BMJ Open

Comparison of the EQ-5D 3L and the SF-6D (SF-36) contemporaneous utility scores in patients with chronic kidney disease in Sri Lanka: a cross-sectional survey

Sanjeewa Kularatna,[1] Sameera Senanayake,[1] Nalika Gunawardena,[2] Nicholas Graves[1]

¹Australian Centre for Health Services Innovation, Queensland University of Technology, Kelvin Grove, Queensland, Australia
²World Health Organization, Colombo, Sri Lanka

Correspondence to
Dr Sanjeewa Kularatna;
sanjeewa.kularatna@qut.edu.au

ABSTRACT

Objectives The aim of this study was to compare utility weights of EuroQoL-five-dimension-3 levels (EQ-5D-3L) and Short-Form six-dimension (SF-6D) in a representative cohort of patients with chronic kidney disease (CKD). A cost–utility analysis (CUA) is designed to report the change to costs required to achieve an estimated change to quality-adjusted life years (QALYs). The quality component of a QALY is measured by utility. Utility represents the preference of general population for a given health state. Classification systems of the multi-attribute utility instruments (MAUIs) are used to define these health states. Utility weights developed from different classification systems can vary and may affect the conclusions from CUAs.

Design A community-based cross-sectional study.

Setting Anuradhapura, a rural district in Sri Lanka.

Participants A representative sample of 1096 patients with CKD, selected using the population-based CKD register, completed the EQ-5D-3L and SF-36. SF-6D was constructed from the SF-36 according to the published algorithm. The study assessed discrimination, correlation and differences across the two instruments.

Results Study participants were predominantly male (62.6%). Mean EQ-5D-3L utility score was 0.540 (SD 0.35) compared with 0.534 (SD 0.09) for the SF-6D (p=0.588). The correlation (r) between the scores was 0.40 (p<0.001). Utility scores were significantly different in both males and females between the two tools, but there was no difference in age and educational categories. Both MAUI scores were significantly lower (p<0.001) among those who were in more advanced stages of the disease and the corresponding utility scores of the two instruments in different CKD stages were also significantly different (p<0.05). The largest effect size was seen among the patients on dialysis.

Conclusions The correlation between the scores was moderate. SF-6D had the lowest floor and ceiling effect and was better at detecting different stages of the disease. Thus, based on the evidence presented in this study, SF-6D appears to be more appropriate to be used among patients with CKD.

Strengths and limitations of this study

► The response rate of the study is very high.

► Both tools used in the study (EuroQoL-five-dimension-3 levels and Short-Form (SF) 36) have been previously validated to the Sri Lankan setting.

► Data collectors were experienced in many local and international studies done among patients with chronic kidney disease in Sri Lanka and further they were trained by the principal investigator to ensure the quality of the data collected.

► Our study was a cross-sectional study; thus, we could not assess how utility scores of the two instruments change over time.

► Some of the information related to quality of life (QOL) in SF-36 is considered to be sensitive in nature and the fact that this information was obtained utilising an interviewer-administered questionnaire could have led to some under-reporting in the assessment of QOL though many measures were taken to minimise this issue.

INTRODUCTION

Chronic kidney disease (CKD) is a substantial public health problem with adverse psychological, physical and economic outcomes. The burden of CKD is increasing globally.[1] World Health Report (2002) and Global Burden of Disease (GBD) project stated that the diseases of the kidney contribute much to the global disease burden with approximately 850 000 deaths every year globally.[2] Furthermore, according to GBD study conducted in 2010, of the top causes of disability-adjusted life year (DALY), CKD is ranked 29th globally, 23rd in Southeast Asia and 14th in Sri Lanka.[3] Due to the progressive and disabling nature of CKD, it poses a substantial impact on the quality of life (QOL) of individuals. It is important to measure QOL indicators for the management of patients with CKD.

Several studies have demonstrated a relationship between reduced QOL and increased morbidity and mortality.[4–7]

All over the world, the importance of including QOL indicators in the clinical management of patients has been highlighted. This has come to the limelight after several studies demonstrated the strong relationship between reduced QOL and increased morbidity and mortality.[5 8] Meantime, economic evaluation has become increasingly popular among researchers and policy-makers during resource allocation in recent years. Due to the relationship between QOL and clinical outcome, during the recent years, QOL has become an important health outcome in economic evaluations. In cost-utility analysis (CUA), a method of economic evaluation, outcomes are usually measured in quality-adjusted life years (QALYs), which is a measure of QOL.

The concept of QALYs was developed in the 1970s. It can measure the changes of an individual's quality and quantity of life and can also aggregate these improvements across the individual.[9 10] The change in the QOL in QALY is measured using a set of weights, called utilities, which reflect different health states. For all possible health states, utilities should be measured on a scale where 1 refers to best imaginable health and 0 refers to death.[11] Measuring utilities for different health states is complex and time-consuming. Thus, multi-attribute utility instruments (MAUIs) such as EuroQol-five-dimension 3 levels (EQ-5D-3L),[12] Short Form-six dimension (SF-6D)[13] or the Health Utility Index (HUI)[14 15] are used to define different health states. The utility scores for different health states in different instruments are derived from methods such as standard gambling method,[16] discrete choice experiments[17] and time trade-off experiments.[18] EQ-5D-3L is the most widely used utility instrument at present.[19] EQ-5D-5L, a newer version of EQ-5D, has also been developed and tested recently.[20]

