

# Preference-Based EQ-5D Index Scores for Chronic Conditions in the United States

Patrick W. Sullivan, PhD, Vahram Ghushchyan, PhD

**Background.** The Panel on Cost-Effectiveness in Health and Medicine has called for an “off-the-shelf” catalogue of nationally representative, community-based preference scores for health states, illnesses, and conditions. A previous review of cost-effectiveness analyses found that 77% did not incorporate community-based preferences, and 33% used arbitrary expert or author judgment. These results highlight the necessity of making a wide array of appropriate, community-based estimates more accessible to cost-effectiveness researchers. **Objective.** To provide nationally representative EQ-5D index scores for chronic ICD-9 codes. **Methods.** The nationally representative Medical Expenditure Panel Survey (MEPS) was pooled (2000–2002) to create a data set of 38,678 adults. Ordinary least squares (OLS), Tobit, and censored least absolute deviations (CLAD) regression methods were used to estimate the marginal disutility of each condition, controlling for age, comorbidity, gender, race, ethnicity, income, and education.

**Results.** Most chronic conditions, age, comorbidity, income, and education were highly statistically significant predictors of EQ-5D index scores. Homoskedasticity and normality assumptions were rejected, suggesting only CLAD estimates are theoretically unbiased. The magnitude and statistical significance of coefficients varied by analytic method. OLS and Tobit coefficients were on average 60% and 143% greater than CLAD, respectively. The marginal disutility of 95 chronic ICD-9 codes as well as unadjusted mean, median, and 25th and 75th percentiles are reported. **Conclusion.** This research provides nationally representative, community-based EQ-5D index scores associated with a wide variety of chronic ICD-9 codes that can be used to estimate quality-adjusted life-years in cost-effectiveness analyses. **Key words:** health-related quality of life; cost-utility analysis; cost-effectiveness analysis; utility; chronic disease; ICD-9; econometric methods. (*Med Decis Making* 2006; 26:410–420)

## BACKGROUND

Advancements in medical technology within the reality of limited societal resources have led to the need to assess the value of competing interventions

through cost-effectiveness analysis. The quality-adjusted life-year (QALY) method incorporates both survival and the impact of health-related quality of life (HRQL) associated with different health states into a single index, thereby facilitating comprehensive assessment of the effectiveness of an intervention.<sup>1,2</sup> Preference weights are used in QALY estimation to adjust for the HRQL impact of different health states.

In 1993, the US Public Health Service convened the Panel on Cost-Effectiveness in Health and Medicine (PCEHM) to standardize the field of cost-effectiveness analysis. The PCEHM recommended the use of community-based, nationally representative preferences for use in cost-effectiveness analyses.<sup>1</sup> In practice, however, adherence to this recommendation is poor. A review of 228 cost-effectiveness analyses found that 77% did not incorporate community-based preferences, and 33% used arbitrary expert or author judgment.<sup>3</sup> These results highlight the necessity of making valid community-based estimates more accessible.

Received 10 March 2005 from the University of Colorado Department of Economics, Boulder, and University of Colorado School of Pharmacy, Pharmaceutical Outcomes Research Program. This research was done without external funding. A preliminary version of this research was presented as a podium at the Society for Medical Decision Making in October 2004. The authors thank Bill Lawrence and an anonymous reviewer for helpful comments on this manuscript.

Revision accepted for publication 5 May 2006.

Address correspondence to Patrick Sullivan, PhD, University of Colorado School of Pharmacy, Pharmaceutical Outcomes Research Program, 4200 East Ninth Avenue, Box C238, Denver, CO 80262; telephone: (303) 315-1560; fax: (303) 315-1797; e-mail: Patrick.Sullivan@UCHSC.edu.

DOI: 10.1177/0272989X06290495

To accomplish this goal, the PCEHM recommended the collection of a national catalogue of preference weights that could be used by cost-effectiveness researchers without the burden of primary data collection. Such a repository would provide “off-the-shelf” access to a wide number of disease and condition-specific preference weights and would encourage comparisons across different studies and health interventions. The PCEHM further recommended that any such catalogue use a theory-based method to derive empirical weights for health states, illnesses, and conditions from a nationally representative, community-based sample of the US population.<sup>1</sup>

Previous systems offer important insight and improve our understanding of the sources of variability in preference weights but do not meet the criteria of the PCEHM. Before the PCEHM was convened, Fryback et al.<sup>4</sup> measured preferences using the time tradeoff (TTO) method and the Quality of Well-being (QWB) scale for 1356 noninstitutionalized residents of Beaver Dam, Wisconsin. Although a seminal contribution to the understanding of condition-specific preferences, the Beaver Dam sample was not nationally representative. Gold et al.<sup>5</sup> developed the Health and Limitations Index (HALex), an HRQL scoring system capable of defining 30 possible health states to provide condition-level weights for the US population. Time and resource constraints prevented the National Center for Health Statistics (NCHS) from directly eliciting empirical community-based preferences for the HALex. HALex scores are not based on empirical weights from a nationally representative community sample. Instead, scores were derived by correspondence analysis of value scores on 2 questions (attributes) in the National Health Interview Survey (NHIS).<sup>6</sup> Tengs et al.<sup>7</sup> and Bell et al.<sup>3</sup> responded to the PCEHM call by collecting numerous preference estimates that had been used in previous studies. The estimates published by Tengs et al. and Bell et al. underscore the eclectic variability in preference weights and highlight one of the underlying motivations for the PCEHM’s call for a national catalogue of standardized preference weights.

More recently, Sullivan et al.<sup>8</sup> developed a preliminary catalogue of preference-based HRQL scores that meets all criteria of the PCEHM. The authors provide EQ-5D index scores of chronic conditions as classified by the Agency for Healthcare Research and Quality (AHRQ) for 58 clinical classification categories (CCCs) and 10 quality priority conditions (QPCs). CCCs are broad clinical groupings of several ICD-9 codes.<sup>9</sup> QPCs are common, serious conditions for

which standards for appropriate clinical care have been developed.<sup>10</sup>

Cost-effectiveness analyses address a wider variety of conditions, illnesses, and health states than those represented in this preliminary catalogue. The most common classification of diseases and conditions used by health researchers and clinicians is the International Classification of Disease (ICD).<sup>11</sup> ICD codes provide more refined and clinically homogeneous classification of distinct conditions. Preference-based HRQL scores associated with ICD-9 codes may be appropriate when QPC or CCC codes are too generic for the desired clinical condition. For example, ICD-9 413 “Angina Pectoris” is 1 of 5 distinct 3-digit ICD-9 codes that comprise CCC 101 “Coronary Atherosclerosis and Other Heart Disease.” In addition, the provision of preference-based scores associated with ICD-9 codes may correspond more directly to available cost data for use in cost-effectiveness models because medical billing data, and hence cost estimates, are often categorized by ICD-9 code.

