domains demonstrated similar trend patterns.

Univariate regression models

In separate, univariate regression analyses, each of OPAQ's four dimensions demonstrated statistically significant relationships with EQ-5D-US and SF-6D HSVs. The relationship between social interaction and EQ-5D-US and SF-6D HSVs was the weakest of the four dimensions. The following variables were also individually statistically significant predictors of EQ-5D-US HSVs: age (as continuous variable, $p=0.004$), BMI ($p=0.006$), alcohol use (categorized as <1/week or \geq 1/week, $p=0.012$), education (categorized as up to high school versus beyond high school, $p<0.001$), current employment ($p=0.001$), and Deyo score (an adaptation of the Charlson comorbidity

Table 2 Spearman correlation coefficients for dimensions of EQ-5D and SF-6D with OPAQ

	OPAQ dimensions			
	Physical function	Emotional status	Symptoms	Social interaction
EQ-5D-US dimensions, $N=275$				
Mobility	–0.60*	–0.47*	–0.36*	0.00
Pain/discomfort	–0.54*	–0.42*	–0.54*	–0.08
Usual activities	–0.61*	–0.47*	–0.46*	–0.09
Self-care	–0.44*	–0.35*	–0.28*	0.11
Anxiety/depression	–0.34*	–0.46*	–0.36*	–0.06
SF-6D dimensions, $N=275$				
Physical function	–0.71*	–0.59*	–0.51*	–0.04
Mental health	–0.27*	–0.47*	–0.32*	–0.17*
Role limitation	–0.57*	–0.55*	–0.45*	–0.15*
Social functioning	–0.50*	–0.52*	–0.42*	–0.24*
Vitality	–0.58*	–0.56*	–0.56*	–0.15
Pain	–0.64*	–0.49*	–0.60*	–0.09*

For both and EQ-5D and SF-6D responses, higher levels indicate worse function (e.g., for EQ-5D1 = no problems; 2 = some problems; 3 = unable)

* $p<0.001$, statistically significant

index, $p < 0.001$) [35, 36]. Race, Hispanic origin, currently smoking, living alone, and taking alendronate or raloxifene were not statistically significant predictors of EQ-5D-US HSVs in individual regressions. For SF-6D, univariate relationships were similar except that BMI did not predict SF-6D HSVs.

Multivariate regression models

Age, alcohol use, education, and employment were no longer statistically significant in the comprehensive regression models. OPAQ’s social interaction dimension was also no longer a significant predictor of EQ-5D-US in the comprehensive model, which demonstrated an adjusted R^2 of 0.63, $p < 0.001$. The final regression model for EQ-5D-US, shown in Table 3, included three of the four OPAQ dimensions (physical function, emotional status, and symptoms). The ability of the model to predict EQ-5D-US HSVs was indicated by the $R^2 = 0.56$, $RMSE = 0.096$, and $p < 0.001$. The range for actual EQ-5D-US HSVs was 0.118 to 1 and the range for predicted values was 0.408 to 0.995. The model for SF-6D included age and all four of the OPAQ dimensions (Table 3). This model demonstrated a slightly stronger predictive ability than the EQ-5D-US model, as shown by the $R^2 = 0.58$, $RMSE = 0.085$, and $p < 0.001$.

Model fit

Figure 1 shows the comparisons between predicted and observed health state values for EQ-5D-US and SF-6D for all subjects and stratified by gender. The observed HSVs were skewed to the right for both, with a ceiling effect noted for EQ-5D-US. Similar patterns were found for male and female subjects. Transformations of the independent and dependent variables did not provide improved fit for the models, and therefore, untransformed data were used in

the final models. Predicted scores were lower than observed at the low end of the range for SF-6D and higher than observed for EQ-5D-US.

Effect of fracture type on health state values

Table 4 summarizes mean OPAQ dimension, EQ-5D-US, and SF-6D scores by fracture type group. The wrist/wrist-like/other fracture group had the highest OPAQ dimension means for physical function, emotional status, and symptoms, while the spine/spine-like fracture group had the highest mean for social interaction. The hip/hip-like group had the lowest means for physical function, emotional status, and social interaction, and the spine/spine-like group had the lowest mean score for symptoms.