Table 1	Demographic distribution of the sample by the EQ-5D-3L and the SF-6D utility scores			
Variable	**N (%)**	**EQ-5D-3L utility mean (SD)**	**SF-6D utility mean (SD)**	**P value†**
All samples	1096	0.540 (0.35)	0.534 (0.09)	<0.001*
Sex				
Male	686 (62.6)	0.561 (0.34)	0.532 (0.10)	<0.001*
Female	410 (37.4)	0.505 (0.37)	0.539 (0.09)	0.342
Age (years)				
Less than 20	07 (0.6)	0.570 (0.44)	0.455 (0.09)	0.235
20–40	45 (4.1)	0.591 (0.34)	0.536 (0.09)	0.103
41–60	562 (51.3)	0.555 (0.35)	0.540 (0.10)	<0.001*
More than 60	482 (44.0)	0.517 (0.35)	0.529 (0.08)	0.006*
Education status				
No formal education	81 (7.4)	0.448 (0.42)	0.508 (0.09)	0.441
5 Grade	413 (37.7)	0.529 (0.35)	0.536 (0.09)	0.001*
6–11 Grade	377 (34.4)	0.556 (0.34)	0.533 (0.10)	<0.001*
GCE O/L passed	190 (17.3)	0.554 (0.34)	0.540 (0.09)	0.007*
GCE A/L passed	35 (3.2)	0.618 (0.34)	0.540 (0.09)	0.225
Employment status				
Employed	380 (34.7)	0.675 (0.25)	0.547 (0.10)	<0.001*
Not employed	716 (65.3)	0.468 (0.37)	0.528 (0.09)	0.417
Comorbidities				
Present	778 (71.0)	0.505 (0.36)	0.532 (0.09)	0.037*
Absent	318 (29.0)	0.625 (0.29)	0.542 (0.10)	<0.001*
CKD stage				
Early stage	254 (24.0)	0.588 (0.30)	0.551 (0.10)	<0.001*
Stage IV	614 (58.1)	0.566 (0.42)	0.536 (0.09)	<0.001*
Stage V	151 (14.3)	0.467 (0.42)	0.523 (0.08)	0.808
Dialysis	38 (3.6)	0.126 (0.39)	0.432 (0.07)	<0.001*

*Significant at p<0.05.
†Wilcoxon signed-rank test.
EQ-5D-3L, EuroQoL-five-dimension-3 levels; GCE A/L, General Certificate of Education Advance level; GCE O/L, General Certificate of Education Ordinary level; SF-6D, Short-Form six-dimension.
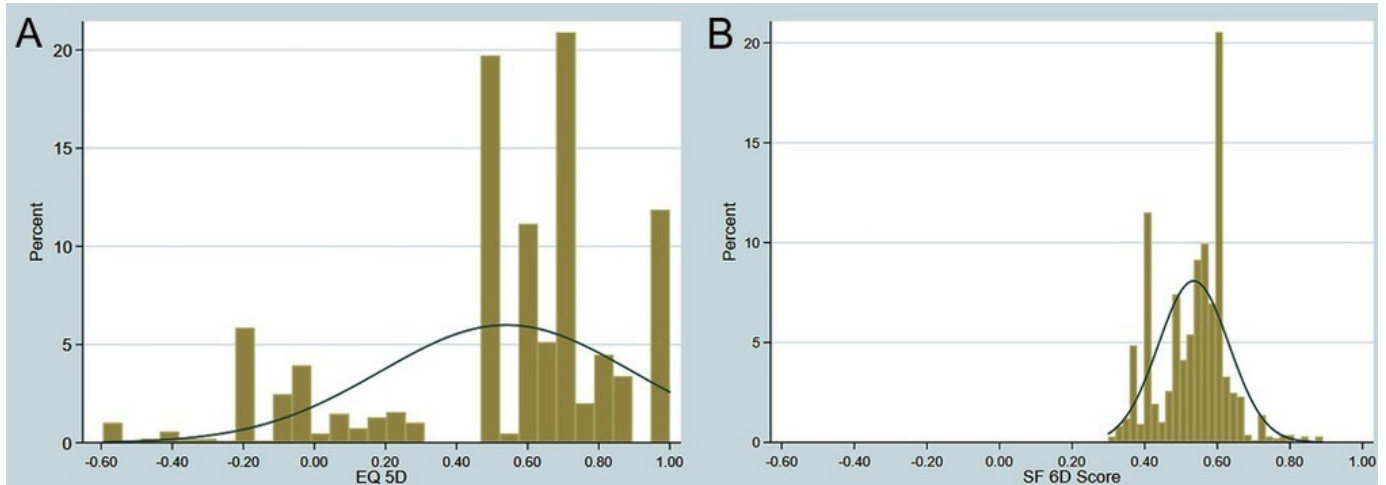
Figure 1 Distribution of EQ-5D-3L (A) and SF-6D (B). EQ-5D-3L, EuroQoL-five-dimension-3 levels; SF-6D, Short-Form six-dimension.

Since all the MAUIs aim at measuring the health state of individuals, all the instruments should generate the same utility value for a particular state of health. However, the evidence indicates that there is an essential difference in the utility scores for a particular health state between different instruments.[19 21–29] This, in turn, indicates that the choice of the MAUI used may adversely influence the results of CUA and thereby the decision-making process.[30] Furthermore, for incremental analyses, use of different MAUIs may lead to different results regarding the magnitude, direction or significance of any change in health-related QOL measure.