The purpose of this research is to augment the preliminary national catalogue of preference-based HRQL scores for QPC and CCC codes developed by Sullivan et al.<sup>8</sup> by providing EQ-5D index scores associated with a variety of chronic ICD-9 codes.

## METHODS

### Data Sources

#### *Medical Expenditure Panel Survey*

The Medical Expenditure Panel Survey (MEPS) is a nationally representative survey of the US civilian noninstitutionalized population with oversampling of Hispanics and blacks, collecting detailed information on demographic characteristics, health conditions, and health status.<sup>10</sup> Medical condition diagnoses in MEPS are based on ICD-9 codes. The current research pooled 2000–2002 MEPS person-level and medical conditions files to construct an individual-level file with information on ICD-9 codes reported by respondents.

#### *EQ-5D*

The EQ-5D consists of a 5-item descriptive system that measures 5 dimensions of health status (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with 3 levels per dimension (no problem, some problems, and extreme problems). The combination of all possible dimensions and levels results in 243 unique health states. A multiattribute value function (MAVF) is used to map preferences for

these health states. From this scoring algorithm, EQ-5D index scores are calculated based on responses to the 5-item questionnaire. The EQ-5D has been used as an HRQL and preference measure in a wide variety of diseases and conditions in more than 600 publications.<sup>12</sup> In addition, it has been used as a general population health measure in numerous countries.<sup>12</sup> The construct validity, reliability, and responsiveness of the EQ-5D have been documented extensively in both general and specific disease populations.<sup>12,13</sup> The scoring algorithm for the EQ-5D index descriptive system used in this research is based on US community preferences.<sup>14</sup>

For the current study, the EQ-5D was administered via a paper-and-pencil self-administered questionnaire (SAQ) given to adults 18 years or greater. The questionnaire was completed by a proxy for individuals unable to respond. There were 38,678 unique individuals 18 years of age or greater with valid EQ-5D scores included in this analysis. Individuals may have reported more than 1 condition. The sample design of the MEPS-HC survey includes stratification, clustering, multiple stages of selection, and oversampling of minority populations.<sup>15</sup> Using the MEPS sampling weights, we adjusted for these factors and for survey nonresponse. Conditions with fewer than 100 observations were excluded from the reported results.

### Chronic Conditions

A total of 693 three-digit ICD-9 codes were mapped by professional coders from the medical conditions reported by survey respondents.<sup>10,11</sup> We have included only chronic (lasting >1 year) conditions to ensure that the condition was experienced while the EQ-5D was administered via the SAQ. (The SAQ was administered during rounds 2 and 4 of the MEPS survey. MEPS data do not provide detailed information about the exact timing and duration of acute conditions and thus cannot be correlated with accuracy to the EQ-5D elicitation time.) Hwang et al.<sup>16</sup> convened 2 physician panels to determine which ICD-9 codes reported in MEPS should be classified as chronic. This list was used to determine which ICD-9 codes could be classified as chronic conditions.

### Estimating Preference-Based HRQL Scores

Preference-based HRQL scores exhibit a ceiling effect with a significant number of respondents rating themselves in full health.<sup>17</sup> The ceiling effect is particularly evident with the EQ-5D index. As discussed elsewhere,<sup>8</sup> in these circumstances, the censored least

absolute deviations (CLAD) estimator developed by Powell et al.<sup>18</sup> has theoretical advantages over ordinary least squares (OLS) regression and the Tobit model, which may produce biased estimates when the assumptions of heteroskedasticity or nonnormality are violated.<sup>18–20</sup> To test for heteroskedasticity in the Tobit model, we used the likelihood ratio test described by Petersen and Waldman.<sup>21</sup> This statistic has a limiting chi-squared distribution. In addition, the normality assumption was examined using the Hausman test.<sup>20</sup>

All regression analyses were done using STATA. CLAD was performed in STATA using an existing add-in (user-written program)<sup>22</sup> and additional programming to incorporate sample weights. Programs were written in STATA to perform the Hausman and the likelihood ratio tests described above. The determination of model specification was guided primarily by the need to be consistent with the specification adopted in the preliminary catalogue.<sup>8</sup>

## RESULTS

Unadjusted mean and median EQ-5D index scores, mean age, and number of chronic conditions reported for the US population in 2000–2002 are shown in Table 1. Table 1 also highlights characteristics that influence EQ-5D index scores. For example, whites are older and report more chronic conditions (possibly due to greater age) than other races. Mean EQ-5D index scores decline by increasing age category, are lower for females than males, are lower for blacks and American Indians compared to whites, are higher for other races and Hispanics compared to whites, and are consistently lower for lower levels of educational attainment and poverty status compared to higher levels.

The homoskedasticity assumption (likelihood ratio test = 401.2;  $P < 0.001$ ) and the normality assumption (Hausman test statistic = 2015;  $P < 0.001$ ) were rejected for the classic Tobit model, suggesting that Tobit estimates may be biased. As with many health status measures in population health surveys, it appears that the EQ-5D index is not normally distributed and exhibits a significant ceiling effect (46% reported EQ-5D index = 1.0 in this sample). Given these factors, only CLAD estimates are theoretically unbiased.