The effect of fracture type on estimates of health state value was assessed by including indicators for fracture type in the model and evaluating models stratified by fracture type. We examined the effect of fracture type using each of the main fracture types and by collapsing fracture types into three categories, combining other fractures (hand/fingers, foot/toes, patella) with wrist and wrist-like fractures. Regardless of categorization, fracture type was not a significant predictor of health state value in multivariate models. Using the three most common fracture types, in Fig. 2, we illustrate that the confidence intervals for all regression coefficient estimates overlapped across fracture types.

Discussion

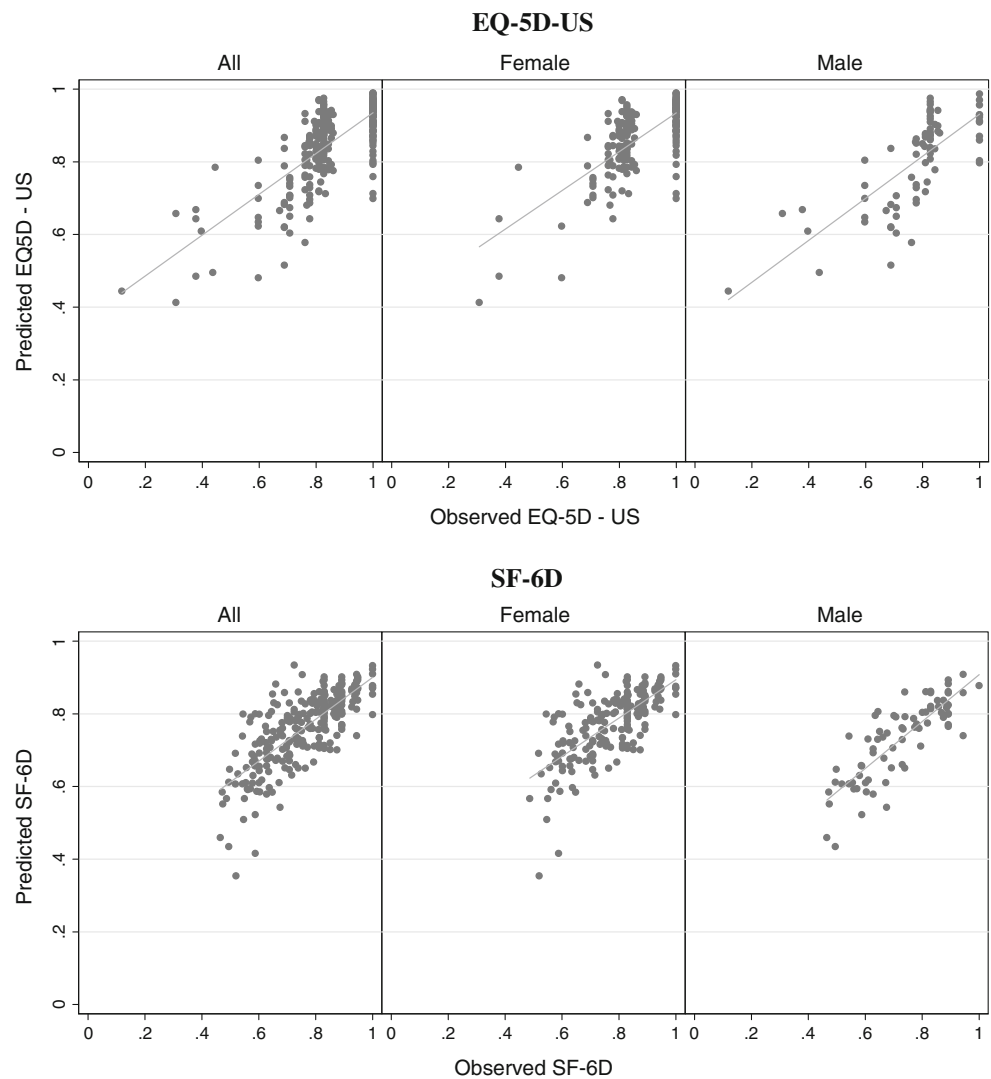
This study of 275 persons with prior osteoporosis-related fracture provided a linear regression-based method to estimate EQ-5D-US and SF-6D HSVs from the widely used osteoporosis-targeted outcome measure, the OPAQ.

