Estimating utilities for chronic kidney disease, using SF-36 and SF-12-based measures: challenges in a population of veterans with diabetes

**Mangala Rajan · Kuan-Chi Lai · Chin-Lin Tseng ·
Shirley Qian · Alfredo Selim · Lewis Kazis ·
Leonard Pogach · Anushua Sinha**

Abstract

Purpose Using transformations of existing quality-of-life data to estimate utilities has the potential to efficiently provide investigators with utility information. We used within-method and across-method comparisons and estimated disutilities associated with increasing chronic kidney disease (CKD) severity.

Methods In an observational cohort of veterans with diabetes (DM) and pre-existing SF-36/SF-12 responses, we used six transformation methods (SF-12 to EQ-5D, SF-36 to HUI2, SF-12 to SF-6D, SF-36 to SF-6D, SF-36 to SF-6D (Bayesian method), and SF-12 to VR-6D) to estimate unadjusted utilities. CKD severity was staged using glomerular filtration rate estimated from serum creatinines, with the modification of diet in renal disease formula. We then used multivariate regression to estimate disutilities specifically associated with CKD severity stage.

Results Of 67,963 patients, 22,273 patients had recent-onset DM and 45,690 patients had prevalent DM. For the recent-onset group, the adjusted disutility associated with CKD derived from the six transformation methods ranged from 0.0029 to 0.0045 for stage 2; −0.004 to −0.0009 for early stage 3; −0.017 to −0.010 for late stage 3; −0.023 to −0.012 for stage 4; −0.078 to −0.033 for stage 5; and −0.012 to −0.001 for ESRD/dialysis.

Conclusion Disutility did not increase monotonically as CKD severity increased. Differences in disutilities estimated using the six different methods were found. Both findings have implications for using such estimates in economic analyses.

M. Rajan · C.-L. Tseng · L. Pogach · A. Sinha
Center for Healthcare Knowledge Management, Veterans Health Administration New Jersey, East Orange, NJ, USA

K.-C. Lai
Robert Wood Johnson Medical School - University of Medicine and Dentistry of New Jersey, New Brunswick, NJ, USA

C.-L. Tseng · L. Pogach · A. Sinha (✉)
Department of Preventive Medicine and Community Health, New Jersey Medical School - University of Medicine and Dentistry of New Jersey, 185 South Orange Avenue, MSB F506, Newark, NJ 07101, USA
e-mail: sinhaan1@umdnj.edu

S. Qian · A. Selim · L. Kazis
Department of Health Policy and Management, Center for the Assessment of Pharmaceutical Practices (CAPP), Boston University School of Public Health, Boston, MA, USA

Introduction

Health economic analyses of new medical interventions require preference-weighted measures of quality of life, utilities, to estimate quality-adjusted life years (QALYs) [1]. However, estimating utilities can be resource-intensive. Using transformations of already existing quality-of-life data that are often descriptive and non-preference-based has the potential to efficiently provide investigators with utility information, based on large samples of responses. Such

methods have been proposed by multiple investigators [2–8]. However, direct comparisons of methods remain sparse in the literature [9–15] and have not been conducted for chronic kidney disease patients with type 2 diabetes mellitus. In addition, the ability of such methods to estimate expected increases in disutility in patients with incremental disease severity has not been tested or compared across methods. Potential biases that may be inherent to using pre-existing data also require exploration and identification.

To address this information gap, we used existing Short Form 36 and 12 Item Health Surveys (SF-36/SF-12) responses from veterans with diabetes mellitus (DM) and chronic kidney disease (CKD) and applied six published transformations to discriminate increasing utility losses (disutilities) as CKD severity increased. Psychometric profile measures such as SF-36 capture information broken down by health dimensions. However, these scores are not easily used for program evaluation or cost-effectiveness analyses. Mathematical models that transform existing psychometric scores into preference-weighted utility scales offer investigators a cost-efficient alternative, relative to administering full-scale preference-based surveys to the population of interest, such as US veterans. Many such transforming methods have been proposed; in our analysis of US veterans with diabetes, we compared six such approaches.

The methods were chosen based on their intention to transform SF-36/SF-12 into well-recognized preference-based, multi-attribute utility scales such as HUI2, SF-6D or VR-6D, and EQ-5D [2, 4, 7, 8, 16–18]. We chose the example of type-2 diabetes and CKD as being of public health and policy importance. The prevalence of type 2 diabetes is projected to increase from 14% in 2010 to 21% in 2050 in the US adult population [19]. Among diabetics, the prevalence of CKD is estimated to be >40% and is a source of increased healthcare utilization and healthcare costs [20].

Investigators who are considering using pre-existing SF-36/SF-12 responses to develop utility estimates using pre-developed transformation methods need to ask themselves: Does each of the available transformation methods perform in a manner that suggests consistency (within-method comparison)? Among the methods that do perform consistently, how will choice of method affect utility estimates (cross-method comparison)? With this in mind, we posed the following analytic questions in conducting this study: (1) Within-method comparison: Did the transformation method under consideration yield a stepwise increase in disutility (when adjusted for covariates), as chronic kidney disease stage increased? and (2) Cross-method comparison: At each stage of chronic kidney disease, were the estimates of disutility generated by the six methods under consideration similar or did they differ?

Methods

Overview

A cross-sectional sample of a well-described longitudinal cohort of veterans with diabetes was used. SF-36/SF-12 responses were available for a subset of cohort members. International Classification of Diseases, 9th edition (ICD-9) codes determined whether respondents were on dialysis or had end-stage renal disease (ESRD). If subjects did not have ESRD/dialysis, serum creatinines were used to stage CKD. Six previously published SF-36/SF-12-to-utility transformations were used to estimate utilities for each subject [2, 4, 7, 8, 16–18]. Generalized linear regression models estimated the disutility associated with each CKD stage, after adjustment for demographics, socio-economics, and comorbid disease conditions.

Population and data sources

US veterans with DM were selected who were both members of a large research cohort of veterans with diabetes and respondents to the Large Veterans Health Survey (LVHS) in 1999 [21]. The research cohort used was the Diabetes Epidemiology Cohort (DEpiC), a well-established longitudinal cohort of veterans with diabetes used in health services research for the past 9 years [22]. DEpiC includes data files from the Veterans Health Administration (VHA inpatient, outpatient, and medication files) and the Center for Medicare and Medicaid Services (CMS) [23] (CMS eligibility, MEDPAR, and Part B files) for all VHA users with diabetes [22]. The diagnosis of diabetes was determined using a validated approach [23], based on having two or more diabetes-specific ICD-9 codes (250.xx, 357.2, 362.0, 366.41) from inpatient or outpatient physician visits.

We selected those US veterans who responded to the LVHS and were members of the DEpiC cohort in 1999, 2000, or 2001. Subjects were categorized as having recent-onset DM (duration of ≤3 years) or prevalent DM (duration >3 years), based on an LVHS survey question regarding duration of diabetes. This categorization was done, as utilities and disutilities experienced by patients may be influenced by duration of illness [24–27].