Though the differences between MAUIs have been evaluated in many disease conditions,[19 21 22 24] there are no evidence in the literature comparing MAUIs using patients with CKD. The aim of our study is to compare contemporaneous EQ-5D-3L and SF-6D utility scores in patients with CKD. Results may be useful for researchers selecting a generic MAUI to estimate utilities for use in economic modelling of treatments for CKD.

METHODS
Participant selection

A population-based descriptive cross-sectional study was conducted in the district of Anuradhapura in the North Central Province (NCP) of Sri Lanka between September and December 2015. The study population consisted of 1162 confirmed patients with CKD, calculated using the appropriate formula,[31] who were above 18 years with

documented evidence of CKD living in the Anuradhapura district. The diagnosis of CKD was made if the glomerular filtration rate (GFR) was less than $60\,mL/min$ per $1.73\,m^2$ of body surface area in two measurements made 3 months apart.

The inclusion criteria were patients above 18 years and those who were diagnosed as having CKD by a specialist in nephrology or a consultant physician. Presence of evidence of such diagnosis was made by way of diagnosis cards, clinic records or any other record issued by a specialist in nephrologist, a consultant physician or a consultant in government hospital. Patients who had previous renal transplantation, who were unable to provide rational information due to any cause (eg, mental retardation) and who were critically ill were excluded from the study.

The study instrument was an interviewer-administered questionnaire to gather information on the socio-demographic information, CKD-related information, EQ-5D-3L and SF-36.

Five Public Health Inspectors working in the CKD unit in the NCP were used for the data collection and all have been working in the unit for more than 5 years and they had experience in functioning as data collectors for many local and international studies done among patients with CKD in the NCP. The data collectors assessed the eligibility of patients by reviewing their clinical records. Informed consent was obtained from those who were eligible for participation in the study before doing the face-to-face interview.

Table 2 Distribution of the sample by the EQ-5D-3L

	Mobility (%)	Self-care (%)	Usual activities (%)	Pain/discomfort (%)	Anxiety/depression (%)
No problem	515 (47.0)	644 (58.8)	473 (43.2)	182 (16.6)	261 (23.8)
Some problem	559 (51.0)	421 (38.4)	587 (53.6)	739 (67.4)	680 (62.0)
Extreme problem	22 (2.0)	31 (2.8)	36 (3.3)	175 (16.0)	155 (14.1)

EQ-5D-3L, EuroQoL-five-dimension-3 levels.

The study was conducted in all 19 Medical Officer of Health (MOH) areas of the Anuradhapura district. The number of participants to be included from each MOH area was based on probability proportionate to the size of patients with CKD registered in each of the MOH areas. The required number of participants from each MOH area was selected using simple random sampling method. The population-based CKD register—which records the patients with a confirmed diagnosis of CKD from renal clinics in hospitals of the NCP since 2003—was used as the sampling frame. The register was obtained from the office of the Provincial Director of Health Services.[32]

Calculation of utility scores

Currently, there is no algorithm based on preferences of the Sri Lankan public to score the SF-6D on a utility scale. Therefore, the UK algorithm was used for this purpose.[13] Though Sri Lankan EQ-5D-3L utility scores are available,[18] the UK utility scores were used for the EQ-5D-3L[33] because of the unavailability of comparable Sri Lankan SF-6D utility scores as mentioned earlier. This allowed the comparison of utility scores from the same country.

The EQ-5D-3L instrument contains five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each domain has one item and each item has three levels: one denoting no problems and three denoting severe problems.[12] Thus, EQ-5D-3L has mutually exclusive 243 different health states.

SF-6D is derived from either SF-36 or SF-12 (Version 1 and Version 2). The current study utilised SF-36 for data collection. SF-36 includes 36 items that measure eight domains: role limitations caused by physical problems (four items), physical function (10 items), role limitations caused by emotional problems (three items), pain (two items), social function (two items), general health perceptions (five items), emotional well-being (five items) and energy/fatigue (four items). Questions have different answer options which range from 2 to 6. While scoring, each question is scored in a scale ranging from 0 (worst health) to 100 (best health). All items in a domain are summed up and averaged to give an average score for each domain which ranges from 0 (worst health) to 100 (best health). To calculate the utility scores of the SF-6D, 11 items are used covering six domains: physical functioning, role limitation, social functioning, pain, mental health and vitality.[13]

The EQ-5D-3L utility calculation was undertaken using the STATA syntax developed by Ramos-Goni et al.[34] The SF-6D scores were computed based on published algorithms.[13] Patients for whom one of the two measurements was missing were excluded from the analysis.

The EQ-5D-3L utility scores range from −0.59, 0=being dead; negative values represent health status considered worse than 'dead', to 1.00 which indicate best imaginable health. The SF-6D utility scores ranged from 0.296 which indicate severely impaired levels in all dimensions to 1.0 which indicates no difficulty in any dimensions.

