

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Comparison of the EQ-5D-3L and the SF-6D (SF-36) contemporaneous utility scores in patients with Chronic Kidney Disease

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2018-024854
Article Type:	Research
Date Submitted by the Author:	18-Jun-2018
Complete List of Authors:	Kularatna, Sanjeewa; Australian Centre For Health Services Innovation, Queensland University of Technology Senanayake, Sameera; Australian Centre For Health Services Innovation, Queensland University of Technology Gunawardena, Nalika; World Health Organization Graves, Nicholas; Australian Centre For Health Services Innovation, Queensland University of Technology
Keywords:	Cost utility analysis, Quality-Adjusted Life Years (QALYs), utility, Chronic Kidney Disease (CKD)

SCHOLARONE™
Manuscripts

Peer Review Only

1 **Comparison of the EQ-5D-3L and the SF-6D (SF-36) contemporaneous utility scores in**
2 **patients with Chronic Kidney Disease**
3

4
5
6 *Sanjeewa Kularatna¹, Sameera Senanayake¹, Nalika Gunawardena², Nicholas Graves¹*
7

8
9 *¹Australian Centre for Health Services Innovation, Queensland University of Technology*

10 *² World Health Organization, Sri Lanka*
11
12
13

14 **Corresponding author;**

15 Sanjeewa Kularatna - *sanjeewa.kularatna@qut.edu.au*
16
17

18
19 **Word count : 3277**
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objectives : The aim of this study was to compare utility weights of EQ-5D-3L and SF-6D in a representative cohort of patients with Chronic Kidney Disease (CKD). A cost-utility analysis 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 (MAUI) are used to define these health states. Utility weights developed from different classification systems can vary and so might affect the conclusions from cost-utility analyses.

Design: A community based cross sectional study

Setting : Anuradhapura a rural district in Sri Lanka.

Participants: A representative sample of 1096 CKD patients completed the EQ-5D-3L and SF-36 from which the SF-6D was constructed 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(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 decreased significantly ($p<0.001$;ANOVA) with advancing CKD stage and the corresponding utility scores of the two instruments in different CKD stages were also significantly ($p<0.05$) different. The largest effect size was seen among the dialysis patients.

Conclusions: The correlation between the scores was moderate. Both tools were able to discriminate advancement of CKD stages. Findings indicate that both the tools cover different spaces in health. Thus, although there was a moderate correlation between the measures, both scores cannot be used interchangeably while assessing QALY during cost utility analysis.

Key words: Cost utility analysis, Quality-Adjusted Life Years (QALYs), utility, Chronic Kidney Disease (CKD)

Article summary

Strengths and limitations of this study

- This is the first study to compare the utility scores arising from the EQ-5D-3L and SF-6D in Chronic Kidney Disease patients
- This is the first study to demonstrate that EQ-5D-3L and SF-6D tools cover different spaces in health among CKD patients.
- Our study was a cross-sectional study, thus we could not assess which utility scores of the two instruments change over time.

For peer review only

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 to the global disease burden with approximately 850,000 deaths every year globally (2). Furthermore, according to GBD study conducted in 2010, out of the top causes of DALY, CKD is ranked 29th globally, 23rd in South East Asia and 14th in Sri Lanka (3). Due to the progressive and disabling nature of CKD it has substantial impact on the quality of life (QOL) of individuals. It is important to measure QOL indicators for the management of chronic kidney disease patients. Several studies demonstrate a relationship between reduced QOL and increased morbidity and mortality (4-7).

World over, 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 individual (9, 10). The change in the quality of life 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 full health and 0 refers to death (11). Measuring utilities for different health states is complex and time-consuming. Thus, Multi Attribute Utility Instruments (MAUI) such as EuroQol-5D (EQ-5D-3L) (12), Short Form-6D (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, which is the newer version of EQ-5D, has also been developed and tested recently (20).