Full results of the CLAD, Tobit, and OLS regressions for the MEPS general population without disease indicator variables are displayed in Table 2. Age, the number of chronic conditions (NCC), NCC<sup>2</sup>, greater levels of income (v. poor), and greater levels of education (v. no degree) were statistically significant in explaining the variation in EQ-5D index scores

**Table 1** EQ-5D index Scores of US Adults by Sociodemographic Characteristic in the 2000–2002 MEPS

	Number	%	Mean Age (yrs)	Mean NCC	Median EQ-5D	Mean EQ-5D	Standard Error EQ-5D
Age group							
18–29	7990	21	24	0.7	1.000	0.922	0.0019
30–39	7867	20	35	1.0	1.000	0.901	0.0021
40–49	8052	21	44	1.6	0.844	0.871	0.0024
50–59	6242	16	54	2.4	0.827	0.842	0.0028
60–69	3980	10	64	3.2	0.827	0.823	0.0034
70–79	2981	8	74	4.0	0.810	0.790	0.0036
≥80	1566	4	84	4.3	0.778	0.736	0.0062
Gender							
Female	20980	52	46	2.2	0.844	0.854	0.0017
Male	17698	48	45	1.5	1.000	0.880	0.0015
Race							
White	31413	83	46	2.0	0.844	0.867	0.0014
Black	5326	11	43	1.5	0.844	0.858	0.0037
American Indian	367	1	42	1.9	0.844	0.835	0.0133
Other	1572	4	42	1.4	1.000	0.892	0.0048
Ethnicity							
Hispanic	8183	11	39	1.2	1.000	0.877	0.0030
Education							
No degree	9299	18	47	2.1	0.827	0.818	
High school diploma	19211	51	45	1.9	0.844	0.863	0.0030
Other degree	2473	7	45	1.9	0.854	0.883	0.0015
Bachelor's degree	5063	16	44	1.7	1.000	0.907	0.0031
Master's or doctorate degree	2481	8	49	1.9	1.000	0.908	0.0021
Family income (as % of poverty line)							0.0031
Poor/negative (<100%)	5193	10	44	2.1	0.827	0.789	
Near poor (100%–124%)	1894	4	48	2.4	0.827	0.801	0.0045
Low income (125%–199%)	5875	13	47	2.1	0.827	0.836	0.0065
Middle income (200%–399%)	12262	32	45	1.8	0.844	0.872	0.0035
High income (≥400%)	13454	42	46	1.8	1.000	0.896	0.0017

MEPS, Medical Expenditure Panel Survey; NCC, number of chronic conditions.

across all 3 methods. Greater comorbidity burden and higher levels of income had a relatively large magnitude impact on EQ-5D index scores.

The relative percentage differences in the coefficients for each analytic method are also displayed in Table 2. It is evident that point estimates vary significantly by choice of analytic method. Coefficients varied by up to 382% for Tobit estimates of bachelor's degree (0.0823) compared to CLAD (0.0171) and 191% for OLS (0.0497) v. CLAD. On average, statistically significant OLS coefficients were 60% greater in magnitude than CLAD, and Tobit estimates were 143% greater than CLAD and 69% greater than OLS.

Table 3 presents EQ-5D index scores for chronic ICD-9 codes resulting from CLAD regression, as well as unadjusted mean, median, and 25th and 75th percentiles. The number of individuals reporting the condition, mean age, and NCC by percentile are also presented. Estimates of the marginal disutility of each

condition after adjusting for age, gender, race, ethnicity, education, income, and comorbidity, are reported for each condition. The complete set of regression coefficients for all ICD-9 codes is available online at the journal's Web site, [mdm.sagepub.com](http://mdm.sagepub.com).

## DISCUSSION

This research augments the preliminary national catalogue of preference-based scores developed by Sullivan et al.<sup>8</sup> by providing EQ-5D index scores associated with a wide range of chronic ICD-9 codes. The preliminary catalogue of chronic conditions was categorized by QPC and CCC codes.<sup>8</sup> The current research adds to the existing catalogue and provides analysts with a standardized and flexible source of nationally representative preference-based scores associated with a variety of chronic ICD-9 codes that can be used in QALY estimation for cost-effectiveness analysis.

**Table 2** Results of the CLAD, Tobit, and OLS Multivariate Regression and Comparative Percentage Differences in the 2000–2002 MEPS

Variable	CLAD Coefficient	CLAD Standard Error	Tobit Coefficient	Tobit v. CLAD (%)	Tobit v. OLS (%)	Tobit Standard Error	OLS Coefficient	OLS v. CLAD (%)	OLS Standard Error
Age	−0.0007*	0.0000	−0.0016*	137	97	0.0001	−0.0008*	20	0.0001
Male	0.0007	0.0008	−0.0011	−271	−70	0.0030	−0.0037***	−669	0.0017
Hispanic	0.0052*	0.0012	0.013***	148	132	0.0051	0.0056	7	0.0029
NCC <sup>a</sup>	−0.0546*	0.0004	−0.0696*	27	84	0.0020	−0.0379*	−31	0.0010
NCC <sup>2a</sup>	0.0031*	0.0000	0.003*	−4	202	0.0002	0.001*	−68	0.0001
Black	−0.0006	0.0013	−0.01***	1573	14	0.0052	−0.0088**	1365	0.0031
American Indian	−0.0043	0.0043	−0.0252	485	44	0.0164	−0.0175	307	0.0112
Other	−0.0007	0.0021	0.0013	−304	−149	0.0089	−0.0027	314	0.0041
Near poor (100%–124%)	0.0123*	0.0023	0.0286**	133	37	0.0092	0.0208**	70	0.0067
Low income (125%–199%)	0.05*	0.0016	0.0625*	25	38	0.0069	0.0452*	−10	0.0047
Middle income (200%–399%)	0.0605*	0.0014	0.0905*	50	39	0.0060	0.0651*	8	0.0043
High income (≥400%)	0.0638*	0.0015	0.1203*	88	47	0.0058	0.0817*	28	0.0042
High school diploma	0.0085*	0.0011	0.0306*	258	28	0.0042	0.0239*	180	0.0027
Other degree	0.0144*	0.0019	0.0541*	275	46	0.0069	0.0369*	156	0.0036
Bachelor's degree	0.0171*	0.0015	0.0823*	382	66	0.0062	0.0497*	191	0.0032
Master's degree or doctorate	0.0236*	0.0019	0.1022*	333	67	0.0071	0.0611*	159	0.0034
Constant	0.9442*	0.0020	1.0036*	6	14	0.0088	0.8791*	−7	0.0054
Pseudo- <i>R</i> <sup>2</sup>	0.19	—	0.35	—	—	—	0.27	—	—

Regression of EQ-5D = constant + Age + NCC + NCC<sup>2</sup> + Male + AmericanIndian + OtherRace + Black + Hispanic + Income2 + Income3 + Income4 + Income5 + Degree2 + Degree3 + Degree4 + Degree5. Full regression coefficients are available in an online appendix.<sup>28</sup> MEPS, Medical Expenditure Panel Survey; OLS, ordinary least squares; CLAD, censored least absolute deviations; NCC, number of chronic conditions.

a. NCC: the total number of chronic ICD-9 codes reported.