Included within the 1999 LVHS was the Veterans RAND 36-item health survey (VR-36), a version of the SF-36 that has been previously validated in US veteran populations [21]. Individual responses to the 36 questions, as well as eight subscales—general health (GH), physical function (PF), role physical (RP), role emotional (RE), vitality (VT), bodily pain (BP), social function (SF), mental health (MH)—and two summary scales—physical component scale (PCS) and mental component scale (MCS)—were utilized for the analysis. In addition, we also calculated the

PCS-12 and MCS-12 summary scales derived from the subset of SF-12 questions. The subscales derived from VR-36 scores were converted to SF-36 scores by using conversion formulas previously validated for the RP and RE subscales, from 5 point ordinal scales to dichotomized yes/no choices for each of the items [18, 28, 29].

Inclusion and exclusion criteria

To be considered for inclusion in the study, subjects had to have responded to the LVHS in 1999 and have utilized the VHA or Medicare system at least once in fiscal year (FY) 2000. Further, they had to be a DEpiC cohort member in FY1999, FY 2000, or FY 2001. Exclusion criteria included failure to respond to LVHS items pertaining to diabetes, absence of demographic data, and absence of creatinine testing (see CKD staging below). Subjects with certain severe illnesses likely to impact quality of life and/or result in anticipated life expectancies less than $\sim$ one year were also excluded from the study. These illnesses were defined as end-stage hepatic disease, active cancer (other than cancers of the skin and prostate), lower extremity amputations, advanced retinopathy, dementia, and other cognitive impairments. These subjects were considered likely to have large disutilities related to these dominant conditions, making discrimination of additional disutility due to CKD problematic. The presence or absence of these conditions was determined by ICD-9 and CPT codes (Supplemental Table 1). However, in a secondary analysis, these subjects with dominant conditions were added back to the study cohort.

Chronic kidney disease staging

First, it was determined whether respondents were on dialysis or had end-stage renal disease (ESRD), using ICD-9 diagnostic and procedure codes (Supplemental Table 1) [30]. If subjects did not have codes for ESRD and/or dialysis, the presence or absence of CKD and the stage of CKD (if present) were assessed using serum creatinine tests. Estimated glomerular filtration rate (eGFR) from serum creatinine was calculated using the modification of diet in renal disease (MDRD) formula [31]. CKD was defined as no CKD or CKD without alteration of eGFR: stage 0/1 (eGFR ≥ 90 mL/min); stage 2 (60 mL/min ≤ eGFR < 90 mL/min); early stage 3 (45 mL/min ≤ eGFR < 60 mL/min); late stage 3 (30 mL/min ≤ eGFR < 45 mL/min); stage 4 (15 mL/min ≤ eGFR < 30 mL/min); or stage 5 (eGFR < 15 mL/min), using an established approach [32, 33].

Two serum creatinine tests in FY 1999 or FY 2000 were required to establish CKD stage. The first test in FY 1999 was flagged as an index test. Subjects had to have a subsequent serum creatinine test at ≥90 days but <365 days after the index test, in order to be staged. If the index eGFR was ≥90 mL/min, then the subject was determined to have no CKD (stage 0) or CKD without altered eGFR (stage 1). If the index eGFR <90 mL/min, CKD was staged based on the subsequent serum creatinine/eGFR using the categories given above.

Other independent variables

Subjects' claims data from DEpiC provided information about comorbid disease conditions. These conditions were classified into diabetes-related and non-diabetes-related categories using ICD-9-CM codes (Supplemental Table 2). The diabetes-related categories included macrovascular diseases (coronary artery disease (CAD), congestive heart failure (CHF), arrhythmia, cerebrovascular disease, peripheral vascular disease, gangrene), microvascular diseases (diabetic retinopathy, lower extremity ulcers, lower extremity amputations), and metabolic diseases (hyperglycemia, hypoglycemia, diabetic ketoacidosis, hyperosmolarity, and diabetic coma). The non-diabetes-related diseases were categorized as gastrointestinal, musculoskeletal, pulmonary, neurological, mental and substance-abuse disorders, and other disorders.

Control variables included in multivariate analysis were demographic variables (age, race, marital status, education, employment), VHA eligibility status (priority code), lifestyle behaviors (smoking, alcohol consumption, exercise frequency), and comorbid conditions (diabetes-related and unrelated, as described above).

Utility models

Six previously published transformations (A = SF-12 to EQ-5D, B = SF-36 to HUI2, C = SF-12 to SF-6D, D = SF-36 to SF-6D, E = SF-36 to SF-6D (Bayesian method), F = SF-12 to VR-6D) were used to estimate utilities. The six methods were chosen based on their intention to transform SF-36/SF-12 into well-recognized, preference-based, multi-attribute utility scales: HUI2, SF-6D (VR-6D), and EQ-5D [4, 6–8, 16–18]. Four methods (C, D, E, F) revalue SF-36/SF-12 items to a descriptive system and then apply weights in order to transform the revalued items into a utility score, whereas the two other methods (A, B) rely on regression-based transformation of SF-36 subscales or summary scales to a utility score. Two of six methods (A, B) were developed using US general populations, three methods were developed using a UK general population (C, D, E), while the final method (F) was developed using US veterans. Table 1 summarizes the models used. In each of the six models, the dependent variable is a utility score that ranges between 0 and 1. This utility score is not specific to CKD but reflects the overall health utility experienced by the respondent.

Descriptive analyses and multivariate analyses
for the estimation of CKD-related disutility

Means, confidence intervals, medians, and interquartile ran-
ges were used to summarize descriptive data, including utility
scores generated by each of the six methods described above.

Generalized linear models were used to estimate the
incremental disutility associated with each CKD stage. We
took a step-wise approach to building nested models. These
were built starting with an unadjusted model; then, we built
models that adjusted for demographics and socio-economic
variables. Finally, we developed a model that adjusted for
demographics, socio-economic variables, and comorbid
disease conditions. Generalized linear models were also
used to test the null hypothesis that utilities estimated by
different methods were equal.

In order to test the impact of mental health disease
burden on the results, we calculated utilities among a
subgroup of subjects with no mental health or substance-
abuse disorders. In order to assess the impact of removing
subjects with dominant conditions, a secondary analysis
was performed including all subjects. SAS version 9.2
(SAS Institute, Cary, NC) was used for all analyses.

Results

Demographics and epidemiology of diabetes mellitus
and chronic kidney disease

A total of 141,747 veterans with diabetes responded to the
LVHS question about diabetes duration. Of these, 3,081
were excluded because of missing demographic informa-
tion (age, sex, race) required for the conversion of serum
creatinine into eGFR. A further 38,143 were excluded for
having severe and/or life-threatening illnesses (dominant
medical conditions), as defined in Methods. Of note, of the
remaining 100,253 patients, a greater proportion of those
both with mental illness and without CKD were found to
have undergone creatinine testing compared to those with
no mental illness and without CKD (29 vs. 22%). 22,341 of
these 100,253 patients did not have a creatinine test done
and were also excluded. Of the remaining, 10,219 had an
index eGFR of <90 mL/min; however, we were unable to
confirm the CKD stage with a qualifying serum creatinine
test. After excluding these individuals, the final study
population consisted of 67,963 respondents.

Of these patients, 22,273 patients had recent-onset DM
and 45,690 patients had prevalent DM. Compared to those
with recent-onset DM, subjects with prevalent DM patients
were older, more likely to have diabetes-related diseases,
and had lower scores on the SF-36 domain of physical
functioning (Tables 2, 3, 4). As CKD severity increased in

both recent-onset and prevalent DM, increasing propor-
tions of subjects had macrovascular conditions such as
CAD, CHF, and arrhythmia. Patients with CKD stage 0/1
had higher rates of mental illness and substance abuse
relative to those with CKD stage 2 through 5. For example,
among those with recent-onset DM, 26.4% of subjects with
CKD stage 0/1 had a mental illness, while 12.5–19.9% of
those with CKD stage 2–5 had a mental illness (Table 2).