1
2
3 Since all the MAUIs aim at measuring the health state of individuals, all the instruments
4 should generate the same utility value for a particular health state. However, the evidence
5 indicates that there is essential difference in the utility scores for a particular health state
6 between different instruments (19, 21-29). This, in turn, indicates that the choice of the
7 MAUI used may adversely influence the results of CUA and thereby the decision-making
8 process (30). Furthermore, for incremental analyses, use of different MAUIs may lead to
9 different results regarding the magnitude, direction or significance of any change in health-
10 related quality of life measure.
11
12
13
14
15

16
17 Though the differences of different MAUIs have been evaluated in many disease conditions
18 (19, 21, 22, 24), there are no evidence in the literature comparing MAUIs using patients with
19 CKD. The aim of our study is to compare contemporaneous EQ-5D-3L and SF-6D utility
20 scores in patients with CKD. Results may be useful for researchers selecting a generic MAUI
21 to estimate utilities for use in economic modelling of treatments for CKD.
22
23
24
25
26

27 **Methods**

28 *Patients selection*

29
30 A population-based descriptive cross-sectional study was conducted in the district of
31 Anuradhapura in the North Central Province (NCP) of Sri Lanka. The study population
32 consisted of 1162 confirmed CKD patients who were over 18 years old with documented
33 evidence of CKD living in the Anuradhapura district. The diagnosis of CKD was made if the
34 Glomerular Filtration Rate (GFR) was less than 60 ml/min per 1.73m² of body surface area
35 in two measurements made three months apart.
36
37
38
39
40

41
42 The inclusion criteria were patients above 18 years of age and those who were diagnosed as
43 having CKD by a specialist nephrologist or a consultant physician. Presence of evidence of
44 such diagnosis was made by way of diagnosis cards, clinic records or any other record issued
45 by a specialist nephrologist, a consultant physician or a government hospital. Patients who
46 had previous renal transplantation, who were unable to provide rational information due to
47 any cause (e.g. mental retardation) and who were critically ill but reliable information cannot
48 be acquired from them were excluded from the study.
49
50
51
52
53

54
55 The data collectors assessed the eligibility of patients by reviewing their clinical records.
56 Informed consent was obtained from those who were eligible for participation in the study.
57
58
59
60

1 The study was conducted in all nineteen Medical Officer of Health (MOH) areas of the
2 Anuradhapura district. The number of participants to be included from each MOH area was
3 based on probability proportionate to the size of CKD patients registered in each of the MOH
4 areas. The required number of participants from each MOH area was selected using the
5 simple random sampling method. The population-based CKD register – which records the
6 patients with a confirmed diagnosis of CKD from renal clinics in hospitals of the NCP since
7 2003 – was used as the sampling frame. The register was obtained from the office of the
8 Provincial Director of Health Services (31).

15 *Calculation of utility scores*

16 Currently, there is no algorithm based on preferences of the Sri Lankan public to score the
17 SF-6D on a utility scale. Therefore, the UK algorithm was used for this purpose (13).
18 Though Sri Lankan EQ-5D-3L utility scores are available (18), the UK utility scores were
19 used for the EQ-5D-3L (32) because of the unavailability of comparable Sri Lankan SF-6D
20 utility scores as mentioned earlier. This allowed the comparison of utility scores from the
21 same country.
22
23
24
25
26

27
28 The EQ-5D-3L instrument contains five domains; mobility, self-care, usual activities, pain /
29 discomfort and anxiety / depression. Each domain has one item and each item has three
30 levels: one denoting no problems and three denoting severe problems (12). Thus, EQ-5D-3L
31 has mutually exclusive 243 different health states.
32
33
34
35

36 SF-6D is derived from SF-36, SF-12 Version 1 and SF-12 Version 2. The current study
37 utilised SF-36 for the data collection. SF-36 includes 36 items that measure eight domains;
38 role limitations caused by physical problems (4 items), physical function (10 items), role
39 limitations caused by emotional problems (3 items), pain (2 items), social function (2 items),
40 general health perceptions (5 items), emotional well-being (5 items) and energy / fatigue (4
41 items). Questions have different answer options which range from two to seven. While
42 scoring, each question is scored in a scale ranging from 0 (worst health) to 100 (best health).
43 All items in a domain are summed up and averaged to give an average score for each domain
44 which ranges from 0 (worst health) to 100 (best health). To calculate the utility scores of the
45 SF-6D, 11 items are used covering six domains; physical functioning, role limitation, social
46 functioning, pain, mental health and vitality (13).
47
48
49
50
51
52
53
54
55

56 The EQ-5D-3L utility calculation was undertaken using the Stata syntax developed by
57 Ramos-Goni et al. (33). The SF-6D scores were computed based on published algorithms
58
59

1 (13). Patients for whom one of the two measurements was missing were excluded from the
2 analysis.
3