Data analysis

STATA V.15.1 software was used for the analysis. Mean utility scores on each instrument were compared by socio-demographic characteristics. Normality of the two distributions was assessed using Kolmogorov-Smirnov and Shapiro-Wilk tests. Wilcoxon signed-rank test was used to assess the difference between the two instruments in each socio-demographic class.[24] Histograms were plotted for the two utility values distribution. Floor effects and ceiling (proportion of patients with the highest and lowest possible scores, respectively) were calculated for the EQ-5D-3L and SF-6D. Ceiling and floor effects were considered small if ≤15% of patients occupy the best or worst health states, but they were considered serious if >15% of patients occupy these states.[35]

Currently, an established methodology to compare different MAUIs is not available. Thus, recently published methodologies, which compared different MAUIs, were followed in the current study.[19 23 35] This included a combination of statistical and psychometric analyses to examine discrimination, agreement, differences and correlation between the two instruments.

Agreement and differences

The Wilcoxon signed-rank test was used to assess the overall difference between the EQ-5D-3L and SF-6D utility scores and the difference of the utility scores according to different socio-demographic and disease-related features. Furthermore, the distribution of the responses to the different domains of the two instruments was tabulated to present the agreement and the differences between the two instruments. Bland-Altman plot was also used to assess the proportional error and the limit of agreement.[36]

Intraclass correlation

The dimensions of the two instruments were compared using intraclass correlation (ICC). The related dimensions between the two MAUIs are role limitation (SF-6D)/usual activities (EQ-5D-3L); physical functioning (SF-6D)/mobility and self-care (EQ-5D-3L); pain (SF-6D)/pain and discomfort (EQ-5D-3L); social functioning (SF-6D)/usual activities (EQ-5D-3L) and mental health (SF-6D)/anxiety and depression (EQ-5D-3L). The vitality dimension of the SF-6D did not have any related dimension with the EQ-5D-3L. The magnitude of the correlation coefficients was interpreted according to Guilford's criteria.[37]

Discrimination

It is important that MAUIs can discriminate correctly among groups of different severity as MAUIs are meant to measure change in QOL due to improvement or worsening of the health, in the condition of interest.

GFR is the most important indicator of kidney function of patients with CKD.[38] Studies have shown that decreased GFR is associated with infection, impaired cognitive and physical function as well as threats to patient safety.[39] Though classifications exist to classify stages of

Table 3 Distribution of the sample by the SF-6D

|	Physical functioning (%)	Role limitation (%)	Social functioning (%)	Pain (%)	Mental health (%)	Vitality (%)
1*	8 (0.7)	173 (15.8)	17 (1.6)	7 (0.6)	6 (0.6)	2 (0.2)
2	22 (2.0)	5 (0.5)	60 (5.5)	5 (0.5)	130 (11.9)	194 (17.7)
3	304 (27.7)	135 (12.3)	482 (44.0)	59 (5.4)	582 (53.1)	498 (45.4)
4	356 (32.5)	783 (71.4)	481 (43.9)	452 (41.2)	364 (33.2)	287 (26.2)
5	87 (7.9)	NA	56 (5.1)	333 (30.4)	14 (1.3)	115 (10.5)
6†	319 (29.1)	NA	NA	240 (21.9)	NA	NA

*No problem.
†Severe problem.
NA, not applicable; SF-6D, Short-Form six-dimension.

CKD, it is evident that, at present, most of the clinical decision-making in CKD is solely based on GFR-based classification.[40 41] Depending on the GFR value, CKD is categorised into five stages: stage I to stage V. For analytical purposes, the CKD stages I to III were categorised as 'early stage' in the present study. It is expected that with advanced stages of the disease, the utility scores should be lower than the early stages.

Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using the nonparametric test, Kruskal-Wallis and effect size. The instrument's ability to discriminate between two adjacent stages was estimated by calculating the effect size. The effect size was calculated by dividing the mean difference of two adjacent CKD stages by the SD of the milder of the two CKD stages.[23 42] Large effect size indicates better discriminating ability of the instrument. The effect size was categorised into small (0.2–0.5), medium (0.5–0.8) and large (more than 0.8).[43]

Test–retest reliability
To assess the test–retest reliability of the study instrument, within a period of 1 week, 30 randomly selected study participants were visited at their households by the data
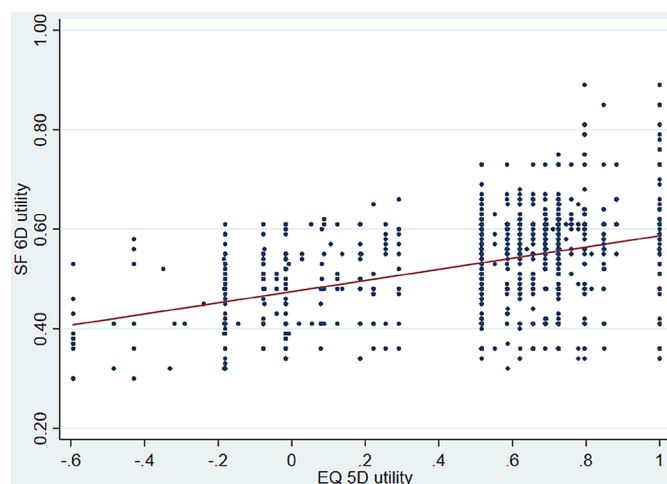
Figure 2 Correlation between the EQ-5D-3L and the SF-6D utilities. EQ-5D-3L, EuroQoL-five-dimension-3 levels; SF-6D, Short-Form six-dimension.

collectors. Test–retest reliability of the utility scores of the two instruments was assessed using ICC and a value of 0.70 or greater was considered as satisfactory reliability.[44]

Patient and public involvement
The main stakeholders, such as consultants, medical officers working in nephrology units, community leaders and the patients living in this area, in the provision of care for the patients with CKD were involved in planning the study. Their concerns were always entertained and if feasible their concerns were incorporated into the study. During the data collection, stage permission was obtained from the respective local officers. The results of the study were communicated to the local-level officials such as MOH, Divisional Secretariat, Regional Director of Health Services and Provincial Director of Health Services.