\**P* < 0.001. \*\**P* < 0.01. \*\*\**P* < 0.05.

Availability of nationally representative, community-based preference scores overcomes one of the most important barriers to appropriate estimation of cost-effectiveness.<sup>1</sup> An “off-the-shelf” catalogue of preference-based scores derived from a theory-based method provides a standardized and consistent source of preference weights.<sup>1</sup> The preference-based scores for chronic ICD-9 codes reported in this research are based on the EQ-5D, a theory-based, previously validated health status classification system that meets all of the criteria established by the PCEHM. Specific examples of how to incorporate these estimates in cost-effectiveness analyses are provided elsewhere (see the online appendix at the journal's Web site, [mdm.sagepub.com](http://mdm.sagepub.com)).

The current list of “off-the-shelf” EQ-5D index scores for ICD-9 codes differs from the previous catalogue for CCC and QPCs. First, the ICD-9 codes in this research provide more refined classification of distinct conditions and may be more appropriate if CCC codes are too generic for the desired clinical conditions. For

example, CCC 089 “Blindness and Other Vision Defects” had no associated disutility in previous research,<sup>8</sup> which *prima facie* contradicts clinical expectation. However, CCC 089 is a combination of 4 distinct ICD-9 codes (367 “Disorders of Refraction,” 368 “Visual Disturbance,” 369 “Blindness and Low Vision,” and V41 “Problems with Specific Functions”). In comparison, ICD-9 369 “Blindness and Low Vision” provides a more clinically homogeneous condition classification and has an associated disutility of −0.0498 (Table 3). Hence, while CCC 089 and ICD-9 369 appear similar in name, there is a significant difference in the specificity of clinical categorization. We recommend that analysts familiarize themselves with the ICD-9 to CCC crosswalk to determine the appropriateness of a given condition categorization.<sup>23</sup> In addition, ICD-9 codes may correspond more practically to available cost data for use in cost-effectiveness models. Estimation of the cost of a given condition in analyses of existing data (i.e., claims data) is often based on ICD-9 codes. In this case, the provision of EQ-5D index scores

**Table 3** EQ-5D Index Scores, Age, and Comorbidity Burden by ICD-9 in the 2000–2002 MEPS

ICD-9 Classification	n	Mean Age	Unadjusted					Regression Results <sup>a</sup>				
			NCC <sup>b</sup> 25%	NCC <sup>b</sup> 50%	NCC <sup>b</sup> 75%	Mean EQ-5D	EQ-5D 25%	EQ-5D 50%	EQ-5D 75%	Disutility of Condition <sup>a</sup>	Condition <sup>a</sup> Standard Error	Statistical Significance (Condition)
ICD-9 173 Other Malign Neopl Skin	453	67	3	4	6	0.812	0.768	0.816	1.000	-0.0137	0.0001	*
ICD-9 174 Malign Neo Female Breast	201	64	3	4	6	0.803	0.708	0.816	1.000	-0.0156	0.0002	*
ICD-9 185 Malign Neopl Prostate	160	70	3	4	6	0.772	0.708	0.810	0.860	-0.0153	0.0002	*
ICD-9 216 Benign Neoplasm of Skin	443	49	2	3	5	0.861	0.800	0.827	1.000	-0.0002	0.0001	*
ICD-9 239 Unspecified Neoplasm	389	58	2	4	6	0.795	0.778	0.827	1.000	-0.0174	0.0001	*
ICD-9 242 Thyrotoxicosis	188	54	2	4	6	0.800	0.761	0.827	1.000	-0.0131	0.0002	*
ICD-9 244 Acquired Hypothyroidism	789	57	3	4	6	0.821	0.778	0.827	1.000	-0.0005	0.0001	*
ICD-9 246 Oth Disorders of Thyroid	1035	61	2	4	6	0.783	0.748	0.810	0.844	-0.0224	0.0001	*
ICD-9 250 Diabetes Mellitus	2778	60	3	4	6	0.751	0.708	0.800	0.844	-0.0351	0.0001	*
ICD-9 259 Oth Endocrine Disorders	552	55	2	4	6	0.806	0.778	0.827	1.000	-0.0168	0.0001	*
ICD-9 272 Dis of Lipoid Metabolism	3433	61	3	4	6	0.810	0.778	0.827	1.000	-0.0049	0.0001	*
ICD-9 274 Gout	412	62	2	4	6	0.771	0.708	0.810	0.844	-0.0221	0.0001	*
ICD-9 278 Obesity/Hyperalimient	131	47	2	4	6	0.744	0.708	0.800	0.844	-0.0500	0.0002	*
ICD-9 285 Anemia Nec/Nos	407	55	3	5	7	0.753	0.708	0.800	0.844	-0.0209	0.0001	*
ICD-9 289 Other Blood Disease	130	64	3	4	6	0.749	0.708	0.800	0.827	-0.0359	0.0002	*
ICD-9 296 Affective Psychoses	247	42	2	3	6	0.693	0.506	0.778	0.844	-0.0897	0.0002	*
ICD-9 300 Neurotic Disorders	1896	49	2	4	6	0.739	0.708	0.800	0.844	-0.0494	0.0001	*
ICD-9 309 Adjustment Reaction	189	48	2	4	6	0.739	0.597	0.800	0.844	-0.0337	0.0002	*
ICD-9 311 Depressive Disorder Nec	3086	48	2	3	5	0.732	0.708	0.800	0.844	-0.0625	0.0001	*
ICD-9 345 Epilepsy	120	45	2	3	5	0.766	0.708	0.827	1.000	-0.0078	0.0002	*
ICD-9 346 Migraine	980	41	2	3	5	0.806	0.778	0.827	1.000	-0.0297	0.0001	*
ICD-9 354 Mononeuritis Upper Limb	351	50	2	3	6	0.771	0.768	0.816	0.827	-0.0379	0.0001	*
ICD-9 355 Mononeuritis Leg	401	53	2	4	6	0.751	0.708	0.800	0.827	-0.0244	0.0001	*
ICD-9 360 Disorders of the Globe	282	51	2	3	5	0.828	0.778	0.827	1.000	-0.0064	0.0001	*
ICD-9 362 Retinal Disorders Nec	223	72	3	5	7	0.792	0.708	0.816	0.844	-0.0156	0.0002	*
ICD-9 365 Glaucoma	634	69	3	5	7	0.782	0.708	0.810	0.860	-0.0130	0.0001	*
ICD-9 366 Cataract	766	73	3	5	7	0.770	0.708	0.810	0.844	-0.0171	0.0001	*
ICD-9 368 Visual Disturbances	175	53	2	4	6	0.810	0.778	0.827	1.000	-0.0046	0.0002	*
ICD-9 369 Blindness and Low Vision	159	61	3	5	7	0.694	0.543	0.778	0.827	-0.0498	0.0002	*
ICD-9 375 Lacrimal System Disorder	127	63	3	5	7	0.808	0.778	0.827	1.000	-0.0002	0.0002	*
ICD-9 379 Eye Disorders Nec	393	57	2	4	6	0.805	0.768	0.827	1.000	-0.0081	0.0001	*
ICD-9 389 Hearing Loss	320	69	4	5	7	0.776	0.708	0.800	0.860	-0.0060	0.0001	*
ICD-9 401 Essential Hypertension	7085	61	2	4	5	0.789	0.761	0.816	1.000	-0.0250	0.0000	*
ICD-9 410 Acute Myocardial Infarct	244	62	3	5	7	0.704	0.575	0.778	0.843	-0.0409	0.0002	*
ICD-9 413 Angina Pectoris	228	69	4	6	9	0.695	0.517	0.768	0.827	-0.0412	0.0002	*
ICD-9 414 Oth Chr Ischemic Hrt Dis	183	66	4	6	8	0.738	0.708	0.794	0.827	-0.0336	0.0002	*
ICD-9 424 Oth Endocardial Disease	214	56	2	4	6	0.789	0.708	0.816	1.000	-0.0059	0.0002	*
ICD-9 427 Cardiac Dysrhythmias	649	67	3	5	7	0.774	0.708	0.810	0.843	-0.0190	0.0001	*
ICD-9 428 Heart Failure	284	71	5	6	9	0.636	0.437	0.708	0.810	-0.0635	0.0002	*
ICD-9 429 Ill-Defined Heart Dis	1204	69	3	5	7	0.716	0.619	0.778	0.827	-0.0492	0.0001	*
ICD-9 436 CVA	340	68	3	5	7	0.650	0.463	0.768	0.816	-0.0524	0.0001	*