Table 3 Regression estimates for linear models to predict EQ-5D-US and SF-6D using OPAQ dimension scores

Predictor	Coefficient	95% CI		<i>p</i>
EQ-5D-US prediction model: $n = 275$, $R^2 = 0.563$, $RMSE = 0.096$				
Age	0.0005555	-0.0006324	0.0017434	0.36
OPAQ physical function	0.0042326	0.0032046	0.0052607	<0.001
OPAQ emotional status	0.0012544	0.0001257	0.0023830	0.030
OPAQ symptoms	0.0014490	0.0005280	0.0023700	0.002
Constant	0.2675196	0.1495928	0.38544650	<0.001
SF-6D prediction model: $n = 275$, $R^2 = 0.580$, $RMSE = 0.085$				
Age	-0.000841	-0.0018959	0.0002139	0.12
OPAQ physical function	0.0020259	0.0011144	0.0029373	<0.001
OPAQ emotional status	0.0027485	0.0017336	0.0037634	<0.001
OPAQ symptoms	0.0013691	0.0005535	0.0021846	0.001
OPAQ social interaction	0.0007735	0.0001570	0.0013900	0.014
Constant	0.3187351	0.2115204	0.4259498	0.001

RMSE root mean square error

Fig. 1 Predicted versus observed health state values ($n=275$)



The strength of the relationships demonstrated between EQ-5D-US, SF-6D, and the OPAQ supports the use of this approach when HSVs are not available from the trial data. The R^2 estimates of 0.57 for the EQ-5D-US model and 0.58 for the SF-6D model are at the high end of the range (0.31–0.66) for published models including those using similar methods reported by Mortimer and Segal [37, 38]. Furthermore, RMSE values found in this study indicated

acceptable prediction of HSV compared to other prediction models [23, 24, 39].

Whenever possible, a preference-based system should be utilized to measure HSVs in clinical studies. However, it is often the case that such measures are not included due to the serious resource challenges faced by investigators. Under such circumstances, the use of societal HSVs estimated from condition-specific measures such as the

Table 4 Mean scores by fracture type for OPAQ dimensions and estimated EQ-5D-US and SF-6D

Scores, mean (SD)	Hip/hip-like ($n=42$)	Spine/spine-like ($n=80$)	Wrist/wrist-like/other ($n=153$)
OPAQ physical function	72.2±25.7	81.93±18.1	91.8±10.4
OPAQ emotional status	61.0±19.6	68.5±17.9	73.4±13.0
OPAQ symptoms	63.0±19.4	61.4±20.3	68.7±17.1
OPAQ social interaction	57.6±20.2	63.3±15.3	59.4±16.5
Estimated EQ-5D-US health state value	0.78±0.15	0.83±0.12	0.88±0.07
Estimated SF-6D health state value	0.70±0.13	0.75±0.11	0.79±0.07