RESULTS
Sample characteristics
Out of 1162 participants selected to be included in the study, 66 (5.6%) did not participate in the study giving a response rate of 94.4%. The mean age of the study population was 58.4 years (SD 10.8). There was a preponderance of males among the study population (62.6%, n=686). The mean estimated GFR of the population was 31.8 (SD 20.2) mL/min/1.73 m^2. The mean number of years since diagnosed with CKD was found to be 4.1 (SD 3.2) years. The majority of participants was in the later stages, stage IV or beyond, of CKD (n=803; 73.2%). In all, 38 participants (3.6%), with stage V of the disease and undergoing dialysis, were on haemodialysis (table 1). CKD of unknown origin was the cause of the CKD in most of the study population (n=489; 43.7%).

Distribution of EQ-5D-3L and SF-6D utility scores
The mean EQ-5D-3L utility score at baseline was 0.540 compared with 0.534 for the SF-6D as summarised in table 1. The EQ-5D-3L utility score ranged from −0.594 to 1, whereas SF-6D ranged from 0.3 to 0.89. The median baseline values have different locations in their respective scoring ranges (figure 1). The EQ-5D-3L showed 1.0%

Table 4 ICC between the EQ-5D-3L and the SF-6D

	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/ depression
Physical functioning	0.381*	0.326*	0.296*	0.382*	0.381*
Role limitation	0.023	−0.003	−0.104	0.016	0.138*
Social functioning	0.517*	0.424*	0.464*	0.566*	0.528*
Pain	0.475*	0.330*	0.355*	0.482*	0.484*
Mental health	0.293*	0.323*	0.295*	0.240*	0.244*
Vitality	0.322*	0.148*	0.255*	0.475*	0.453*

*Significant at P<0.05 level.
EQ-5D-3L, EuroQoL-five-dimension-3 levels; ICC, intraclass correlation; SF-6D, Short-Form six-dimension.

floor effect and 11.8% ceiling effect, whereas SF-6D had 0.0% floor and ceiling effects.

Agreement, differences and correlation between the two utility scores

Analyses revealed non-normal distribution of the utility scores of both the instruments; thus, Wilcoxon signed-rank test was used to compare the two utility scores. There was a significant difference (p<0.001) between overall scores of the two utility instruments. Furthermore, the two utility scores were significantly different among males (p<0.001), age more than 40 years groups, those who were employed, among both who had and did not have comorbidities, up to stage IV of CKD and among patients on dialysis (table 1). The SD of the EQ-5D-3L was considerably larger than that of the SF-6D among all subgroups.

Significant proportion of the patients reported 'no problem' in any of the EQ-5D-3L than the SF-6D. However, fewer patients reported 'extreme problems' in the EQ-5D-3L than in the SF-6D (tables 2 and 3). Patients reported different results for the related dimensions of the two MAUIs (tables 2 and 3). Nearly half of the patients reported 'no problem' in mobility domain of the EQ-5D-3L, whereas only 0.7% reported 'no problem' with the physical functioning of the SF-6D. Nearly a quarter

(23.8%) of patients reported 'no problem' for the anxiety/depression dimension in the EQ-5D-3L, whereas only 0.6% reported the same for the mental health dimension of the SF-6D.

The correlation between EQ-5D-3L and SF-6D was 0.408, which was statistically significant at p<0.001 level (figure 2). Regarding the ICC between different domains of the two instruments, according to the Guilford's criteria, moderate correlation (0.4–0.6) was evident between social functioning and mobility (0.517); social functioning and self-care (0.424); social functioning and usual activities (0.464); social functioning and pain/ discomfort (0.566); social functioning and anxiety/ depression (0.528); pain and mobility (0.475); pain and pain/discomfort (0.482); pain and anxiety/depression (0.484); vitality and pain/discomfort (0.475) and vitality and anxiety/depression (0.453) (table 4). The Bland-Altman plot showed proportional error and wide limits of agreement (figure 3).

Discrimination

With both MAUIs, utility scores decreased with increasing severity (as measured by CKD stage) (table 5). In both MAUIs, the utility differences across CKD stages were statistically significant (p<0.05) indicating good discrimination. Figure 4 indicated the box-plots present the median, quartiles and extreme values for the EQ-5D-3L and SF-6D utility scores for CKD stage. Furthermore, the calculated effect size between CKD early stage and stage IV was 0.071 and 0.141 for EQ-5D-3L and SF-6D, respectively. The highest effect size was observed between CKD stage V and dialysis group, which was 0.807 for EQ-5D-3L and 1.098 for SF-6D.