4
5
6 The EQ-5D utility scores range from -0.59, 0=being dead; negative values represent health
7 status considered worse than “dead”, to 1.00 which indicate good health status. Values close
8 to zero indicate worse conditions, while 1.00 represents perfect health status. The SF-6D
9 utility scores ranged from 1.0 which indicates no difficulty in any dimensions to 0.296 which
10 indicate severely impaired levels in all dimensions.
11
12
13

14 *Data analysis*

15
16 Stata 15.1 software was used for the analysis. Distribution of the socio-demographic
17 characteristics of the study population was compared with their mean utility scores. Paired t-
18 test was used to assess the difference between the two instruments in each socio-demographic
19 class (34). Histograms were plotted for the two utility values distribution. Floor effects and
20 ceiling (proportion of patients with the highest and lowest possible scores respectively) were
21 calculated for the EQ-5D and SF-6D. Ceiling and floor effects were considered small if
22 $\leq 15\%$ of patients occupy the best or worst health states, but they were considered serious if
23 $> 15\%$ of patients occupy these states (35).
24
25
26
27
28
29
30

31
32 Currently, an established methodology to compare different MAUIs is not available. Thus,
33 recently published methodologies, which compared different MAUIs, were followed in the
34 current study (19, 23, 34). This included a combination of statistical and psychometric
35 analyses to examine discrimination, agreement, differences and correlation between the two
36 instruments.
37
38
39
40

41 *Agreement and differences*

42
43 The paired t-test was used to assess the difference between the EQ-5D-3L and SF-6D utility
44 scores. Overall difference of the two utility scores as well as the difference of the utility
45 scores according to different socio-demographic and disease related features were assessed.
46
47 Furthermore, the distribution of the responses to the different domains of the two instruments
48 was tabulated to present the agreement and the differences between the two instruments.
49
50

51 *Correlation*

52
53 The dimensions of the two instruments were compared using Spearman correlation
54 coefficient. The related dimensions between the two MAUIs are role limitation (SF-6D)/usual
55 activities (EQ-5D-3L), physical functioning (SF-6D)/mobility and self-care (EQ-5D-3L),
56 pain (SF-6D)/pain and discomfort (EQ-5D-3L), social functioning (SF-6D)/usual activities
57
58
59

(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 were interpreted according to Guilford's criteria (36).

Discrimination

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

Glomerular Filtration Rate (GFR) is the most important indicator of kidney function of patients with CKD (37). Studies have shown that decreased GFR is associated with infection, impaired cognitive and physical function as well as threats to patient safety (38). Though classifications exist to classify stages of CKD, it is evident that at present most of the clinical decision making in CKD is solely based on GFR base classification (39, 40). 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 ANOVA 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 standard deviation of the milder of the two CKD stages (41). 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) (42).

Test-retest reliability

To assess the test-retest reliability of the study instrument, within a period of one week, 30 randomly selected study participants were visited at their households by the data collectors. Test re-test reliability of the utility scores of the two instruments was assessed using Spearman's r correlation coefficient and a value of 0.70 or greater was considered as satisfactory reliability (43).

Results

Sample characteristics

1 Out of 1162 participants selected to be included in the study, 66 (5.6%) did not participate in
2 the study giving a response rate of 94.4%. The mean age of the study population was 58.4
3 years (Standard Deviation (SD) 10.8). There was a preponderance of males among the study
4 population (62.6%, N=686). The mean eGFR of the population was 31.8 (SD 20.2)
5 ml/min/1.73 m². The mean number of years since diagnosed with CKD was found to be 4.1
6 (SD 3.2) years. The majority of participants was in the later stages, stage 4 or beyond, of
7 CKD (n=803; 73.2%). 38 participants (3.6%), with stage 5 of the disease and undergoing
8 dialysis, were on haemodialysis (Table 1). Chronic Kidney Disease of Unknown origin
9 (CKDu) was the cause of the CKD in most of the study population (n=489; 43.7%).
10
11
12
13
14
15

16 *Distribution of EQ-5D-3L and SF-6D utility scores*

17 The mean EQ-5D-3L utility score at baseline was 0.540 compared with 0.534 for the SF-6D
18 as summarised in Table 1. The EQ-5D-3L utility score ranged from -0.594 to 1, while SF-6D
19 ranged from 0.3 to 0.89. The median baseline values have different locations in their
20 respective scoring ranges (Fig 1). The EQ-5D-3L showed 1.0% floor effect and 11.8% ceiling
21 effect, while SF-6D had 0.0% floor and ceiling effects.
22
23
24
25
26
27