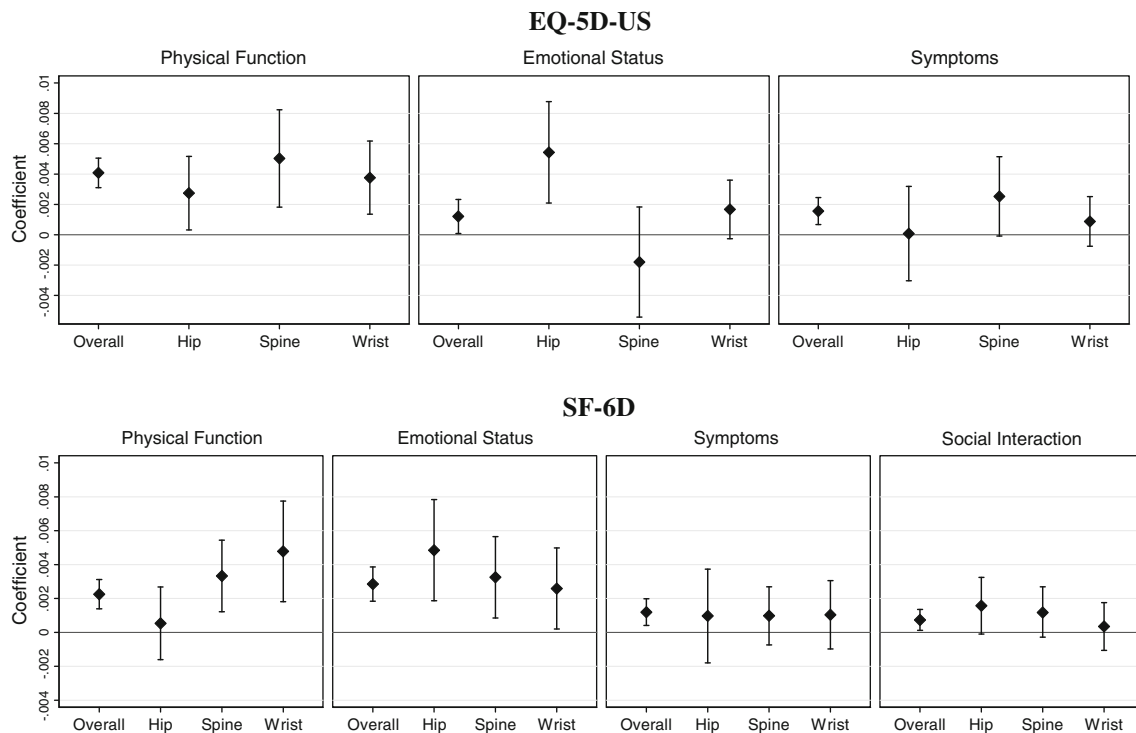


Fig. 2 Impact of fracture type on OPAQ coefficients showing regression coefficients for fracture type compared to the overall coefficient for $n=275$ participants. Fracture types: hip = hip/hip-like ($n=42$), spine = spine/spine-like ($n=80$), wrist = wrist/wrist-like/other ($n=153$)

OPAQ may provide the opportunity to conduct cost-utility analyses that would not otherwise be possible.

The important advantages of leveraging existing data to address pressing questions about the relative value of interventions must be considered against the limitations of this approach compared to more direct measurement. Limited variability of estimates and diminished measurement precision have been associated with regression-based “mapping” approaches [40–43]. Our data provide support for some concern about limited variability. The range of predicted HSVs was smaller for both EQ-5D-US and SF-6D, though this was less apparent for SF-6D than for EQ-5D-US. This could potentially translate into difficulty in detecting meaningful differences in health outcomes in studies that use this approach to estimate health state values. On the other hand, it could be argued that estimating HSVs from disease-targeted instruments such as the OPAQ may provide an additional degree of sensitivity or responsiveness by detecting change in clinically important areas where a generic instrument may not. It is unlikely that any improvements in measurement of change over time would outweigh the loss of variation and precision associated with predicting HSVs; however, future research could prospectively compare the performance of these approaches.

Another challenge presented in the application of estimation methods is in the interpretation of study results

using different methods. Several approaches have been utilized to convert from health status data to HSVs, including regression modeling, effect size translation, and response mapping [37], and there is evidence to support the significant variation in the results [40–42]. One review in this area was unable to identify a superior approach [37]. From the perspective of policy decision-making, comparability of various methods used to estimate societal HSVs is the subject of ongoing discussion and investigation [40, 44–50]. Meaningful differences in estimates from different approaches may undermine the fundamental strength of cost-utility analysis—the ability to compare outcomes across diseases and interventions. While there is some indication that variation exists between estimates from different approaches in some cases, how often and how much is yet unknown. This serious concern must be weighed against the need to optimize the use of limited health care research resources.