SF-36 and SF-12 scores

Among those with recent-onset DM, as CKD severity
increased, the SF-36 physical function (PF) and role
physical (RP) scores declined (Table 4). This pattern was
also seen in subjects with prevalent DM. SF-36 mental
health (MH) scores among recent-onset DM ranged from a
mean ($\pm$sd) of 59.4 ± 25.6 in those with CKD stage 0/1,
increasing to 65.0 ± 23.0 among those with CKD early
stage 3, then decreasing to a mean of 58.0 ± 24.0 for those
with ESRD/dialysis (Table 4).

Utility and disutility estimates

The range of estimated utilities for individual respondents
varied by transformation method used (Fig. 1). These
ranges were SF-36 to HUI2, 0.249–1.023; SF-36 to EQ-5D,
0.377–1.000; SF-12 to SF-6D, 0.345–1.000; SF-36
to SF-6D (non-Bayesian approach), 0.301–1.000; SF-36
to SF-6D (Bayesian approach), 0.203–1.000; and SF-12
to VR-6D, 0.297–0.997. Mean utility values followed an
"S" shaped curve, with improvements in utility at CKD
stage 2 compared to stage 0/1 and an improvement in
utility from stage 5 to ESRD/dialysis, although these latter
estimates were associated with wide confidence intervals
due to small numbers. Distributions were not overly
skewed, and, generally, the mean utility for each CKD
stage was centered, relative to extreme values (Supple-
mentary Table 3). Furthermore, the utilities estimated at
each CKD stage (stage 0/1, stage 2, early/late stage 3, stage
4/5 (no dialysis), ESRD/dialysis) were found to differ
across transformation methods ($p < 0.0001$).

Figure 2 summarizes results of multivariate analyses
estimating disutilities associated with CKD itself. Estimated
disutilities by all six methods fell into different but over-
lapping ranges. For the recent-onset group, CKD-related
disutility ranged from −0.004 to −0.001 for early stage 3;
−0.017 to −0.010 for late stage 3; −0.023 to −0.012
for stage 4; −0.078 to −0.033 for stage 5; and −0.012 to
−0.001 for ESRD/dialysis (full models available upon
request to author). CKD stage 2 was associated with a
positive disutility, relative to stage 0/1. At each CKD stage,
the absolute value of the associated disutility was less for the
prevalent diabetes group than for the recent-onset diabetes

Table 1 Models used to transform SF-36 and SF-12 health survey responses to utility scores

Model	Transformation	Mathematical approach	SF-36 measures used in model	Coefficients
A	SF-12 → EQ-5D	Censored least absolute deviation (CLAD) regression	PCS-12 and MCS-12[a]	EQ-5D = 0.057867 + 0.010367*PCS-12 + 0.00822*MCS-12 − 0.000034*PCS-12 × MCS-12 − 0.01067
B	SF-36 → HUI2	Ordinary least square (OLS) regression	SF-36 eight domain scores[b] and age	HUI2 = 0.0009*GH + 0.00046*RP + 0.0043*BP + 0.0042*MH + 0.0018*PF + 0.0018*VT + 0.0015*RE + 0.0015*SF − 0.0006*AGE + 0.045
C	SF-12 → SF-6D	Revaluation using standard gamble with ordinary least square model restricted to unity	SF-6D six dimension levels from SF-12[c]	SF-6D = 1.000 − 0.045*PF3 − 0.063*RL234 − 0.063 *SF2 − 0.066*SF3 − 0.081*SF4 − 0.093*SF5 − 0.042*PAIN3 − 0.077*PAIN4 − 0.137*PAIN5 − 0.059*MH23 − 0.113*MH4 −0.134*MH5 − 0.078*VIT234 − 0.106*VIT5 − 0.077*MOST
D	SF-36 → SF-6D	Revaluation using standard gamble with ordinary least square model restricted to unity	SF-6D six dimension levels from SF-36[c]	SF-6D = 1.000 − 0.035*PF23 − 0.044*PF4 − 0.056*PF5 − 0.117*PF6 − 0.053*RL234 − 0.057*SF2 − 0.059*SF3 − 0.072*SF4 − 0.087*SF5 − 0.042*PAIN23 − 0.065*PAIN4 − 0.102*PAIN5 − 0.171*PAIN6 − 0.042*MH23 − 0.100*MH4 − 0.118*MH5 − 0.071*VIT234 − 0.092*VIT5 − 0.061*MOST
E	SF-36 → SF-6D	Revaluation using standard gamble with nonparametric Bayesian model	SF-6D six dimension levels from SF-36[c]	Each health state (N = 18,000) corresponds to a specific utility, requested from author
F	SF-12 → VR-6D	Uses predicted SF-6D from the eight domain scores with nonlinear transformation and missing data adjustment to estimate VR-6D	SF-36 eight domain scores[b]	Multi-step calculation, requested from author

[a] PCS-12 = SF-12 physical component score; MCS-12 = SF-12 mental component score

[b] SF-36 domains: GH general health, RP role limitation due to physical problem, BP bodily pain, MH mental health, PF physical functioning, VT vitality, RE role limitation due to emotional problem, SF social functioning

[c] SF-6D dimensions have various levels of severity indicated numerically and coded as dummy variables: PF physical functioning, RL role limitations, SF social functioning, PAIN pain, MH mental health, VIT vitality, MOST an interaction term that is defined when one or more dimensions are at one of the most severe levels

Table 2 Demographics and comorbid conditions for subjects with recent-onset diabetes mellitus, by CKD stage