28 *Agreement, differences and correlation between the two utility scores*

29 There was no significant difference (p=0.588) between overall mean scores of the two utility
30 instruments as well as different age categories (p>0.05) and different education statuses
31 (p>0.05). Compared to SF-6D, the mean EQ-5D-3L utility scores were significantly higher
32 among males (p=0.016), which included those who were employed (p<0.001), had no
33 comorbidities (p<0.001) and had CKD stages earlier than stage V (p=0.042 and 0.015). The
34 mean SF-6D utility scores were significantly higher among females (p=0.045), which
35 included those who were not employed (p<0.001), had comorbidities (p=0.028) and were on
36 dialysis (p<0.001) (Table 1). The standard deviation of the EQ-5D-3L was considerably
37 larger than that of the SF-6D among all sub groups.
38
39
40
41
42
43
44
45

46 Significant proportion of the patients reported “no problem” in any of the EQ-5D-3L
47 dimensions than the SF-6D. However, fewer patients reported “extreme problems” in the EQ-
48 5D-3L than in the SF-6D (Tables 2 and 3). Patients reported different results for the related
49 dimensions of the two MAUIs (Tables 2 and 3). Nearly half of the patients reported “no
50 problem” in Mobility domain of the EQ-5D-3L, while only 0.7% reported “no problems”
51 with the physical functioning of the SF-6D. Nearly quarter (23.8%) of patients reported “no
52 problems” for the anxiety / depression dimension in the EQ-5D-3L, whereas only 0.6%
53 reported the same for the mental health dimension of the SF-6D.
54
55
56
57
58
59

1
2
3 The correlation between EQ-5D-3L and SF-6D was 0.408, which was statistically significant
4 at $p < 0.001$ level (Figure 2). Regarding the correlation between different domains of the two
5 instruments, according to the Guilford's criteria, low level of correlation (0.2-0.4) was seen
6 between Mobility and Physical functioning (0.3249), Social functioning (0.3672) and Pain
7 (0.3607); between Usual activities and Social functioning (0.3152); between Pain/ discomfort
8 and Physical functioning (0.3123), Pain (0.3567) and Vitality (0.3420); between Anxiety/
9 depression and Social functioning (0.3656), Pain (0.3495) and Vitality (0.3136). Also,
10 moderate correlation (0.4-0.6) was evident between Pain / discomfort and Social functioning
11 (0.4090). All other domains were poorly correlated between the two instruments (Table 4).
12
13
14
15
16
17

18 *Discrimination*

19
20 With both MAUIs, utility scores decreased with increasing severity (as measured by CKD
21 stage) (Table 5). In both MAUIs, the utility differences across CKD stages were statistically
22 significant ($p < 0.05$; ANOVA) indicating good discrimination. Figure 3 indicated the box-
23 plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility
24 scores for CKD stage. Furthermore, the calculated effect size between CKD early stage and
25 stage IV was 0.071 and 0.141 for EQ-5D-3L and SF-6D respectively. The highest effect size
26 was observed between CKD stage V and dialysis group, which was 0.807 for EQ-5D-3L and
27 1.098 for SF-6D.
28
29
30
31
32
33

34 *Test-retest reliability*

35
36 The test re-test Spearman's correlations was more than 0.9 for both the instruments indicating
37 good test re-test reliability.
38
39
40

41 **Discussion**

42
43 The findings from this study are a comparison of utility scores arising from the EQ-5D-3L
44 and SF-6D in CKD patients. Comparisons between utility scores of EQ-5D-3L and SF-6D are
45 scarce in the literature. Moreover, this is the first such comparison among CKD patients.
46 According to the current study, the correlation between the scores was moderate. Both tools
47 were able to discriminate advancement of CKD stages. Effect size, which denoted the
48 discriminating ability of different CKD stages, is highest when disease condition is advanced
49 and the highest effect size was seen in SF-6D.
50
51
52
53
54
55

56 At present, there is no consensus on the methodology to compare the utility scores of
57 different MAUIs. The present study adopted the methodologies used by Kularatna et al.
58
59

1 (2017) and Lamers et al. (2006) (19, 34). Only one time assessment of the utilities was done
2 in the present study. Thus, the responsiveness of the two instruments was not assessed.
3
4 Though Sri Lankan EQ-5D-3L utility scores are available (18), yet we used the UK utility
5 scores for the EQ-5D-3L (32) because of the unavailability of comparable Sri Lankan SF-6D
6 utility scores. This is an accepted method of calculating the utility scores in the absence of
7 country specific utilities. Two studies conducted in Netherlands (24) and Italy (21),
8 comparing the utility scores of the two instruments, had used the UK derived EQ-5D-3L and
9 SF-6D utility scores.
10
11
12
13
14