(continued)

Table 3 (continued)

ICD-9 Classification	n	Mean Age	Unadjusted					Regression Results <sup>a</sup>				
			NCC <sup>b</sup> 25%	NCC <sup>b</sup> 50%	NCC <sup>b</sup> 75%	Mean EQ-5D	EQ-5D 25%	EQ-5D 50%	EQ-5D 75%	Disutility of Condition <sup>a</sup>	Condition <sup>a</sup> Standard Error	Statistical Significance (Condition)
ICD-9 436 CVA	340	68	3	5	7	0.650	0.463	0.768	0.816	-0.0524	0.0001	*
ICD-9 444 Arterial Embolism	201	66	3	5	7	0.734	0.708	0.778	0.827	-0.0198	0.0002	*
ICD-9 453 Oth Venous Thrombosis	100	65	3	5	7	0.727	0.597	0.778	0.827	-0.0380	0.0002	*
ICD-9 455 Hemorrhoids	246	51	2	3	6	0.826	0.800	0.827	1.000	-0.0008	0.0002	*
ICD-9 459 Oth Circulatory Disease	367	64	3	5	7	0.732	0.708	0.778	0.827	-0.0376	0.0001	*
ICD-9 473 Chronic Sinusitis	2047	45	2	3	5	0.853	0.800	0.827	1.000	-0.0022	0.0001	*
ICD-9 477 Allergic Rhinitis	3326	46	2	3	5	0.853	0.800	0.827	1.000	-0.0009	0.0001	*
ICD-9 492 Emphysema	296	68	3	5	7	0.663	0.508	0.708	0.810	-0.0667	0.0001	*
ICD-9 493 Asthma	1604	46	2	3	5	0.797	0.761	0.827	1.000	-0.0213	0.0001	*
ICD-9 518 Other Lung Diseases	167	59	3	5	7	0.714	0.597	0.778	0.827	-0.0428	0.0002	*
ICD-9 519 Oth Resp System Diseases	284	52	2	3	5	0.816	0.768	0.827	1.000	-0.0277	0.0001	*
ICD-9 525 Other Dental Disorder	567	43	1	2	4	0.853	0.800	0.833	1.000	-0.0030	0.0001	*
ICD-9 530 Diseases of Esophagus	1500	55	3	4	6	0.785	0.761	0.816	0.844	-0.0216	0.0001	*
ICD-9 531 Gastric Ulcer	244	57	3	5	7	0.727	0.597	0.800	0.827	-0.0269	0.0002	*
ICD-9 535 Gastritis and Duodenitis	224	49	2	4	7	0.765	0.708	0.800	0.844	-0.0116	0.0002	*
ICD-9 536 Stomach Function Disord	1211	53	2	4	6	0.773	0.708	0.810	1.000	-0.0228	0.0001	*
ICD-9 537 Oth Gastroduodenal Dis	392	56	2	4	6	0.750	0.708	0.800	0.843	-0.0311	0.0001	*
ICD-9 553 Other Abdominal Hernia	480	57	2	4	6	0.770	0.708	0.810	0.827	-0.0326	0.0001	*
ICD-9 558 Oth Noninf Gastroenterit	117	53	2	4	6	0.777	0.768	0.827	0.844	-0.0429	0.0002	*
ICD-9 562 Diverticula of Intestine	195	64	4	5	7	0.735	0.708	0.778	0.827	-0.0302	0.0002	*
ICD-9 564 Funct Digestive Dis Nec	396	56	3	5	7	0.728	0.708	0.800	0.833	-0.0346	0.0001	*
ICD-9 569 Oth Intestinal Disorders	254	57	2	4	6	0.794	0.761	0.816	1.000	-0.0193	0.0002	*
ICD-9 573 Oth Liver Disorders	113	52	2	4	6	0.729	0.597	0.778	0.827	-0.0567	0.0002	*
ICD-9 574 Cholelithiasis	148	52	3	4	6	0.782	0.708	0.816	1.000	-0.0288	0.0002	*
ICD-9 575 Oth Gallbladder Disorder	216	51	2	4	6	0.793	0.761	0.827	1.000	-0.0063	0.0002	*
ICD-9 586 Renal Failure Nos	102	58	4	6	8	0.651	0.506	0.708	0.816	-0.0603	0.0002	*
ICD-9 592 Renal/Ureteral Calculus	299	49	2	3	5	0.832	0.800	0.827	1.000	-0.0159	0.0001	*
ICD-9 593 Oth Renal & Ureteral Dis	229	60	3	5	8	0.712	0.708	0.778	0.827	-0.0527	0.0002	*
ICD-9 596 Other Bladder Disorders	259	62	4	6	7	0.706	0.597	0.778	0.827	-0.0430	0.0002	*
ICD-9 599 Oth Urinary Tract Disor	761	49	2	3	5	0.826	0.800	0.827	1.000	-0.0093	0.0001	*
ICD-9 602 Oth Prostatic Disorders	315	69	3	4	6	0.788	0.759	0.810	1.000	-0.0163	0.0001	*
ICD-9 610 Benign Mammary Dysplasia	143	51	2	3	5	0.852	0.810	0.827	1.000	-0.0107	0.0002	*
ICD-9 611 Other Breast Disorders	407	49	2	3	5	0.838	0.800	0.827	1.000	-0.0008	0.0001	*
ICD-9 620 Noninflam Dis Ovar/Adnex	204	35	1	2	4	0.859	0.800	0.827	1.000	-0.0044	0.0002	*
ICD-9 623 Noninflam Dis Vagina	116	49	2	4	6	0.822	0.778	0.827	0.860	-0.0285	0.0002	*
ICD-9 626 Disorder of Menstruation	317	37	2	3	4	0.872	0.800	0.844	1.000	-0.0005	0.0001	*
ICD-9 627 Menopausal Disorders	1003	56	2	4	6	0.817	0.778	0.827	1.000	-0.0061	0.0001	*
ICD-9 696 Psoriasis/Like Disorders	176	51	2	3	5	0.834	0.778	0.827	1.000	-0.0150	0.0002	*
ICD-9 698 Pruritus & Like Cond	106	60	3	6	7	0.730	0.527	0.800	0.844	-0.0304	0.0002	*
ICD-9 700 Corns and Callosities	153	64	3	4	7	0.789	0.761	0.800	0.844	-0.0214	0.0002	*