Test–retest reliability

The test–retest ICC was 0.943 in EQ-5D-3L while it was 0.921 in SF-6D, indicating good test–retest reliability in both the instruments.
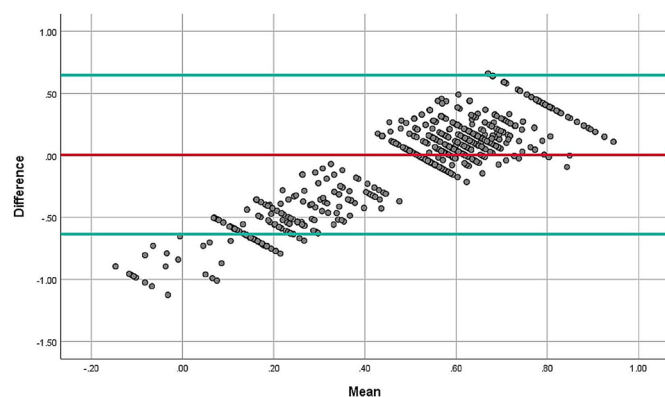
Figure 3 Bland and Altman plot of differences between EQ-5D-3L and SF-6D for patients with CKD. CKD, chronic kidney disease; EQ-5D-3L, EuroQoL-five-dimension-3 levels; SF-6D, Short-Form six-dimension.

DISCUSSION

This is the first study to compare the utility scores arising from the EQ-5D-3L and SF-6D in patients with CKD.

Table 5 Discrimination across clinical severity groups

CKD stage	EQ-5D-3L					SF-6D				
	N	Mean (SD)	Median	Sig*	ES	N	Mean (SD)	Median	Sig*	ES
Early stage	254 (24.0)	0.588 (0.30)	0.656	<0.001		254 (24.0)	0.551 (0.10)	0.570	<0.001	
IV	614 (58.1)	0.566 (0.42)	0.620		0.071	614 (58.1)	0.536 (0.09)	0.560		0.141
V	151 (14.3)	0.467 (0.42)	0.585		0.305	151 (14.3)	0.523 (0.08)	0.550		0.138
Dialysis	38 (3.6)	0.126 (0.39)	−0.016		0.807	38 (3.6)	0.432 (0.07)	0.410		1.098

*Kruskal-Wallis test.
CKD, chronic kidney disease; EQ-5D-3L, EuroQoL-five-dimension-3 levels; ES, effect size; SF-6D, Short-Form six-dimension.

According to the current study, the correlation between the scores was moderate. Both tools were able to discriminate advancement of CKD stages. Effect size, which denoted the discriminating ability of different CKD stages, is highest when disease condition is advanced and the highest effect size was seen in SF-6D. Furthermore, the lowest ceiling effect and the floor effect were seen in SF-6D.

Evidence indicate that the choice of MAUI (eg, EQ-5D or SF-6D) has an impact on the results of the CUA.[45 46] Sack et al[45] compared the results of cost–utility estimates using both EQ-5D and SF-6D. Results indicated contrasting results for the two instruments and authors concluded that the choice of the instrument does matter in CUA.[45] Thus, from an economic perspective, it is important to know the most suitable MAUI to be used among patients with CKD.

At present, there is no consensus on the methodology to compare the utility scores of different MAUIs.[19 35] The present study adopted the methodologies used by Kularatna et al (2017) and Lamers et al (2006).[19 35] Only one time assessment of the utilities was done in the present study. Thus, the responsiveness of the two instruments to changes in kidney function over time was not

Figure 4 The box-plots present the median, quartiles and extreme values for the EQ-5D-3L and SF-6D utility scores for CKD stage. CKD, chronic kidney disease; EQ-5D-3L, EuroQoL-five-dimension-3 levels; SF-6D, Short-Form six-dimension.

assessed. Though Sri Lankan EQ-5D-3L utility scores are available,[18] yet we used the UK utility scores for the EQ-5D-3L[33] because of the unavailability of comparable Sri Lankan SF-6D utility score values. This is an accepted method of calculating the utility scores in the absence of country-specific utilities. Two studies conducted in Netherlands[24] and Italy,[21] comparing the utility scores of the two instruments, had used the UK-derived EQ-5D-3L and SF-6D utility scores.

The present study did not find any difference (p=0.588) between the overall mean scores of the two utility instruments. This was similar to a study conducted among a group of patients with HIV/AIDS,[28] but different from other studies available in the literature where different results have been reported. Significantly higher utility values for EQ-5D-3L were found among general population,[29 47] patients with cardiovascular disease,[19] patients with rheumatoid arthritis[21] and patients with stable angina.[16] However, in a study conducted among a group of patients with psychiatric disorders, significantly higher utility values were obtained for SF-6D instrument.[24] These varying results could be due to different recall periods of the two instruments. EQ-5D-3L assessed the health status of the day of instrument administration while SF-6D, which was derived from SF-36, assessed the health status of the past 30 days.