15 The present study did not find any difference ($p=0.588$) between the overall mean scores of
16 the two utility instruments. This was similar to a study conducted among a group of
17 HIV/AIDS patients (28), but different to several other studies available in the literature where
18 different results have been reported. Significantly higher utility values for EQ-5D-3L were
19 found among general population (29), cardio-vascular disease patients (19), rheumatoid
20 arthritis patients (21) and patients with stable angina (16). However, in a study conducted
21 among a group of patients with psychiatric disorders, significantly higher utility values were
22 obtained for SF-6D instrument (24). These varying results could be due to different recall
23 periods of the two instruments. EQ-5D-3L assessed the health status of the day of instrument
24 administration while SF-6D, which was derived from SF-36, assessed the health status of the
25 past 30 days.
26
27
28
29
30
31
32
33

34
35 Though overall ceiling and floor effects of both instruments were small, significant ceiling
36 effect was evident in the EQ-5D-3L. This was consistent with several other studies conducted
37 elsewhere (16, 19, 44-46). This is mainly due to the fact that the EQ-5D-3L has limited
38 response levels and the five level newer version of EQ-5D expected to improve the properties
39 of the three-level in terms of reduced ceiling effects, increased reliability and improved
40 ability to discriminate between different levels of health (47).
41
42
43
44
45

46 Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using
47 ANOVA and effect size. In both MAUIs, the utility differences across CKD stages were
48 statistically significant ($p<0.05$; ANOVA) indicating good discrimination. However, the
49 effect size was small for both the tools until the dialysis stage. At the dialysis stage, the effect
50 size is large and this was highest in the SF-6D instrument. It could be due to the fact that
51 CKD is considered asymptomatic until the later stages of the disease (48, 49) so that the
52 instruments cannot discriminate different stages. According to a recent study conducted by
53
54
55
56
57
58
59
60

1 Jesky et al. (2016), the EQ-5D-3L utility scores of the adjacent pre-dialysis CKD stages were
2 not found to be statistically significant (50).
3
4
5

6 **Limitations**

7 Some of the information related to QOL in SF-36 is considered to be sensitive in nature and
8 the fact that this information was obtained utilising an interviewer-administered questionnaire
9 could have led to some under-reporting in the assessment of QOL though many measures
10 were taken to minimize this issue. Our study was a cross-sectional study, thus we could not
11 assess which utility scores of the two instruments change over time.
12
13
14
15

16 **Conclusions**

17 The correlation between the scores was moderate. Both tools were able to discriminate
18 advancement of CKD stages. Effect size, which denoted the discriminating ability of the
19 different CKD stages, is highest when disease condition is advanced and the highest effect
20 size was seen in SF-6D. Findings indicate that both tools cover different spaces in health.
21 Thus, although there was a moderate correlation between the measures, both scores cannot be
22 used interchangeably while assessing QALY during cost utility analysis.
23
24
25
26
27
28
29

30 *Ethics approval and consent to participate*

31 The study is in accordance with Helsinki Declaration. The study protocol has been approved
32 by the Ethics Committee of Colombo Medical Faculty. Permission was obtained from the
33 Provincial Director of Health Service, to assess the CKD register available at his office.
34 Participants gave their informed consent.
35
36
37
38
39

40 *Funding*

41 This study was funded by the Ministry of Health, Sri Lanka
42
43
44

45 *Competing interests:* None declared.
46
47

48 *Patient consent :* Obtained.
49
50

51 *Author Contributions*

52 SK and SS : Research idea, study design, statistical analysis and drafting of the manuscript.
53 NG²: study design, data analysis/interpretation. NG¹: participated in study design, data
54 interpretation and supervision. All authors read and approved the final manuscript.
55
56
57
58
59

1 *Provenance and peer review* : Not commissioned; externally peer reviewed.