Given that our models account for 56% and 58% of the variance in EQ-5D-US and SF-6D HSVs, respectively, it is important to consider possible sources of the unexplained variation. For each instrument, a portion of the unexplained variation would be attributable to measurement error. In addition, variation in coverage of constructs is a likely contribution. Although SF-6D, EQ-5D, and OPAQ cover similar key dimensions of health including physical function, emotional status, symptoms, and social interac-

tion, there are important differences. SF-6D dimensions include role limitations and vitality and EQ-5D describes usual activities and self-care. Considering widely used models of health and disability such as the ICF model [51, 52], OPAQ, EQ-5D, and SF-6D appear to cover the level of the person completing specific tasks or activities (e.g., walking, bending, or lifting) and the level of the person participating in society (e.g., grocery shopping, using public transportation, going to a movie). However, the content within these levels of functioning is different across the instruments, in the number of questions covering the construct and in the range of functions addressed, with OPAQ appearing to emphasize activity more heavily than participation. Differences of coverage between OPAQ and the preference-based instruments could explain why both models leave a similar level of variation in HSVs unexplained and support the argument for inclusion of a preference-based measure in clinical studies when cost-utility analysis may be of interest.

There are limitations to this study that should be considered. First, although strong linear relationships were evident between the EQ-5D-US and SF-6D HSVs and the OPAQ dimension scores, particularly for EQ-5D-US, the data were skewed to the right. This would have some impact on the accuracy of EQ-5D-US coefficients and therefore the individual HSV estimates. However, this caveat may be tempered by the consideration that the purpose of this model is to estimate mean HSVs for cost-utility analysis and not individual HSVs.

This sample of persons with a history of osteoporosis-related fractures was largely white with HSVs centered at 0.75 for SF-6D and 0.85 for EQ-5D-US, and age centered at 68 years. Our OPAQ and HSV estimates are similar to those reported in the literature. Studies investigating the impact of osteoporosis on quality of life have published various HSV estimates. Although none have provided the link from OPAQ to HSV and only one presents EQ-5D-US HSV as in this study, results of other studies may be useful to relate to ours. Compared to the study of Silverman et al. of prevalent and incident vertebral fractures, the mean OPAQ dimension scores were very similar to the mean OPAQ scores found for our sample [53]. In regard to HSV, in their prospective study, Strom et al. reported lower mean EQ-5D-UK HSV for vertebral fracture and similar mean HSV for hip and wrist fracture groups at 1 year [54]. EQ-5D-UK HSV published from the European MORE and EPOS studies were slightly lower than those in our study [55, 56]. Hagino et al. reported similar mean EQ-5D HSV at 1-year post fracture to our mean EQ-5D HSV by fracture type using Japanese preference weights [57]. In studies of more severely affected patients, the EQ-5D-UK HSVs were lower than the EQ-5D-US HSVs reported here [58–60]. Other studies reporting mean EQ-5D-UK HSVs for

osteoporosis-related health states from community-based samples found slightly higher HSVs overall [61, 62]. For example, van Schoor et al. reported mean EQ-5D-UK HSV of 0.80 (0.69–0.85) for severe osteoporosis based on radiographic evidence. Only one study provided estimates of mean EQ-5D-US HSVs related to osteoporosis for comparison [63]. Sullivan et al. analyzed data from the Medical Expenditure Panel survey and reported on chronic conditions based on “quality priority conditions” coding and “clinical classification categories.” Preference-based scores for osteoporosis and other chronic conditions were included in this study, and the mean EQ-5D-US score for osteoporosis of 0.753 and median of 0.80 were reported. Our mean for those with osteoporotic fracture, at 0.85, is slightly higher. Taking these studies into consideration, our sample appears to be a reasonably representative one for many clinical situations; however, for trials involving severe osteoporosis-related health states or predominantly older persons, the estimates from this model may not be generalizable.

The strength of associations between dimensions of EQ-5D, SF-6D, and OPAQ was as expected and consistent with the nature of the constructs measured. It should be noted that multiple correlations were assessed for statistical significance in this study, which would allow for the detection of statistically significant relationships due to chance alone. However, given the nature of the constructs and the level of statistical significance ($p < 0.001$ level), we believe that these relationships are not likely to be due to chance.

While the subjects in this study were included if they had experienced a fracture within 5 years, recency is not a focus of this project and details on recency of a fracture are not available for analysis. Our aim was to link the specific OPAQ scores with health state values at the same point in time. Therefore, we used a cross-sectional approach, analyzing the OPAQ, EQ-5D, and SF-6D scores representing the same health state in time for that subject.