Though overall ceiling and floor effects of both instruments were small, relatively higher ceiling effect was evident in the EQ-5D-3L. This was consistent with several other studies conducted elsewhere, where EQ-5D-3L reported a relatively higher ceiling effect compared with SF-6D.[16 19 48–50] This is mainly due to the fact that the EQ-5D-3L has limited response levels and the five-level newer version of EQ-5D-3L expected to improve the properties of the three-level version in terms of reduced ceiling effects, increased reliability and improved ability to discriminate between different levels of health.[51] Furthermore, the current study reported relatively lower ceiling effect, for the EQ-5D, compared with results obtained among patients with Parkinson's disease (13.5%) and stable angina (15.5%). However, our result was higher compared with the ceiling effect observed among patients with systemic sclerosis (7.0%). Among many other factors that could contribute to these differences, the level of morbidity of a disease is said to be one of the factors that
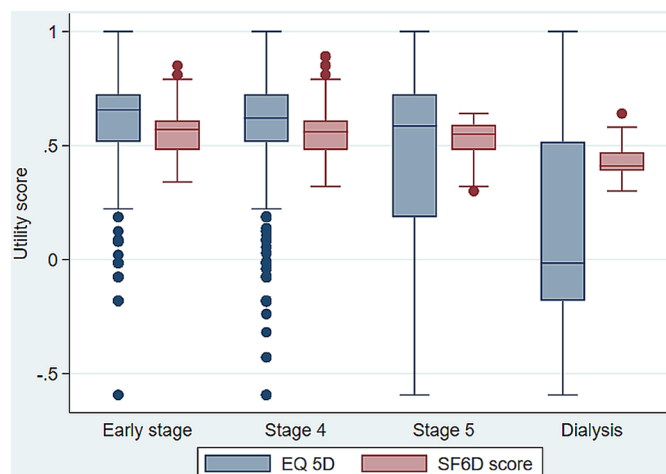
could influence the ceiling effect observed in EQ-5D.[50] Thus, the diseases with lower morbidity are expected to have higher ceiling effects.

Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using analysis of variance (ANOVA) and effect size. In both MAUIs, the utility differences across CKD stages were statistically significant (p<0.05; ANOVA) indicating good discrimination. However, the effect size was small for both the tools until the dialysis stage. At the dialysis stage, the effect size is large and this was highest in the SF-6D instrument. It could be because CKD is considered asymptomatic until the later stages of the disease,[52 53] not allowing the instruments to discriminate the different stages. According to a recent study conducted by Jesky *et al*,[54] the EQ-5D-3L utility scores of the adjacent pre-dialysis CKD stages were not found to be statistically significant.[54]

Limitations

Some of the information related to QOL in SF-36 is considered to be sensitive in nature and the fact that this information was obtained utilising an interviewer-administered questionnaire could have led to some under-reporting in the assessment of QOL though many measures were taken to minimise this issue. Our study was a cross-sectional study; thus, we could not assess how utility scores of the two instruments change over time.

CONCLUSIONS

The correlation between the scores was moderate. Both tools were able to discriminate advancement of CKD stages. Effect size, which denoted the discriminating ability of the different CKD stages, is highest when disease condition is advanced. Findings indicate that both tools cover different aspects of health. Thus, although there was a moderate correlation between the measures, both scores cannot be used interchangeably while assessing QALY during CUA. Finally, SF-6D had the lowest floor and ceiling effect and was better at detecting different stages of the disease. Thus, based on the evidence presented in this study, SF-6D appears to be more appropriate to be used among patients with CKD.

REFERENCES

1. Senanayake S, Gunawardena N, Palihawadana P, *et al*. Validity and reliability of the Sri Lankan version of the kidney disease quality of life questionnaire (KDQOL-SF™). *Health Qual Life Outcomes* 2017;15:119.
2. Schieppati A, Remuzzi G. Chronic renal diseases as a public health problem: epidemiology, social, and economic implications. *Kidney Int Suppl* 2005;68:S7–S10.
3. Evaluation IfHMa. Global Burden of Disease 2010 Arrow Diagram: Institue for Health Metrics and Evaluation. 2010 http://vizhub.healthdata.org/irank/arrow.php
4. Rebollo-Rubio A, Morales-Asencio JM, Pons-Raventos ME, *et al*. Review of studies on health related quality of life in patients with advanced chronic kidney disease in Spain. *Nefrologia* 2015;35:92–109.
5. Mujais SK, Story K, Brouillette J, *et al*. Health-related quality of life in CKD patients: correlates and evolution over time. *Clin J Am Soc Nephrol* 2009;4:1293–301.
6. Tsai YC, Hung CC, Hwang SJ, *et al*. Quality of life predicts risks of end-stage renal disease and mortality in patients with chronic kidney disease. *Nephrol Dial Transplant* 2010;25:1621–6.
7. DeOreo PB. Hemodialysis patient-assessed functional health status predicts continued survival, hospitalization, and dialysis-attendance compliance. *Am J Kidney Dis* 1997;30:204–12.
8. Gussous Y, Theologis AA, Demb JB, *et al*. Correlation Between Lumbopelvic and Sagittal Parameters and Health-Related Quality of Life in Adults With Lumbosacral Spondylolisthesis. *Global Spine J* 2018;8:17–24.
9. Howren MB. *Quality-Adjusted Life Years (QALYs). Encyclopedia of Behavioral Medicine*: Springer, 2013:1605–6.
10. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care* 1989;5:559–75.
11. Drummond MF, Sculpher MJ, Claxton K, *et al*. *Methods for the economic evaluation of health care programmes*: Oxford university press, 2015.
12. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.
13. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;21:271–92.
14. 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.
15. Feeny D, Furlong W, Boyle M, *et al*. Multi-attribute health status classification systems. *Pharmacoeconomics* 1995;7:490–502.
16. Wu J, Han Y, Zhao FL, *et al*. Validation and comparison of EuroQoL-5 dimension (EQ-5D) and Short Form-6 dimension (SF-6D) among stable angina patients. *Health Qual Life Outcomes* 2014;12:156.
17. Stolk EA, Oppe M, Scalone L, *et al*. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. *Value Health* 2010;13:1005–13.
18. Kularatna S, Whitty JA, Johnson NW, *et al*. Valuing EQ-5D health states for Sri Lanka. *Qual Life Res* 2015;24:1785–93.
19. Kularatna S, Byrnes J, Chan YK, *et al*. Comparison of the EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with cardiovascular disease. *Quality of Life Research* 2017;26:3399–408.
20. Herdman M, Gudex C, Lloyd A, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
21. Salaffi F, Carotti M, Ciapetti A, *et al*. A comparison of utility measurement using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis. *Clin Exp Rheumatol* 2011;29:661–71.
22. Kwakkenbos L, Fransen J, Vonk MC, *et al*. A comparison of the measurement properties and estimation of minimal important differences of the EQ-5D and SF-6D utility measures in patients with systemic sclerosis. *Clin Exp Rheumatol* 2013;31(2 Suppl 76):50–6.