2
3
4 *Data sharing statement*

5 The datasets used and/or analysed during the current study available from the corresponding
6 author on reasonable request.
7
8
9

10
11 **References**

- 12 1. Senanayake S, Gunawardena N, Palihawadana P, Kularatna S, Peiris T. Validity and reliability
13 of the Sri Lankan version of the kidney disease quality of life questionnaire (KDQOL-SF™).
14 Health and quality of life outcomes. 2017;15(1):119.
- 15 2. Schieppati A, Remuzzi G. Chronic renal diseases as a public health problem: epidemiology,
16 social, and economic implications. *Kidney International*. 2005;68:S7-S10.
- 17 3. Evaluation IfHMA. Global Burden of Disease 2010 Arrow Diagram: Institute for Health Metrics
18 and Evaluation; 2010 [Available from: <http://vizhub.healthdata.org/irank/arrow.php>.
19
- 20 4. Rebollo-Rubio A, Morales-Asencio JM, Pons-Raventos ME, Mansilla-Francisco JJ. Review of
21 studies on health related quality of life in patients with advanced chronic kidney disease in
22 Spain. *Nefrologia*. 2015;35(1):92-109.
- 23 5. Mujais SK, Story K, Brouillette J, Takano T, Soroka S, Franek C, et al. Health-related quality
24 of life in CKD patients: correlates and evolution over time. *Clinical Journal of the American*
25 *Society of Nephrology*. 2009;4(8):1293-301.
- 26 6. Tsai Y-C, Hung C-C, Hwang S-J, Wang S-L, Hsiao S-M, Lin M-Y, et al. Quality of life
27 predicts risks of end-stage renal disease and mortality in patients with chronic kidney disease.
28 *Nephrology dialysis transplantation*. 2009;25(5):1621-6.
- 29 7. DeOreo PB. Hemodialysis patient-assessed functional health status predicts continued survival,
30 hospitalization, and dialysis-attendance compliance. *American Journal of Kidney Diseases*.
31 1997;30(2):204-12.
- 32 8. Gussous Y, Theologis AA, Demb JB, Tangtiphaiboontana J, Berven S. Correlation Between
33 Lumbopelvic and Sagittal Parameters and Health-Related Quality of Life in Adults With
34 Lumbosacral Spondylolisthesis. *Global spine journal*. 2018;8(1):17-24.
- 35 9. Howren MB. Quality-Adjusted Life Years (QALYs). *Encyclopedia of Behavioral Medicine*:
36 Springer; 2013. p. 1605-6.
- 37 10. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *International journal of*
38 *technology assessment in health care*. 1989;5(4):559-75.
- 39 11. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the*
40 *economic evaluation of health care programmes*: Oxford university press; 2015.
- 41 12. Brooks R. EuroQol: the current state of play. *Health policy*. 1996;37(1):53-72.
- 42 13. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from
43 the SF-36. *Journal of health economics*. 2002;21(2):271-92.
- 44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59

- 1 14. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and
2 single-attribute utility functions for the health utilities index mark 3 system. *Medical care*.
3 2002;40(2):113-28.
- 4 15. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification
5 systems. *Pharmacoeconomics*. 1995;7(6):490-502.
- 6 16. Wu J, Han Y, Zhao F-L, Zhou J, Chen Z, Sun H. Validation and comparison of EuroQoL-5
7 dimension (EQ-5D) and Short Form-6 dimension (SF-6D) among stable angina patients. *Health
8 and quality of life outcomes*. 2014;12(1):156.
- 9 17. Stolk EA, Oppe M, Scalone L, Krabbe PF. Discrete choice modeling for the quantification of
10 health states: the case of the EQ-5D. *Value in Health*. 2010;13(8):1005-13.
- 11 18. Kularatna S, Whitty JA, Johnson NW, Jayasinghe R, Scuffham PA. Valuing EQ-5D health
12 states for Sri Lanka. *Quality of Life Research*. 2015;24(7):1785-93.
- 13 19. Kularatna S, Byrnes J, Chan YK, Ski CF, Carrington M, Thompson D, et al. Comparison of the
14 EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with
15 cardiovascular disease. *Quality of Life Research*. 2017;26(12):3399-408.
- 16 20. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and
17 preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life
18 research*. 2011;20(10):1727-36.
- 19 21. Salaffi F, Carotti M, Ciapetti A, Gasparini S, Grassi W. A comparison of utility measurement
20 using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid
21 arthritis. *Clinical and experimental rheumatology*. 2011;29(4):661-71.
- 22 22. Kwakkenbos L, Franssen J, Vonk MC, Becker ES, Jeurissen M, van den Hoogen FH, et al. A
23 comparison of the measurement properties and estimation of minimal important differences of
24 the EQ-5D and SF-6D utility measures in patients with systemic sclerosis. *Clinical and
25 experimental rheumatology*. 2013;31(2 Suppl 76):50-6.
- 26 23. Rowen D, Young T, Brazier J, Gaugris S. Comparison of generic, condition-specific, and
27 mapped health state utility values for multiple myeloma cancer. *Value in Health*.
28 2012;15(8):1059-68.
- 29 24. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D
30 and SF-6D utilities in mental health patients. *Health economics*. 2006;15(11):1229-36.
- 31 25. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired
32 people as estimated by three different utility measures Un comparación de la calidad de vida de
33 personas con trastornos auditivos estimada por tres diferentes medidas de utilidad. *International
34 Journal of Audiology*. 2005;44(3):157-63.
- 35 26. Barton GR, Bankart J, Davis AC, Summerfield QA. Comparing utility scores before and after
36 hearing-aid provision. *Applied health economics and health policy*. 2004;3(2):103-5.
- 37 27. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary
38 heart disease. *Health and Quality of Life Outcomes*. 2006;4(1):20.