Conclusions

Estimation of EQ-5D-US and SF-6D health state values from OPAQ using the regression model under investigation appears to be a reasonable approach when actual values are not available and may facilitate cost-effectiveness analyses in trials where OPAQ is used. Future research should test the performance of this model prospectively.

Acknowledgments This study was supported by the National Institutes of Health (R01 AG12262, P60-AR048094, and 1F32HD056763) and a New Investigator Fellowship Training Initiative (NIFTI) in Health Services Research Award from the Foundation for Physical Therapy.

Conflicts of interest None.

References

- Bloom BS (2004) Use of formal benefit/cost evaluations in health system decision making. *Am J Manage Care* 10(5):329–335, see comment
- Gold M, Franks P, Erickson P (1996) Assessing the health of the nation. The predictive validity of a preference-based measure and self-rated health. *Med Care* 34(2):163–177
- Randell AG, Bhalariao N, Nguyen TV, Sambrook PN, Eisman JA, Silverman SL (1998) Quality of life in osteoporosis: reliability, consistency, and validity of the Osteoporosis Assessment Questionnaire. *J Rheumatol* 25(6):1171–1179
- Torrance GW (1986) Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 5(4):1–30
- Torrance GW, Thomas WH, Sackett DL (1972) Utility maximization model for evaluation of health care programs. *Health Serv Res* 7(2):118–133
- Gold MR, Siegel JE, Russell LB, Weinstein MC (1996) Cost-effectiveness in health and medicine. Oxford University Press, New York
- Kaplan RM, Anderson JP (1988) A general health policy model: update and applications. *Health Serv Res* 23(2):203–235
- Brooks R (1996) EuroQol: the current state of play. *Health Policy* 37(1):53–72
- Brazier J, Roberts J, Deverill M (2002) The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 21(2):271–292
- Torrance GW, Furlong W, Feeny D, Boyle M (1995) Multi-attribute preference functions. Health Utilities Index. *Pharmacoeconomics* 7(6):503–520
- Fryback DG, Dasbach EJ, Klein R, Klein BEK, Dorn N, Peterson K et al (1993) The Beaver Dam Health Outcomes Study: initial catalogue of health-state quality factors. *Med Decis Mak* 13(2):89–102
- Hollingsworth W, Deyo RA, Sullivan SD, Emerson SS, Gray DT, Jarvik JG (2002) The practicality and validity of directly elicited and SF-36 derived health state preferences in patients with low back pain. *Health Econ* 11(1):71–85
- Lee TA, Hollingsworth W, Sullivan SD (2003) Comparison of directly elicited preferences to preferences derived from the SF-36 in adults with asthma. *Med Decis Mak* 23:323–334
- Fryback DG, Lawrence WF, Martin PA, Klein R, Klein BE (1997) Predicting quality of well-being scores from the SF-36: results from the Beaver Dam Health Outcomes Study. *Med Decis Mak* 17(1):1–9
- Lundberg L, Johannesson M, Isacson DG, Borgquist L (1999) The relationship between health-state utilities and the SF-12 in a general population. *Med Decis Mak* 19(2):128–140
- Franks P, Lubetkin EI, Gold MR, Tancredi DJ, Jia H (2004) Mapping the SF-12 to the Euroqol EQ-5D Index in a national US sample. *Med Decis Mak* 24:247–254
- Sengupta N, Nichol MB, Wu J, Globe D (2004) Mapping the SF-12 to the HUI3 and VAS in a managed care population. *Med Care* 42(9):927–937