	Stage 0/1	Stage 2	Early stage 3	Late stage 3	Stage 4	Stage 5	ESRD/dialysis	Total
Total N	8,689	9,315	2,898	926	217	16	212	22,273
Percent	39.0%	41.8%	13.0%	4.2%	1.0%	0.1%	1.0%	100.0%
Age (years)								
Under 55	37.8%	14.7%	5.2%	4.6%	6.0%	6.3%	15.1%	4,893 (22.0%)
55–64	31.6%	21.0%	13.9%	12.1%	15.7%	25.0%	16.0%	5,284 (23.7%)
65–69	16.1%	21.9%	21.5%	15.4%	17.1%	12.5%	17.0%	4,278 (19.2%)
70–79	12.9%	36.2%	48.9%	49.8%	47.0%	50.0%	45.3%	6,575 (29.5%)
≥ 80	1.6%	6.3%	10.5%	18.0%	14.3%	6.3%	6.6%	1,243 (5.6%)
Gender								
Male	97.5%	97.9%	97.7%	97.8%	97.2%	100.0%	96.2%	21,759 (97.7%)
Race								
White	68.3%	82.3%	85.6%	87.1%	82.0%	62.5%	67.9%	17,223 (77.3%)
African American	19.8%	10.8%	8.7%	7.8%	10.6%	37.5%	19.8%	3,119 (14.0%)
Other/missing	11.9%	6.9%	5.8%	5.1%	7.4%	0.0%	12.3%	1,931 (8.7%)
Marital status								
Married	55.5%	63.5%	65.4%	65.0%	65.0%	75.0%	56.1%	13,509 (60.7%)
Cigarettes								
Current smoker	18.3%	10.4%	7.2%	6.9%	5.1%	12.5%	14.6%	8,766 (39.4%)
Diabetes-related comorbidity								
Total[a]	59.0%	68.4%	78.8%	90.0%	96.3%	100.0%	97.6%	15,040 (67.5%)
Microvascular	20.5%	23.0%	34.1%	61.1%	85.7%	100.0%	92.9%	5,882 (26.4%)
Macrovascular	35.9%	51.1%	62.7%	73.5%	74.7%	81.3%	85.8%	10,730 (48.2%)
Metabolic	24.3%	23.0%	24.4%	23.7%	22.1%	18.8%	28.3%	5,284 (23.7%)
Non-diabetes-related comorbidity								
Total[a]	69.0%	68.9%	69.3%	73.8%	70.0%	56.3%	84.4%	15,446 (69.3%)
Gastrointestinal	18.8%	20.7%	21.7%	24.7%	16.6%	25.0%	37.3%	4,541 (20.4%)
Pulmonary	20.8%	24.3%	28.1%	31.2%	31.8%	37.5%	35.8%	5,326 (23.9%)
Musculoskeletal	40.4%	43.7%	45.1%	48.2%	46.5%	31.3%	44.3%	9,542 (42.8%)
Neurological	2.8%	1.5%	1.4%	1.1%	1.8%	6.3%	6.6%	453 (2.0%)
Mental illness	26.4%	19.9%	18.6%	16.4%	14.3%	12.5%	31.6%	4,939 (22.2%)
Substance abuse	6.1%	2.5%	1.7%	2.3%	1.4%	0.0%	8.0%	851 (3.8%)

[a] Sum across subgroups of conditions (e.g., microvascular, macrovascular, metabolic) exceeds total as individual subjects may have more than one comorbid condition

group. For both recent-onset and prevalent diabetes groups, differences in utility values were significant ($p < 0.01$) at each CKD stage and for each contrasted pair of methods.

Secondary analyses

In the first of these analyses, we calculated utilities among a subgroup of subjects with no mental health or substance-abuse disorders, as it was standard practice in VHA mental health clinics in 1999/2000 to send serum creatinine as part of a screening panel of tests and because this practice may have influenced results. The subgroup consisted of 53,060 patients divided into 17,004 subjects with recent-onset DM and 36,056 subjects with prevalent DM. In this analysis, for those with prevalent diabetes, utility decreased from stage

0/1 to stage 5 CKD and then increased for those subjects with ESRD/dialysis (Fig. 3). This was not the case for those with recent-onset diabetes, although the median utility increase from CKD stage 0/1 to stage 2 was 0.0083, less than the median increase in utility from stage 0/1 to stage 2 seen in the primary analysis (0.0166).

We conducted another secondary analysis without excluding patients with dominant illnesses ($n = 88,503$). In this analysis, 27,675 had recent-onset DM. Patients with dominant illnesses were slightly older than those without (37 vs. 35% over 70 years). Mean PCS and MCS scale scores were similar, with differences in MCS and PCS mean scores between those with and without dominant conditions of −0.62 and −0.54, respectively. Estimated CKD-related disutilities in the analysis including patients

Table 3 Demographics and comorbid conditions for subjects with prevalent diabetes mellitus, by CKD stage

	Stage 0/1	Stage 2	Early stage 3	Late stage 3	Stage 4	Stage 5	ESRD/dialysis	Total
Total N	14,753	17,889	7,232	2,957	975	160	1,724	45,690
Percent	32.3%	39.2%	15.8%	6.5%	2.1%	0.4%	3.8%	100.0%
Age (years)								
Under 55	26.8%	8.7%	3.8%	3.5%	5.0%	14.4%	10.2%	6,147 (13.5%)
55–64	31.6%	17.4%	12.5%	12.9%	14.4%	26.9%	21.7%	9,631 (21.1%)
65–69	19.5%	22.3%	22.2%	16.4%	18.5%	20.6%	21.4%	9,532 (20.9%)
70–79	19.2%	43.0%	50.9%	49.8%	46.7%	28.8%	39.6%	16,853 (36.9%)
≥80	2.9%	8.5%	10.6%	17.4%	15.5%	9.4%	7.1%	3,527 (7.7%)
Gender								
Male	97.8%	98.3%	98.2%	97.7%	97.3%	97.5%	98.7%	44,806 (98.1%)
Race								
White	66.6%	79.4%	82.9%	80.8%	74.7%	55.0%	65.5%	34,348 (75.2%)
African American	20.9%	11.9%	10.3%	11.4%	15.2%	33.8%	23.4%	6,892 (15.1%)
Other/missing	12.5%	8.8%	6.8%	7.8%	10.2%	11.3%	11.1%	4,450 (9.7%)
Marital status								
Married	61.2%	66.7%	67.5%	66.1%	64.3%	67.5%	63.9%	29,626 (64.8%)
Cigarettes								
Current smoker	15.1%	7.6%	6.0%	5.2%	7.7%	7.5%	7.8%	16,064 (35.2%)
Diabetes-related comorbidity								
Total[a]	71.5%	77.9%	86.4%	92.7%	98.3%	94.4%	99.5%	36,305 (79.5%)
Microvascular	36.3%	39.2%	53.5%	72.8%	90.9%	93.8%	98.4%	21,117 (46.2%)
Macrovascular	43.1%	56.6%	68.2%	74.6%	77.9%	60.6%	90.7%	26,043 (57.0%)
Metabolic	29.1%	28.4%	30.6%	30.1%	30.7%	18.1%	34.2%	13,396 (29.3%)
Non-diabetes-related comorbidity								
Total[a]	65.0%	65.9%	67.5%	69.1%	66.4%	46.9%	81.6%	30,429 (66.6%)
Gastrointestinal	18.0%	19.7%	21.7%	20.7%	20.8%	14.4%	36.8%	9,212 (20.2%)
Pulmonary	18.1%	21.5%	24.8%	27.5%	28.1%	10.6%	36.3%	10,051 (22.0%)
Musculoskeletal	39.1%	41.4%	43.5%	43.7%	42.6%	23.1%	42.3%	18,783 (41.1%)
Neurological	2.9%	2.2%	1.9%	2.1%	3.1%	1.3%	12.8%	1,270 (2.8%)
Mental Illness	23.4%	18.5%	17.0%	17.0%	15.4%	12.5%	28.0%	9,136 (20.0%)
Substance abuse	4.4%	1.8%	1.2%	1.0%	1.1%	1.9%	3.1%	1,162 (2.5%)

[a] Sum across subgroups of conditions (e.g., microvascular, macrovascular, metabolic) exceeds total as individual subjects may have more than one comorbid condition

with dominant conditions were also similar to disutilities estimated in the primary analysis. As in the original analysis, each transformation resulted in CKD stage 2 being associated with a positive disutility, relative to stage 0/1 (CKD stage 2 disutility range across methods, 0.0037–0.0053, for those with recent-onset diabetes). As in the primary analysis, the magnitude of mean disutilities at each CKD stage was less for the prevalent diabetes group than for the recent-onset diabetes group (data not shown).