- 1 28. Stavem K, Frøland SS, Hellum KB. Comparison of preference-based utilities of the 15D, EQ-
2 5D and SF-6D in patients with HIV/AIDS. *Quality of Life Research*. 2005;14(4):971-80.
- 3
- 4 29. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D
5 based on hypothetical preferences in a general population. *Health economics*.
6 2005;14(11):1169-89.
- 7
- 8 30. McDonough CM, Tosteson AN. Measuring preferences for cost-utility analysis.
9 *Pharmacoeconomics*. 2007;25(2):93-106.
- 10
- 11 31. Senanayake S, Gunawardena N, Palihawadana P, Bandara P, Haniffa R, Karunaratna R, et al.
12 Symptom burden in chronic kidney disease; a population based cross sectional study. *BMC*
13 *nephrology*. 2017;18(1):228.
- 14
- 15 32. Dolan P. Modeling valuations for EuroQol health states. *Medical care*. 1997;35(11):1095-108.
- 16
- 17 33. Ramos-Goni JM, Rivero-Arias O. eq5d: A command to calculate index values for the EQ-5D
18 quality-of-life instrument. *Stata Journal*. 2011;11(1):120.
- 19
- 20 34. Lamers L, Bouwmans C, van Straten A, Donker M, Hakkaart L. Comparison of EQ-5D and
21 SF-6D utilities in mental health patients. *Health economics*. 2006;15(11):1229-36.
- 22
- 23 35. Harrison M, Davies L, Bansback N, McCoy M, Verstappen S, Watson K, et al. The
24 comparative responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory
25 arthritis. *Quality of Life Research*. 2009;18(9):1195.
- 26
- 27 36. Tredoux C, Durrheim K. Numbers, hypotheses and conclusions: A course in statistics for the
28 social sciences. Cape Town: University of Cape Town Press[Links]. 2002.
- 29
- 30 37. Levey AS, Stevens LA, Coresh J. Conceptual model of CKD: applications and implications.
31 *American journal of kidney diseases*. 2009;53(3):S4-S16.
- 32
- 33 38. Seliger SL, Zhan M, Walker LD, Fink JC. Chronic kidney disease adversely influences patient
34 safety. *Journal of the American Society of Nephrology*. 2008;19(12):2414-9.
- 35
- 36 39. Stevens PE, Levin A. Evaluation and management of chronic kidney disease: synopsis of the
37 kidney disease: improving global outcomes 2012 clinical practice guideline. *Annals of internal*
38 *medicine*. 2013;158(11):825-30.
- 39
- 40 40. Tseng C-L, Lafrance J-P, Lu S-E, Soroka O, Miller DR, Maney M, et al. Variability in
41 estimated glomerular filtration rate values is a risk factor in chronic kidney disease progression
42 among patients with diabetes. *BMC nephrology*. 2015;16(1):34.
- 43
- 44 41. Kularatna S, Byrnes J, Chan YK, Ski CF, Carrington M, Thompson D, et al. Comparison of the
45 EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with
46 cardiovascular disease. *Quality of life research : an international journal of quality of life*
47 *aspects of treatment, care and rehabilitation*. 2017;26(12):3399-408.
- 48
- 49 42. Cohen J. A power primer. *Psychological bulletin*. 1992;112(1):155.
- 50
- 51 43. Litwin MS, Fink A. How to measure survey reliability and validity: Sage; 1995.
- 52
- 53 44. Zhao Fl, Yue M, Yang H, Wang T, Wu Jh, Li SC. Validation and comparison of EuroQol and
54 short form 6D in chronic prostatitis patients. *Value in Health*. 2010;13(5):649-56.
- 55
- 56
- 57
- 58
- 59
- 60