- Sanderson K, Andrews G, Corry J, Lapsley H (2004) Using the effect size to model change in preference values from descriptive health status. *Qual Life Res* 13(7):1255–1264
- Brazier JE, Roberts J (2004) The estimation of a preference-based measure of health from the SF-12. *Med Care* 42(9):851–859
- Gray AM, Rivero-Arias O, Clarke PM (2006) Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. *Med Decis Mak* 26(1):18–29
- Sullivan PW, Ghushchyan V (2006) Mapping the EQ-5D Index from the SF-12: US general population preferences in a nationally representative sample. *Med Decis Mak* 26:401–409
- Lenert LA, Sturley AP, Rapaport MH, Chavez S, Mohr PE, Rupnow M (2004) Public preferences for health states with schizophrenia and a mapping function to estimate utilities from positive and negative symptom scale scores. *Schizophr Res* 71(1):155–165
- Kind P, Macran S (2005) Eliciting social preference weights for Functional Assessment of Cancer Therapy-Lung health states. *Pharmacoeconomics* 23(11):1143–1153
- Brazier J, Kolotkin RL, Crosby RD, Williams GR (2004) Estimating a preference-based single index for the Impact of Weight on Quality of Life-Lite (IWQOL-Lite) instrument from the SF-6D. *Value Health* 7(4):490–498
- Buxton M, Lacey L, Feagan B, Niecko T, Miller D, Townsend R (2007) Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease Questionnaire and Crohn's Disease Activity Index in Crohn's disease and measures of utility. *Value Health* 10(3):214–220
- Erickson P (1998) Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Qual Life Res* 7:101–114
- Gabriel SE, Tosteson AN, Leibson CL, Crowson CS, Pond GR, Hammond CS et al (2002) Direct medical costs attributable to osteoporotic fractures. *Osteoporos Int* 13(4):323–330
- Tosteson AN, Gabriel SE, Grove MR, Moncur MM, Kneeland TS, Melton LJ 3rd et al (2001) Impact of hip and vertebral fractures on quality-adjusted life years. *Osteoporos Int* 12(12):1042–1049
- (1998) Osteoporosis: Review of the evidence for prevention, diagnosis, and treatment and cost-effectiveness analysis. *Osteoporos Int* 8(Suppl 4):S1–S88
- Silverman SL (2000) The Osteoporosis Assessment Questionnaire (OPAQ): a reliable and valid disease-targeted measure of health-related quality of life (HRQOL) in osteoporosis. *Qual Life Res* 9(s1):767–774
- Silverman SL, Mason J, Greenwald M (1993) The Osteoporosis Assessment Questionnaire (OPAQ): a reliable and valid self assessment measure of quality of life in osteoporosis. *J Bone Miner Res* 8(Suppl 1):s343
- Ware JE Jr, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30(6):473–483
- Brazier J, Usherwood T, Harper R, Thomas K (1998) Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol* 51(11):1115–1128
- Shaw JW, Johnson JA, Coons SJ (2005) US valuation of the EQ-5D health states. Development and testing of the D1 valuation model. *Med Care* 43(3):203–220
- Deyo RA, Cherkin DC, Ciol MA (1992) Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 45:613–619
- Romano PS, Roos LL, Jollis JG (1993) Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *J Clin Epidemiol* 46:1075–1079
- Mortimer D, Segal L (2008) Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY weights. *Med Decis Mak* 28(1):66–89
- Revicki D, Kawata A, Harnam N, Chen W, Hays R, Cella D (2009) Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. *Qual Life Res* 18:783–791