ICD-9 703 Diseases of Nail	291	59	3	5	6	0.775	0.708	0.810	0.843	-0.0211	0.0001	*
ICD-9 707 Chronic Ulcer of Skin	318	56	3	5	7	0.694	0.597	0.778	0.827	-0.0272	0.0001	*
ICD-9 710 Diff Connective Tiss Dis	111	51	3	5	7	0.742	0.708	0.800	0.827	-0.0235	0.0002	*
ICD-9 714 Oth Inflamm Polyarthrop	231	59	2	4	7	0.662	0.446	0.778	0.816	-0.0844	0.0002	*
ICD-9 715 Osteoarthritis et al.	348	62	3	5	7	0.703	0.689	0.778	0.816	-0.0642	0.0001	*
ICD-9 716 Arthropathies Nec/Nos	3353	63	3	4	6	0.715	0.689	0.778	0.827	-0.0601	0.0001	*
ICD-9 719 Joint Disorder Nec & Nos	1730	54	2	3	5	0.773	0.748	0.810	0.827	-0.0402	0.0001	*
ICD-9 721 Spondylosis et al.	193	59	3	5	7	0.702	0.597	0.778	0.827	-0.0489	0.0002	*
ICD-9 722 Intervertebral Disc Dis	736	49	2	3	6	0.716	0.689	0.778	0.827	-0.0737	0.0001	*
ICD-9 723 Other Cervical Spine Dis	456	48	2	3	5	0.795	0.778	0.827	0.854	-0.0283	0.0001	*
ICD-9 724 Back Disorder Nec & Nos	3158	49	2	3	5	0.786	0.761	0.816	0.844	-0.0455	0.0001	*
ICD-9 726 Periph Enthesopathies	757	51	2	4	6	0.787	0.768	0.816	0.827	-0.0433	0.0001	*
ICD-9 733 Oth Bone & Cartilage Dis	976	66	3	5	7	0.764	0.708	0.810	0.844	-0.0182	0.0001	*
ICD-9 737 Curvature of Spine	111	45	2	4	5	0.758	0.708	0.816	0.827	-0.0526	0.0003	*
ICD-9 797 Senility w/o Psychosis	131	82	4	6	7	0.545	0.397	0.597	0.708	-0.1997	0.0002	*
NCC = 2	5704	48	—	—	—	0.853	0.800	0.827	1.000	-0.0942	0.0001	*
NCC = 3	3800	53	—	—	—	0.815	0.778	0.827	1.000	-0.0876	0.0001	*
NCC = 4	2514	57	—	—	—	0.788	0.768	0.810	0.844	-0.0711	0.0001	*
NCC = 5	1657	61	—	—	—	0.755	0.708	0.800	0.827	-0.0547	0.0001	*
NCC = 6	1079	63	—	—	—	0.721	0.708	0.778	0.827	-0.0419	0.0002	*
NCC = 7	688	63	—	—	—	0.684	0.573	0.768	0.816	-0.0350	0.0002	*
NCC = 8	397	63	—	—	—	0.661	0.508	0.768	0.810	-0.0344	0.0002	*
NCC = 9	256	65	—	—	—	0.661	0.467	0.708	0.800	0.0026	0.0003	*
NCC ≥ 10	362	66	—	—	—	0.585	0.397	0.708	0.778	0.0097	0.0003	*