Discussion

We found that all six methods of transforming the psychometric profile measure SF-36 or SF-12 into a preference-based utility scale were able to discriminate among utilities associated with progressive stages of chronic kidney disease (CKD), in veterans with recent-onset or prevalent diabetes. The disutilities associated with each CKD stage, estimated from multivariate regression models using stage 0/1 as baseline, were all statistically significant and, for CKD stages 3 onwards, less than zero.

However, in the unadjusted utility analysis, CKD stage 2 had a higher mean utility than stage 0/1 CKD, in each of the six transformation methods. In the adjusted model, the disutility associated with stage 2 was also positive. Discussion with Veterans Health Administration clinicians suggests that this result may be a consequence of the routine practice of sending serum creatinine as part of the screening panel of laboratory studies for veterans receiving

Table 4 SF-36 domain and subscale scores, by CKD stage

| SF-36 domain or subscale mean (SD)	Stage 0/1	Stage 2	Early stage 3	Late stage 3	Stage 4	Stage 5	ESRD/dialysis
Recent-onset DM							
PF—Physical function	46.4 (29.3)	46.4 (28.6)	42.1 (27.3)	37.8 (26.3)	36.2 (26.1)	26.5 (22.7)	33.3 (26.2)
RP—Role physical	44.4 (32.2)	44.0 (30.9)	39.8 (29.4)	35.8 (29.2)	32.9 (27.6)	37.5 (35.2)	31.1 (27.8)
BP—Bodily pain	43.0 (26.7)	46.0 (26.5)	44.5 (26.5)	42.1 (26.3)	42.1 (24.8)	34.8 (26.7)	39.0 (25.3)
GH—General health	40.1 (24.0)	42.9 (23.3)	40.6 (22.4)	37.8 (20.9)	33.9 (21.7)	28.6 (18.2)	31.6 (21.1)
VT—Vitality	37.9 (24.6)	39.3 (23.9)	37.6 (23.5)	34.5 (22.0)	32.2 (21.4)	25.9 (24.7)	33.1 (21.4)
SF—Social functioning	53.2 (31.9)	58.7 (31.1)	57.6 (30.6)	53.7 (30.3)	52.2 (30.1)	42.2 (28.5)	43.5 (29.0)
RE—Role emotional	57.5 (34.3)	60.6 (33.3)	57.9 (33.4)	53.5 (34.6)	51.1 (33.7)	50.0 (33.8)	50.0 (33.9)
MH—Mental health	59.4 (25.6)	64.6 (23.6)	65.0 (23.0)	63.4 (22.3)	62.0 (23.1)	56.8 (29.9)	58.0 (24.0)
Standard physical component scale (PCS)	33.0 (11.0)	32.7 (10.9)	31.1 (10.4)	29.6 (9.7)	28.8 (9.4)	27.6 (9.6)	28.1 (8.9)
Standard mental component scale (MCS)	42.2 (14.3)	44.8 (13.5)	44.8 (13.1)	43.7 (12.6)	42.8 (12.9)	39.7 (16.4)	40.9 (12.7)
Prevalent DM							
PF—Physical function	45.0 (29.4)	43.8 (28.2)	38.8 (27.5)	33.5 (26.2)	31.7 (25.6)	30.5 (27.4)	26.4 (25.0)
RP—Role physical	43.1 (31.8)	41.7 (30.2)	37.3 (29.2)	32.7 (27.9)	30.5 (28.0)	27.7 (29.4)	24.0 (26.4)
BP—Bodily pain	42.4 (26.5)	43.8 (25.9)	42.0 (25.6)	40.3 (25.0)	38.6 (25.8)	42.0 (29.4)	36.1 (25.3)
GH—General health	37.6 (22.7)	39.6 (22.3)	37.4 (21.7)	34.0 (20.6)	30.9 (20.1)	25.8 (19.2)	25.2 (17.9)
VT—Vitality	37.7 (24.0)	38.1 (23.6)	35.5 (22.9)	32.9 (22.1)	31.0 (22.8)	31.9 (23.7)	28.2 (21.3)
SF—Social functioning	53.7 (31.4)	56.9 (30.6)	54.5 (30.3)	50.5 (29.8)	48.1 (30.0)	42.2 (32.2)	39.5 (28.9)
RE—Role emotional	57.5 (34.2)	58.2 (33.5)	55.6 (33.8)	51.7 (34.3)	49.8 (35.3)	47.3 (37.0)	44.0 (35.8)
MH—Mental health	60.5 (24.9)	64.1 (23.4)	64.3 (22.6)	62.9 (22.2)	62.3 (22.5)	60.2 (23.3)	58.7 (22.9)
Standard physical component scale (PCS)	32.1 (10.9)	31.5 (10.5)	29.7 (10.1)	27.9 (9.5)	26.8 (9.0)	26.7 (9.6)	25.1 (8.2)
Standard mental component scale (MCS)	42.7 (13.9)	44.4 (13.3)	44.1 (12.9)	43.2 (12.8)	42.6 (12.9)	41.6 (12.8)	40.4 (12.8)
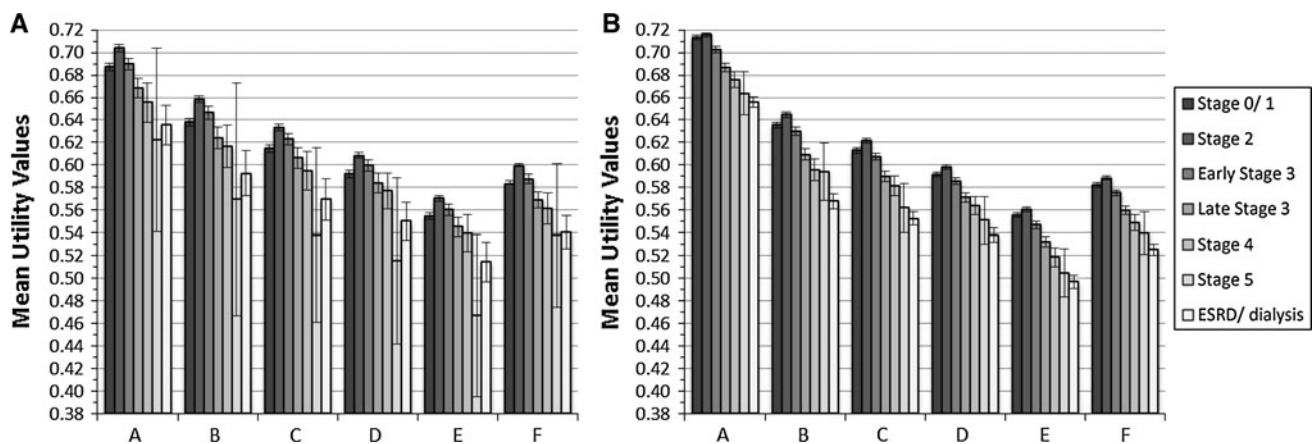
Fig. 1 Mean utility values (± 95% confidence interval) by CKD stage by transformation method. **a** Recent-onset DM. **b** Prevalent DM. Methods: A = SF-12 to EQ-5D, B = SF-36 to HUI2, C = SF-12 to SF-6D, D = SF-36 to SF-6D (non-Bayesian), E = SF-36 to SF-6D (Bayesian), F = SF-12 to VR-6D

care in mental health clinics within the VHA system, during the time period under consideration. This explanation is supported by the finding that the subgroup with CKD stage 0/1 had not only a higher proportion of patients having mental illness but also lower SF-36 MCS scores,