39. Bansback N, Marra C, Tsuchiya A, Anis A, Guh D, Hammond T et al (2007) Using the Health Assessment Questionnaire to estimate preference-based single indices in patients with rheumatoid arthritis. *Arthritis Rheum* 57(6):963–971
40. Feeny D (2006) As good as it gets but good enough for which applications? *Med Decis Mak* 26:307–309
41. Kaplan R, Groessl EJ, Sengupta N, Sieber WJ, Ganiats TG (2005) Comparison of measured utility scores and imputed scores from the SF-36 in patients with rheumatoid arthritis. *Med Care* 43(1):79–87
42. Sherbourne C, Unutzer J, Schoenbaum M, Duan N, Lenert L, Sturm R et al (2001) Can utility-weighted health-related quality-of-life estimates capture health effects of quality improvement for depression? *Med Care* 39(11):1246–1259
43. McDonough CM, Grove MR, Tosteson TD, Lurie JD, Hilibrand AS, Tosteson AN (2005) Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among spine patient outcomes research trial (SPORT) participants. *Qual Life Res* 14(5):1321–1332
44. Pickard AS, Wang Z, Walton SM, Lee TA (2005) Are decisions using cost-utility analyses robust to choice of SF-36/SF-12 preference-based algorithm? *Health Qual Life Outcomes* 3(1):11
45. Thomas KJ, MacPherson H, Ratcliffe J et al (2005) Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain. *Health Technol Assess* 9(32):1–109
46. Wee H-L, Machin D, Loke W-C, Li S-C, Cheung Y-B, Luo N et al (2007) Assessing differences in utility scores: a comparison of four widely used preference-based instruments. *Value Health* 10(4):256–265
47. Burns AW, Bourne RB, Chesworth BM, MacDonald SJ, Rorabeck CH (2006) Cost effectiveness of revision total knee arthroplasty. *Clin Orthop Relat Res* 446:29–33
48. McDonough C, ANA T (2007) Measuring preferences for cost-utility analysis: how choice of method may influence decision-making. *Pharmacoeconomics* 25(2):93–106
49. Stiggelbout AM (2006) Health state classification systems: how comparable are our ratios? *Med Decis Mak* 25:223–225
50. Fryback DG, Palta M, Cherepanov D, Bolt D, Kim J-S (2010) Comparison of 5 health-related quality of life indexes using item response theory analysis. *Med Decis Mak* 30:5–15
51. Lima VD, Kopec JA (2005) Quantifying the effect of health status on health care utilization using a preference-based health measure. *Soc Sci Med* 60(3):515–524
52. World Health Organization (2002) Towards a common language for functioning, disability, and health ICF: the International Classification of Function, Disability and Health. World Health Organization, Geneva
53. Silverman S, Minshall ME, Shen W, Harper KD, Xie S (2001) The relationship of health-related quality of life to prevalent and incident vertebral fractures in postmenopausal women with osteoporosis. Results from the Multiple Outcomes of Raloxifene Evaluation Study. *Arthritis Rheum* 44(11):2611–2619
54. Strom O, Borgstrom F, Zethraeus N, Johnell O, Lidgren L, Ponzer S et al (2008) Long-term cost and effect on quality of life of osteoporosis-related fractures in Sweden. *Acta Orthop* 79(2):269–280
55. Oleksik A, Lips P, Dawson A, Minshall ME, Shen W, Cooper C et al (2000) Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *J Bone Miner Res* 15(7):1384–1392
56. van Schoor NM, Ewing SK, O'Neill TW, Lunt M, Smit JH, Lips P et al (2008) Impact of prevalent and incident vertebral fractures on utility: results from a patient-based and a population-based sample. *Qual Life Res* 17(1):159–167
57. Hagino H, Nakamura T, Fujiwara S, Oeki M, Okano T, Teshima R et al (2009) Sequential change in quality of life for patients with incident clinical fractures: a prospective study. *Osteoporos Int* 20(5):695–702
58. Cooper C, Jakob F, Chinn C, Martin-Mola E, Fardellone P, Adami S et al (2008) Fracture incidence and changes in quality of life in women with an inadequate clinical outcome from osteoporosis therapy: the Observational Study of Severe Osteoporosis (OSSO). *Osteoporos Int* 19(4):493–501
59. Jakob F, Marin F, Martin-Mola E, Torgerson D, Fardellone P, Adami S et al (2006) Characterization of patients with an inadequate clinical outcome from osteoporosis therapy: the Observational Study of Severe Osteoporosis (OSSO). *QJM* 99(8):531–543
60. Rajzbaum G, Jakob F, Karras D, Ljunggren O, Lems WF, Langdahl BL et al (2008) Characterization of patients in the European Forsteo Observational Study (EFOS): postmenopausal women entering teriparatide treatment in a community setting. *Curr Med Res Opin* 24(2):377–384
61. Cockerill W, Lunt M, Silman AJ, Cooper C, Lips P, Bhalla AK et al (2004) Health-related quality of life and radiographic vertebral fracture. *Osteoporos Int* 15(2):113–119
62. van Schoor NM, Smit JH, Twisk JW, Lips P, van Schoor NM, Smit JH et al (2005) Impact of vertebral deformities, osteoarthritis, and other chronic diseases on quality of life: a population-based study. *Osteoporos Int* 16(7):749–756
63. Sullivan PW, Lawrence WF, Ghushchyan V, Sullivan PW, Lawrence WF, Ghushchyan V (2005) A national catalog of preference-based scores for chronic conditions in the United States. *Med Care* 43(7):736–749