American Indian includes Aleut, Eskimo, or Alaskan Native. Income (as % of poverty line): poor/negative = Income1; near poor = Income2; low income = Income3; middle income = Income4; high income = Income5. No degree = Degree1; GED or high school diploma = Degree2; other (e.g., associate's degree) = Degree3; bachelor = Degree4; master or PhD = Degree5. Only ICD-9 codes with  $n = 100$  are displayed. Other ICD-9 codes with  $n < 100$  and full regression coefficients are available in an online appendix.<sup>28</sup> MEPS, Medical Expenditure Panel Survey; CLAD, censored least absolute deviations; NCC, number of chronic conditions.

a. Regression of EQ-5D = constant + ICD-9 042 + ... + ICD-9 797 + Age + Male + AmericanIndian + Black + OtherRace + Hispanic + Income2 + Income3 + Income4 + Income5 + Degree2 + Degree3 + Degree4 + Degree5 + NCC2 + NCC3 + ... + NCC = 10 using CLAD regression (median) specification.

b. NCC: the total number of chronic conditions (including the condition of interest).

\*Condition coefficient statistically significant ( $P < 0.05$ ).



associated with ICD-9 codes would provide researchers with the flexibility of having scores that more directly correspond to the costs of the condition. Second, the 10 QPCs reported in the previous research are based on a different question frame. For the QPCs reported previously, patients were asked if they had *ever* been diagnosed with the condition of interest (e.g., stroke),<sup>8</sup> whereas ICD-9 and CCC condition classifications are based on reports of experiencing a given condition within a given year. Hence, ICD-9 436 "Cerebral Vascular Attack (CVA)" corresponds to an individual who, at some point within the year that the EQ-5D was administered, experienced a CVA (Table 3), and CCC 109 "Acute Cerebrovascular Disease" from the preliminary catalogue<sup>8</sup> corresponds to an individual who, at some point within the year, reported any 1 of 6 three-digit ICD-9 codes (430, 431, 432, 433, 434, or 436). In contrast, QPC "Stroke" from the previous research<sup>8</sup> corresponds to an individual who has previously had a stroke (not necessarily within the past year).

The decision to use broader QPC or CCC codes v. more clinically homogeneous ICD-9 codes in QALY estimation will depend on the requirements of the end-user. However, where a broader approach to the clinical condition of interest is desired (e.g., the generic QPC "Coronary Heart Disease"), or when the sample size of individuals reporting a specific ICD-9 code is too small to ensure accurate estimation of EQ-5D index scores, the CCC or QPC estimates provided in the previous research may be more appropriate.

## Limitations

Administration of the EQ-5D in MEPS could not be correlated to the exact timing of reported conditions. Both were reported within a given year. Hence, the current research has restricted the scope of the analysis to chronic ICD-9 codes (lasting  $\geq 1$  year), as classified by Hwang et al.<sup>16</sup> Some of the ICD-9 codes used in this research may not have a chronic impact for all individuals reporting the condition. There were 5 ICD-9 condition coefficients from the regression results that were positive and were omitted. The following 4 were statistically significant and positive: ICD-9 199 malignant neoplasm, 367 disorders of refraction, 600 hyperplasia of prostate, and other skin hypertrophy/atrophy. Positive disutility from chronic conditions is not consistent with clinical expectation. It is possible that some of these conditions were not truly chronic and their impact was not captured in the EQ-5D index score when elicited, that there are unobserved variables that are positively associated with EQ-5D index

scores for these conditions, and/or that the EQ-5D health status instrument is not sensitive to these conditions. Condition coefficients that were not statistically significant should be used with extreme caution, incorporating appropriate assessment of the greater uncertainty in the estimates.

The ability of survey respondents to report accurate condition data may be a source of inaccuracy for the conditions listed, and this bias may be exacerbated in blacks and Hispanics. Although the ICD-9 codes were verified and error rates for professional coders did not exceed 2.5% in MEPS,<sup>10</sup> they were based on self-report. Previous research has shown that blacks, whites, and Hispanics differ in reporting of disease labeling, levels of illness, and disability, and there is evidence that self-reported conditions may be underreported in general.<sup>24-26</sup>

As discussed, MEPS data do not contain information on severity of disease. Future research examining the impact of disease severity on preference scores for different conditions is needed. In addition, future research validating the condition estimates in this general population is necessary.

The construct validity, reliability, and responsiveness of the EQ-5D index have been documented extensively in both general and specific disease populations. Although considered a major advantage of the EQ-5D, the parsimony in the items and levels of the questionnaire may result in a potential lack of discrimination. Previous research has suggested that EQ-5D index scores based on the US scoring algorithm exhibit strong ceiling effects.<sup>8</sup> This characteristic has necessitated the use of censored regression methods to derive unbiased estimates. The minimum possible decrement in the EQ-5D index score is 0.140 for the US weights. The inability to quantify a health state between 0.86 and 1.0 may lead to lack of discrimination for mild health states. This has also been suggested as a limitation of the UK scoring algorithm for the EQ-5D.<sup>27-29</sup> Similar to the QWB and the SF-6D, the scoring function of the EQ-5D is based on a linear additive model, which assumes no interactions in preferences between attributes and may be a limitation of the EQ-5D index. In contrast, the HUI relies on a multiplicative functional form.<sup>30</sup> In addition, the US scoring algorithm uses TTO values to estimate preferences based on a MAVF rather than standard gamble (SG)-derived utilities based on a multiattribute utility function (MAUF). Although not free of criticism,<sup>1,31,32</sup> the SG results in utilities that are more consistent with the inherent uncertainty in health decisions required by Von Neuman Morganstern expected utility theory.<sup>33,34</sup> Despite these limitations, there is no

consensus that other preference-based health status classification systems are superior to the EQ-5D index.<sup>1,29,35,36</sup> Each has its own advantages and limitations, and the EQ-5D is the only theory-based health status classification system with a scoring function based on US general population preferences that is currently available in a nationally representative data set.