compared to all other stages except for ESRD/dialysis. In a sub-analysis in which patients with ICD-9 codes indicative of mental illness were removed, utilities decreased from CKD stages 0/1 through CKD stage 5 in both adjusted and unadjusted analyses, for those with prevalent diabetes. This
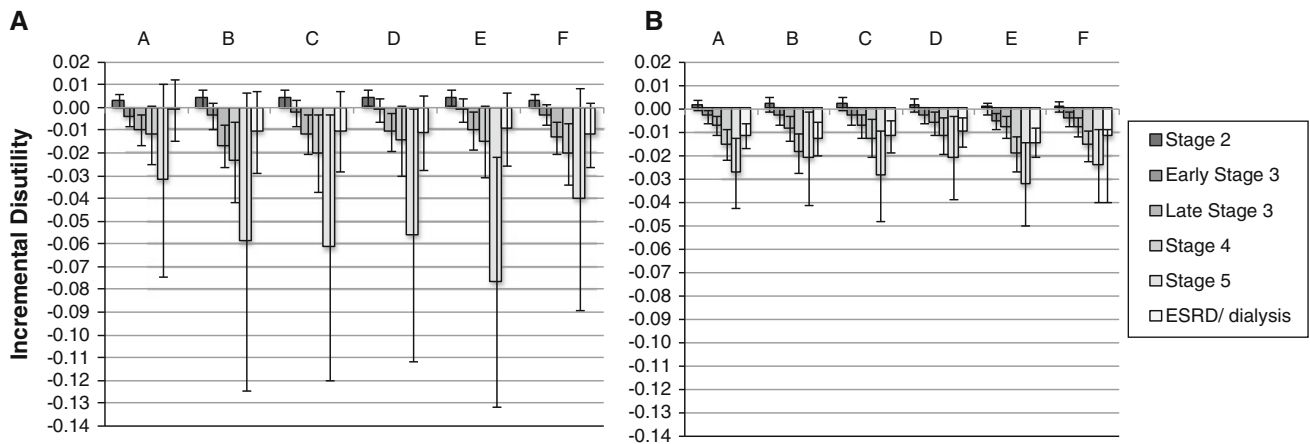
Fig. 2 Incremental (adjusted) disutility (± 95% confidence interval) associated with increasing CKD severity by transformation method. Comparator group was CKD stage 0/1. **a** Recent-onset DM.

b Prevalent DM. Methods: *A* = SF-12 to EQ-5D, *B* = SF-36 to HUI2, *C* = SF-12 to SF-6D, *D* = SF-36 to SF-6D (non-Bayesian), *E* = SF-36 to SF-6D (Bayesian), *F* = SF-12 to VR-6D
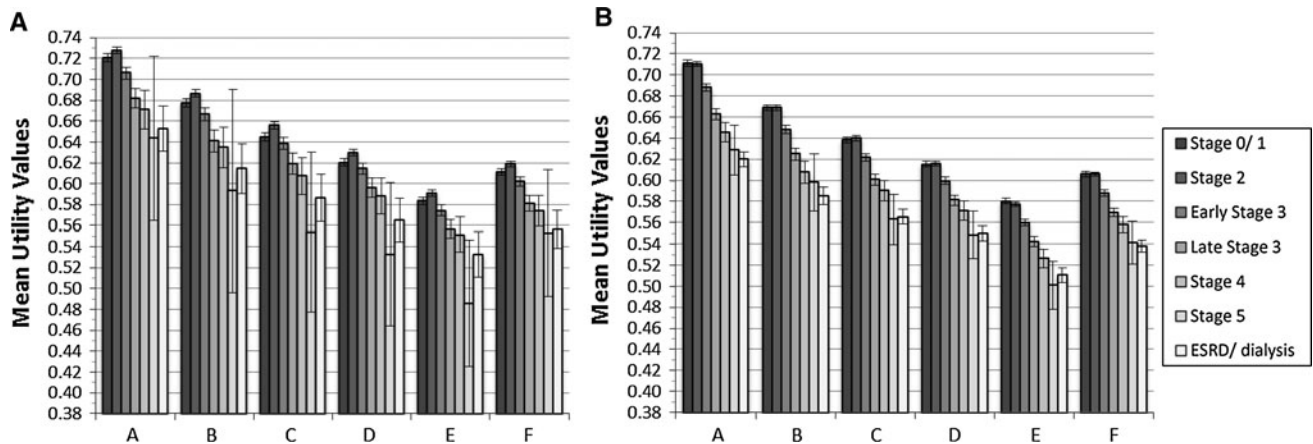

Fig. 3 Mean utility values (± 95% confidence interval) by CKD stage by transformation method, for subset of patients without mental illness or substance-abuse disorders. **a** Recent-onset DM. **b** Prevalent

DM. Methods: *A* = SF-12 to EQ-5D, *B* = SF-36 to HUI2, *C* = SF-12 to SF-6D, *D* = SF-36 to SF-6D (non-Bayesian), *E* = SF-36 to SF-6D (Bayesian), *F* = SF-12 to VR-6D

finding indicates that care must be taken when using pre-existing quality-of-life data in the development of utilities, with a careful search for potential biases introduced by subject selection.

Utilities were also found to increase in patients with an established diagnosis of end-stage renal disease ± dialysis relative to those with stage 5 CKD not yet in receipt of the ESRD diagnosis or dialysis. This was the result of improved scores on most SF-36 subscales for those in the former group, relative to those in the latter. The improved utility for patients on dialysis may represent improved quality of life relative to that experienced by patients with advanced CKD who have not transitioned to renal replacement therapy, although estimates were associated with wide confidence intervals, a result of relatively small numbers of subjects in these groups.

Other investigators have found the SF-36 to SF-6D transformations to be associated with a floor effect in a

variety of populations [3, 34, 35], perhaps due to its standard gamble valuation method [36] and a descriptive system that may not cover the most severe states [37]. The tendency to overestimate utilities at the lower end of the utility scale may be most problematic in the SF-12-based SF-6D, which may not be as suitable to assess utility in severely ill patients [4, 6]. The higher predicted utilities in the SF-36 to EQ-5D transformation method may be due to ceiling effects associated with EQ-5D's valuation method [38]. However, unlike other investigators, we did not find that the estimated utility distributions experienced severe floor or ceiling effects, nor were distributions overly skewed. The range of utility values for the SF-36 to HUI2 transformation did exceed one using the published method, suggesting that some modification of the transformation algorithm at the extreme is needed for this population.

The advantages of our work include a large dataset of greater than 67,000 subjects, representative of the US

veteran population with diabetes. To the best of our knowledge, we are the first group to report the discriminant validity of the SF-6D Bayesian non-parametric model as well as the VR-6D model, compared to other transformation methods.

There have only been a few studies publishing utilities for health states associated with CKD [39]. One such study investigated utilities by CKD stage in 205 renal clinic patients, and our results fall in the general range of HUI3 utilities estimated [40]. Another study of 62 dialysis patients reported a mean SF-6D utility of 0.62 [41], higher than SF-6D results for the current study's subjects. The difference may result from the younger, non-diabetic dialysis population used, with fewer non-renal comorbid disease conditions relative to the veterans with diabetes in the current analysis. Both of these studies had small sample sizes and were not specific to diabetes, gaps in the literature that our study addresses.