23. Rowen D, Young T, Brazier J, *et al*. Comparison of generic, condition-specific, and mapped health state utility values for multiple myeloma cancer. *Value Health* 2012;15:1059–68.
24. Lamers LM, Bouwmans CA, van Straten A, *et al*. Comparison of EQ-5D and SF-6D utilities in mental health patients. *Health Econ* 2006;15:1229–36.
25. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures Un comparación de la calidad de vida de personas con trastornos auditivos estimada por tres diferentes medidas de utilidad. *Int J Audiol* 2005;44:157–63.
26. Barton GR, Bankart J, Davis AC, *et al*. Comparing utility scores before and after hearing-aid provision. *Appl Health Econ Health Policy* 2004;3:103–5.
27. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease. *Health Qual Life Outcomes* 2006;4:20.
28. Stavem K, Frøland SS, Hellum KB. Comparison of preference-based utilities of the 15D, EQ-5D and SF-6D in patients with HIV/AIDS. *Qual Life Res* 2005;14:971–80.
29. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Econ* 2005;14:1169–89.
30. McDonough CM, Tosteson ANA. Measuring preferences for cost-utility analysis. *Pharmacoeconomics* 2007;25:93–106.
31. Lwanga SK, Lemeshow S. *Sample size determination in health studies: a practical manual*: WHO, 1991.
32. Senanayake S, Gunawardena N, Palihawadana P, *et al*. Symptom burden in chronic kidney disease; a population based cross sectional study. *BMC Nephrol* 2017;18:228.
33. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108.
34. Ramos-Goñi JM, Rivero-Arias O. Eq5d: a command to calculate index values for the EQ-5D quality-of-life instrument. *Stata J* 2011;11:120–5.
35. Harrison MJ, Davies LM, Bansback NJ, *et al*. The comparative responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. *Qual Life Res* 2009;18:1195–205.
36. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
37. Tredoux C, Durrheim K. *Numbers, hypotheses and conclusions: a course in statistics for the social sciences*. Cape Town: University of Cape Town Press[Links], 2002.
38. Levey AS, Stevens LA, Coresh J. Conceptual model of CKD: applications and implications. *Am J Kidney Dis* 2009;53:S4–16.
39. Seliger SL, Zhan M, Hsu VD, *et al*. Chronic kidney disease adversely influences patient safety. *J Am Soc Nephrol* 2008;19:2414–9.
40. Stevens PE, Levin A. Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group Members. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann Intern Med* 2013;158:825–30.
41. Tseng CL, Lafrance JP, Lu SE, *et al*. Variability in estimated glomerular filtration rate values is a risk factor in chronic kidney disease progression among patients with diabetes. *BMC Nephrol* 2015;16:34.
42. Kularatna S, Byrnes J, Chan YK, *et al*. Comparison of the EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with cardiovascular disease. *Qual Life Res* 2017;26:3399–408.
43. Cohen J. A power primer. *Psychol Bull* 1992;112:155–9.
44. Litwin MS, Fink A. *How to measure survey reliability and validity*: Sage, 1995.
45. Sach TH, Barton GR, Jenkinson C, *et al*. Comparing cost-utility estimates: does the choice of EQ-5D or SF-6D matter? *Med Care* 2009;47:889–95.
46. Grieve R, Grishchenko M, Cairns J. SF-6D versus EQ-5D: reasons for differences in utility scores and impact on reported cost-utility. *Eur J Health Econ* 2009;10:15–23.
47. Stewart ST, Cutler DM, Rosen AB. Comparison of trends in US health-related quality of life over the 2000s using the SF-6D, HALex, EQ-5D, and EQ-5D visual analog scale versus a broader set of symptoms and impairments. *Med Care* 2014;52:1.
48. Zhao FL, Yue M, Yang H, *et al*. Validation and comparison of EuroQol and short form 6D in chronic prostatitis patients. *Value Health* 2010;13:649–56.
49. Cunillera O, Tresserras R, Rajmil L, *et al*. Discriminative capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey. *Qual Life Res* 2010;19:853–64.
50. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. *Value Health* 2006;9:262–71.
51. Janssen MF, Pickard AS, Golicki D, *et al*. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res* 2013;22:1717–27.
52. de Boer IH, Olan GA, Patel UD. American Society of Nephrology Chronic Kidney Disease Advisory Group. Screening for chronic kidney disease. *Ann Intern Med* 2013;158(5 Pt 1):362–3.
53. Hummel A. [Chronic renal failure, a frequent and asymptomatic pathology]. *Soins* 2010;745:24–6.
54. Jesky MD, Dutton M, Dasgupta I, *et al*. Health-related quality of life impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic kidney disease: a prospective observational study. *PLoS One* 2016;11:e0165675.