Although consensus guidelines exist on the appropriate use of preference-based HRQL scores, inappropriate use continues to be widespread. The wide variability and inconsistencies across estimates of preference-based HRQL scores in prior cost-effectiveness analyses<sup>3,7</sup> underscore the need for consistency in reference-case cost-effectiveness evaluations. This need for consistency lead the PCEHM to call for a national catalogue of preference-based HRQL scores<sup>1</sup> as an important step toward promoting comparability of cost-effectiveness analyses. Although there are limitations in the current research, it provides an important addition to the preliminary catalogue developed by Sullivan et al.<sup>8</sup> and moves toward the consistency called for by the PCEHM. Perhaps the availability of accessible EQ-5D index scores associated with chronic ICD-9 codes will provide analysts with the flexibility and breadth of “off-the-shelf” preference-based HRQL scores needed to improve the consistency of QALY estimation in cost-effectiveness analysis. Additional studies are necessary to validate these scores within condition-specific populations.

## CONCLUSION

The PCEHM has called for an “off-the-shelf” catalogue of nationally representative, community-based preference scores for health states as well as illnesses and conditions. This research augments the preliminary catalogue developed by Sullivan et al.<sup>8</sup> by providing EQ-5D index scores associated with a wide variety of chronic ICD-9 codes that can be used by researchers to estimate QALYs in cost-effectiveness analyses.

## REFERENCES

- Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996.
- Drummond MF. *Methods for the Economic Evaluation of Health Care Programmes*. 2nd ed. Oxford, UK: Oxford University Press; 1997.
- Bell CM, Chapman RH, Stone PW, Sandberg EA, Neumann PJ. An off-the-shelf help list: a comprehensive catalog of preference scores from published cost-utility analyses. *Med Decis Making*. 2001;21:288–94.
- 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.
- Gold MR, Franks P, McCoy KI, Fryback DG. Toward consistency in cost-utility analyses: using national measures to create condition-specific values. *Med Care*. 1998;36:778–92.
- Erickson P, Wilson R, Shannon I. *Years of Healthy Life*. Atlanta, GA: Centers for Disease Control and Prevention; 1995.
- Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. *Med Care*. 2000;38:583–637.
- Sullivan PW, Lawrence WF Jr, Ghushchyan V. A national catalogue of preference-based scores for chronic conditions in the U.S. *Med Care*. 2005;43:736–49.
- Elixhauser A, Andrews RM, Fox S. *Clinical Classifications for Health Policy Research: Discharge statistics by principal diagnosis and procedure*. AHCPR Pub. No. 93-0043. Rockville, MD: Agency for Health Policy and Research; 1993.
- Agency for Healthcare Research and Quality (AHRQ). *Puf Documentation Files*. Agency for Healthcare Research and Quality (AHRQ). Available at: <http://www.meps.ahrq.gov/PUFFiles/H52/H52doc.htm>. Accessed 10 November 2003.
- US Health Care Financing Administration. *The International Classification of Diseases, 9th revision, Clinical Modification: ICD-9-CM*. 2nd ed. Washington, DC: US Department of Health and Human Services, Public Health Service; 1980.
- The EuroQol Group. Available at: <http://www.euroqol.org/download/ref.pdf>. Accessed 11 November 2003.
- McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 2nd ed. New York: Oxford University Press; 1996.
- Shaw JW, Johnson JA, Coons SJ. U.S. valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005;43:203–20.
- Agency for Healthcare Research and Quality. *Computing Standard Errors for MEPS Estimates*. Rockville, MD: Agency for Healthcare Research and Quality. Accessed September 2003.
- Hwang W, Weller W, Ireys H, Anderson G. Out-of-pocket medical spending for care of chronic conditions. *Health Aff (Millwood)*. 2001;20:267–78.
- Austin PC. Bayesian extensions of the Tobit model for analyzing measures of health status. *Med Decis Making*. 2002;22:152–62.
- Powell J. Least absolute deviations estimation for the censored regression model. *J Econometrics*. 1984;25:303–25.
- Maddala GS. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press; 1999.
- Green WH. *Econometric Analysis*. 4th ed. Upper Saddle River, NJ: Prentice Hall; 2000.
- Petersen D, Waldman D. The treatment of heteroscedasticity in the limited dependent variable model. Mimeo. November 1981.
- <http://www.stata.com/stb/stb58>
- The ICD-9 to CCC crosswalk is available in the Medical Conditions Public Use Files (MEPS HC-069) at [www.meps.ahrq](http://www.meps.ahrq) (appendix 3, p A3-1).
- Berkanovic E, Telesky C. Mexican-American, black-American and white-American differences in reporting illnesses, disability and physician visits for illnesses. *Soc Sci Med*. 1985;20:567–77.

25. Stewart AL, Napoles-Springer A. Health-related quality-of-life assessments in diverse population groups in the United States. *Med Care*. 2000;38(9 Suppl):II102–24.
26. Johnson TP, O'Rourke D, Chavez N, Sudman S. Cultural variations in the interpretation of health survey questions. In: Warnecke RB, ed. *Health Survey Research Methods*. Hyattsville, MD: National Center for Health Statistics; 1996. p 57–62.
27. Glick HA, Polsky D, Willke RJ, Schulman KA. A comparison of preference assessment instruments used in a clinical trial: responses to the visual analog scale from the EuroQol EQ-5D and the Health Utilities Index. *Med Decis Making*. 1999;19:265–75.
28. Wu AW, Jacobson KL, Frick KD, et al. Validity and responsiveness of the Euroqol as a measure of health-related quality of life in people enrolled in an AIDS clinical trial. *Qual Life Res*. 2002;11:273–82.
29. Macran S, Weatherly H, Kind P. Measuring population health: a comparison of three generic health status measures. *Med Care*. 2003;41:218–31.
30. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the Health Utilities Index Mark 3 system. *Med Care*. 2002;40:113–28.
31. Richardson J. Cost utility analysis: what should be measured? *Soc Sci Med*. 1994;39:7–21.
32. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. The measurement of patients' values in medicine. *Med Decis Making*. 1982;2:449–62.
33. Von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior*. 2nd ed. Princeton, NJ: Princeton University Press; 1947.
34. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ*. 1986;5:1–30.
35. Luo N, Chew LH, Fong KY, et al. A comparison of the EuroQol-5D and the Health Utilities Index Mark 3 in patients with rheumatic disease. *J Rheumatol*. 2003;30:2268–74.
36. Bosch JL, Hunink MG. Comparison of the Health Utilities Index Mark 3 (HUI3) and the EuroQol EQ-5D in patients treated for intermittent claudication. *Qual Life Res*. 2000;9:591–601.