Limitations to this study included missing data in VR-36/SF-36 responses. Some questions were not completed by survey respondents, preventing estimation of values for necessary independent variables used in the transformation models. For example, because the Bayesian model required most survey item questions answered to define explicit SF-6D health states, it had the greatest removal of respondents due to missing responses to certain items. In contrast, since the VR-6D method has a built-in algorithm for imputing missing data, it used most of the available patient surveys and had very limited removal of respondents. Because of missing data, including absence of creatinine testing, we assessed utilities in only a subset of total LVHS respondents with diabetes and thus may have introduced selection bias in either direction.

In addition, we had only a single survey response per subject and therefore can only report group (between patient) differences but not longitudinal (within patient) differences among these transformation methods. Although SF-6D has been shown to detect longitudinal differences [35], whether or not this will be true in veterans with diabetes and CKD cannot be tested using our study design. We used data from 1999/2000 and therefore classification of comorbid conditions based on ICD-9 coding may not reflect contemporary practice. Our study used surveys from CKD patients to estimate CKD utilities. Societal preferences are often desired in cost-utility analysis for resource allocation purposes. Because patient populations and general populations may value the same health state differently [42, 43], future work needs to take this into consideration. Finally, our work was not intended to identify a "best" transformation method or to rank methods. It was instead descriptive and evaluative in nature.

Other transformation methods have been published [44–54]. We chose not to include them in our analysis because either they may have inconsistent/paradoxical coefficient estimates [47], have been found to differentiate poorly between different health status states [12, 14], used UK weights for EQ-5D calculation [44–46, 50, 53], suffered a floor effect [10, 11], did not result in improvements in precision [6, 48], have been difficult to implement [51], or the target utility measure is not widely used in the US [49]. Some of our selected methods have already been shown to demonstrate construct and empirical validity [11, 13, 55, 56].

We have shown that all six selected methods transforming SF-36 into a preference-based utility scale discriminated disutilities associated with increasingly severe CKD stage well in a large US veteran population with diabetes, but that values differed by method and that increases in utility (decreases in disutility) between stage 1 and stage 2 were observed. Directly elicited utilities are preferable as transformed measures may be less accurate [12] and indirect methods such as SF-6D systematically yielded lower utilities than direct methods (SG or TTO) [34, 57, 58]. However, as direct measures are not always available, investigators often turn to indirect measures. Our results suggest that careful evaluation of such methods is necessary before incorporating indirectly derived disutilities for chronic kidney disease and diabetes, and possibly other disease conditions, into health economic models.

References

1. Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. Oxford: Oxford University Press.
2. 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*(2), 271–292.
3. Brazier, J., Roberts, J., Tsuchiya, A., & Busschbach, J. (2004). A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics, 13*(9), 873–884.
4. Brazier, J. E., & Roberts, J. (2004). The estimation of a preference-based measure of health from the SF-12. *Medical Care, 42*(9), 851–859.

5. Brazier, J. E., Yang, Y., Tsuchiya, A., & Rowen, D. L. (2010). A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *European Journal of Health Economics, 11*(2), 215–225.

6. Mortimer, D., & Segal, L. (2008). Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. *Medical Decision Making, 28*(1), 66–89.

7. Nichol, M. B., Sengupta, N., & Globe, D. R. (2001). Evaluating quality-adjusted life years: Estimation of the Health Utility Index (HUI2) from the SF-36. *Medical Decision Making, 21*(2), 105–112.

8. Sullivan, P. W., & Ghushchyan, V. (2006). Mapping the EQ-5D index from the SF-12: US general population preferences in a nationally representative sample. *Medical Decision Making, 26*(4), 401–409.

9. Donald-Sherbourne, C., Unutzer, J., Schoenbaum, M., Duan, N., Lenert, L. A., Sturm, R., et al. (2001). Can utility-weighted health-related quality-of-life estimates capture health effects of quality improvement for depression? *Medical Care, 39*(11), 1246–1259.

10. Glasziou, P., Alexander, J., Beller, E., & Clarke, P. (2007). Which health-related quality of life score? A comparison of alternative utility measures in patients with type 2 diabetes in the ADVANCE trial. *Health Quality of Life Outcomes, 5,* 21.

11. Hollingworth, W., Deyo, R. A., Sullivan, S. D., Emerson, S. S., Gray, D. T., & Jarvik, J. G. (2002). The practicality and validity of directly elicited and SF-36 derived health state preferences in patients with low back pain. *Health Economics, 11*(1), 71–85.

12. Kaplan, R. M., Groessl, E. J., Sengupta, N., Sieber, W. J., & Ganiats, T. G. (2005). Comparison of measured utility scores and imputed scores from the SF-36 in patients with rheumatoid arthritis. *Medical Care, 43*(1), 79–87.

13. Lobo, F. S., Gross, C. R., & Matthees, B. J. (2004). Estimation and comparison of derived preference scores from the SF-36 in lung transplant patients. *Quality of Life Research, 13*(2), 377–388.

14. McDonough, C. M., Grove, M. R., Tosteson, T. D., Lurie, J. D., Hilibrand, A. S., & Tosteson, A. N. (2005). Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among spine patient outcomes research trial (SPORT) participants. *Quality of Life Research, 14*(5), 1321–1332.

15. Pickard, A. S., Wang, Z., Walton, S. M., & Lee, T. A. (2005). Are decisions using cost-utility analyses robust to choice of SF-36/SF-12 preference-based algorithm? *Health Quality of Life Outcomes, 3,* 11.

16. SF-6D Bayesian programme. (2011). Retrieved May 25, 2011, from http://www.sheffield.ac.uk/scharr/sections/heds/mvh/sf-6d/bayesian.html.

17. Kharroubi, S. A., Brazier, J. E., Roberts, J., & O'Hagan, A. (2007). Modelling SF-6D health state preference data using a nonparametric Bayesian method. *Journal of Health Economics, 26*(3), 597–612.

18. Selim, A. J., Rogers, W., Qian, S. X., Brazier, J., & Kazis, L. E. (2011). A preference-based measure of health: The VR-6D derived from the veterans RAND 12-item health survey. *Quality of Life Research, 20*(8), 1337–1347.

19. Boyle, J. P., Thompson, T. J., Gregg, E. W., Barker, L. E., & Williamson, D. F. (2010). Projection of the year 2050 burden of diabetes in the US adult population: Dynamic modeling of incidence, mortality, and prediabetes prevalence. *Population Health Metrics, 8,* 29.

20. US Renal Data System. (2010). *USRDS 2010 annual data report: Atlas of chronic kidney disease and end-stage renal disease in the United States.* National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2010.

21. Kazis, L. E., Ren, X. S., Lee, A., Skinner, K., Rogers, W., Clarke, J., et al. (1999). Health status in VA patients: Results from the Veterans Health Study. *American Journal of Medical Quality, 14*(1), 28–38.

22. Miller, D. R., & Pogach, L. (2008). Longitudinal approaches to evaluate health care quality and outcomes: The Veterans Health Administration diabetes epidemiology cohorts. *Journal of Diabetes Science Technology, 2*(1), 24–32.

23. Miller, D. R., Safford, M. M., & Pogach, L. M. (2004). Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. *Diabetes Care, 27*(Suppl 2), b10–b21.

24. Damschroder, L. J., Zikmund-Fisher, B. J., & Ubel, P. A. (2008). Considering adaptation in preference elicitations. *Health Psychology, 27*(3), 394–399.

25. Kostopoulou, O. (2006). The transient nature of utilities and health preferences. *Medical Decision Making, 26*(4), 304–306.

26. Menzel, P., Dolan, P., Richardson, J., & Olsen, J. A. (2002). The role of adaptation to disability and disease in health state valuation: A preliminary normative analysis. *Social Science and Medicine, 55*(12), 2149–2158.

27. Dolan, P. (1996). Modelling valuations for health states: The effect of duration. *Health Policy, 38*(3), 189–203.

28. Kazis, L. E., Miller, D. R., Clark, J. A., Skinner, K. M., Lee, A., Ren, X. S., et al. (2004). Improving the response choices on the veterans SF-36 health survey role functioning scales: Results from the Veterans Health Study. *The Journal of Ambulatory Care Management, 27*(3), 263–280.

29. Kazis, L. E., Miller, D. R., Skinner, K. M., Lee, A., Ren, X. S., Clark, J. A., et al. (2004). Patient-reported measures of health: The Veterans Health Study. *The Journal of Ambulatory Care Management, 27*(1), 70–83.

30. Tiwari, A., Tseng, C. L., Kern, E. F., Maney, M., Miller, D. R., & Pogach, L. (2007). Facility variation in utilization of angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers in patients with diabetes mellitus and chronic kidney disease. *The American Journal of Managed Care, 13*(2), 73–79.

31. Miller, W. G. (2009). Estimating glomerular filtration rate. *Clinical Chemistry and Laboratory Medicine, 47*(9), 1017–1019.

32. Levey, A. S., Coresh, J., Balk, E., Kausz, A. T., Levin, A., Steffes, M. W., et al. (2003). National Kidney Foundation practice guidelines for chronic kidney disease: Evaluation, classification, and stratification. *Annals of Internal Medicine, 139*(2), 137–147.

33. Tseng, C. L., Kern, E. F., Miller, D. R., Tiwari, A., Maney, M., Rajan, M., et al. (2008). Survival benefit of nephrologic care in patients with diabetes mellitus and chronic kidney disease. *Archives of Internal Medicine, 168*(1), 55–62.

34. Khanna, D., Ahmed, M., Yontz, D., Ginsburg, S. S., Park, G. S., Leonard, A., et al. (2008). The disutility of chronic gout. *Quality of Life Research, 17*(5), 815–822.

35. McDonough, C. M., & Tosteson, A. N. (2007). Measuring preferences for cost-utility analysis: How choice of method may influence decision-making. *Pharmacoeconomics, 25*(2), 93–106.

36. Green, C., Brazier, J., & Deverill, M. (2000). Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics, 17*(2), 151–165.

37. Tsuchiya, A., Brazier, J., & Roberts, J. (2006). Comparison of valuation methods used to generate the EQ-5D and the SF-6D value sets. *Journal of Health Economics, 25*(2), 334–346.

38. Kontodimopoulos, N., Pappa, E., Papadopoulos, A. A., Tountas, Y., & Niakas, D. (2009). Comparing SF-6D and EQ-5D utilities across groups differing in health status. *Quality of Life Research, 18*(1), 87–97.

39. Dale, P. L., Hutton, J., & Elgazzar, H. (2008). Utility of health states in chronic kidney disease: A structured review of the literature. *Current Medical Research and Opinion, 24*(1), 193–206.

40. Gorodetskaya, I., Zenios, S., McCulloch, C. E., Bostrom, A., Hsu, C. Y., Bindman, A. B., et al. (2005). Health-related quality of life and estimates of utility in chronic kidney disease. *Kidney International, 68*(6), 2801–2808.

41. Davison, S. N., Jhangri, G. S., & Feeny, D. H. (2009). Comparing the Health Utilities Index Mark 3 (HUI3) with the Short Form-36 preference-based SF-6D in chronic kidney disease. *Value Health, 12*(2), 340–345.

42. Chapman, B. P., Franks, P., Duberstein, P. R., & Jerant, A. (2009). Differences between individual and societal health state valuations: Any link with personality? *Medical Care, 47*(8), 902–907.

43. Mann, R., Brazier, J., & Tsuchiya, A. (2009). A comparison of patient and general population weightings of EQ-5D dimensions. *Health Economics, 18*(3), 363–372.

44. Ara, R., & Brazier, J. (2008). Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). *Value Health, 11*(7), 1131–1143.

45. Franks, P., Lubetkin, E. I., Gold, M. R., & Tancredi, D. J. (2003). Mapping the SF-12 to preference-based instruments: Convergent validity in a low-income, minority population. *Medical Care, 41*(11), 1277–1283.

46. Franks, P., Lubetkin, E. I., Gold, M. R., Tancredi, D. J., & Jia, H. (2004). Mapping the SF-12 to the EuroQol EQ-5D index in a national US sample. *Medical Decision Making, 24*(3), 247–254.

47. Fryback, D. G., Lawrence, W. F., Martin, P. A., Klein, R., & Klein, B. E. (1997). Predicting Quality of Well-being scores from the SF-36: Results from the Beaver Dam Health Outcomes Study. *Medical Decision Making, 17*(1), 1–9.

48. Gray, A. M., Rivero-Arias, O., & Clarke, P. M. (2006). Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. *Medical Decision Making, 26*(1), 18–29.

49. Hawthorne, G., Densley, K., Pallant, J. F., Mortimer, D., & Segal, L. (2008). Deriving utility scores from the SF-36 health instrument using Rasch analysis. *Quality of Life Research, 17*(9), 1183–1193.

50. Lawrence, W. F., & Fleishman, J. A. (2004). Predicting EuroQoL EQ-5D preference scores from the SF-12 health survey in a nationally representative sample. *Medical Decision Making, 24*(2), 160–169.

51. Le, Q. A., & Doctor, J. N. (2011). Probabilistic mapping of descriptive health status responses onto health state utilities using Bayesian networks: An empirical analysis converting SF-12 into EQ-5D utility index in a national US sample. *Medical Care, 49*(5), 451–460.

52. Lundberg, L., Johannesson, M., Isacson, D. G., & Borgquist, L. (1999). The relationship between health-state utilities and the SF-12 in a general population. *Medical Decision Making, 19*(2), 128–140.

53. Rowen, D., Brazier, J., & Roberts, J. (2009). Mapping SF-36 onto the EQ-5D index: How reliable is the relationship? *Health Quality of Life Outcomes, 7*, 27.

54. Sengupta, N., Nichol, M. B., Wu, J., & Globe, D. (2004). Mapping the SF-12 to the HUI3 and VAS in a managed care population. *Medical Care, 42*(9), 927–937.

55. Ferreira, L. N., Ferreira, P. L., Pereira, L. N., Brazier, J., & Rowen, D. (2010). A Portuguese value set for the SF-6D. *Value Health, 13*(5), 624–630.

56. Petrou, S., & Hockley, C. (2005). An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Economics, 14*(11), 1169–1189.

57. Arnold, D., Girling, A., Stevens, A., & Lilford, R. (2009). Comparison of direct and indirect methods of estimating health state utilities for resource allocation: Review and empirical analysis. *British Medical Journal, 339*, b2688.

58. Schmidlin, M., Fritsch, K., Matthews, F., Thurnheer, R., Senn, O., & Bloch, K. E. (2010). Utility indices in patients with the obstructive sleep apnea syndrome. *Respiration, 79*(3), 200–208.