- 1 45. Cunillera O, Tresserras R, Rajmil L, Vilagut G, Brugulat P, Herdman M, et al. Discriminative
2 capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health
3 survey. *Quality of Life Research*. 2010;19(6):853-64.
- 4 46. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess
5 their ceiling effects in the US general population. *Value in Health*. 2006;9(4):262-71.
- 6 47. Janssen M, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement
7 properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-
8 country study. *Quality of Life Research*. 2013;22(7):1717-27.
- 9 48. de Boer IH, Olan GA, Patel UD. Screening for chronic kidney disease. *Ann Intern Med*.
10 2013;158(5 Pt 1):362-3.
- 11 49. Hummel A. Chronic renal failure, a frequent and asymptomatic pathology. *Soins; la revue de*
12 *reference infirmiere*. 2010(745):24-6.
- 13 50. Jesky MD, Dutton M, Dasgupta I, Yadav P, Ng KP, Fenton A, et al. Health-related quality of
14 life impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic
15 kidney disease: a prospective observational study. *PloS one*. 2016;11(11):e0165675.
- 16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 : Demographic distribution of the sample by the EQ-5D-3L and the SF-6D utility scores

p value significant<0.05, paired t test

Variable	N (%)	EQ-5D-3L utility mean (SD)	Sf-6D utility mean (SD)	p value [#]
All sample	1096	0.540 (0.35)	0.534 (0.09)	0.588
Sex				
Male	686 (62.6)	0.561 (0.34)	0.532 (0.10)	0.016
Female	410 (37.4)	0.505 (0.37)	0.539 (0.09)	0.045
Age (years)				
Less than 20	07 (0.6)	0.570 (0.44)	0.455 (0.09)	0.486
20-40	45 (4.1)	0.591 (0.34)	0.536 (0.09)	0.259
41-60	562 (51.3)	0.555 (0.35)	0.540 (0.10)	0.282
More than 60	482 (44.0)	0.517 (0.35)	0.529 (0.08)	0.440
Education status				
No formal education	81 (7.4)	0.448 (0.42)	0.508 (0.09)	0.154
5 Grade	413 (37.7)	0.529 (0.35)	0.536 (0.09)	0.681
6-11 Grade	377 (34.4)	0.556 (0.34)	0.533 (0.10)	0.159
GCE O/L passed	190 (17.3)	0.554 (0.34)	0.540 (0.09)	0.513
GCE A/L passed	35 (3.2)	0.618 (0.34)	0.540 (0.09)	0.267
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)	<0.001
Comorbidities				
Present	778 (71.0)	0.505 (0.36)	0.532 (0.09)	0.028
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)	0.042
Stage IV	614 (58.1)	0.566 (0.42)	0.536 (0.09)	0.015
Stage V	151 (14.3)	0.467 (0.42)	0.523 (0.08)	0.076
Dialysis	38 (3.6)	0.126 (0.39)	0.432 (0.07)	<0.001

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

	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)

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

	Physical functioning (%)	Role limitation (%)	Social functioning (%)	Pain (%)	Mental health (%)	Vitality (%)
1 ^a	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 ^b	319 (29.1)	NA	NA	240 (21.9)	NA	NA

^a No problems^b Severe problems**Table 4 : Correlation between the EQ-5D-3L and the SF-6D**

	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression
Physical functioning	0.3249*	0.2644*	0.2347*	0.3123*	0.2988*
Role limitation	0.0145	-0.0021	0.2615*	0.0099	0.0875*
Social functioning	0.3672*	0.2811*	0.3152*	0.4090*	0.3656*
Pain	0.3607*	0.2258*	0.2459*	0.3567*	0.3495*
Mental health	0.1770*	0.1975*	0.1771*	0.1387*	0.1401*
Vitality	0.2242*	0.0889*	0.1629*	0.3420*	0.3136*

Table 5 : Discrimination across clinical severity groups

CKD stage	EQ-5D-3L				SF-6D			
	N	Mean (SD)	Median	ES	N	Mean (SD)	Median	ES
Early stage	254 (24.0)	0.588 (0.30)	0.656		254 (24.0)	0.551 (0.10)	0.570	
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

Between CKD stage utility differences are significant (<0.001) within EQ-5D-3L and SF-6D (ANOVA; P<0.05)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

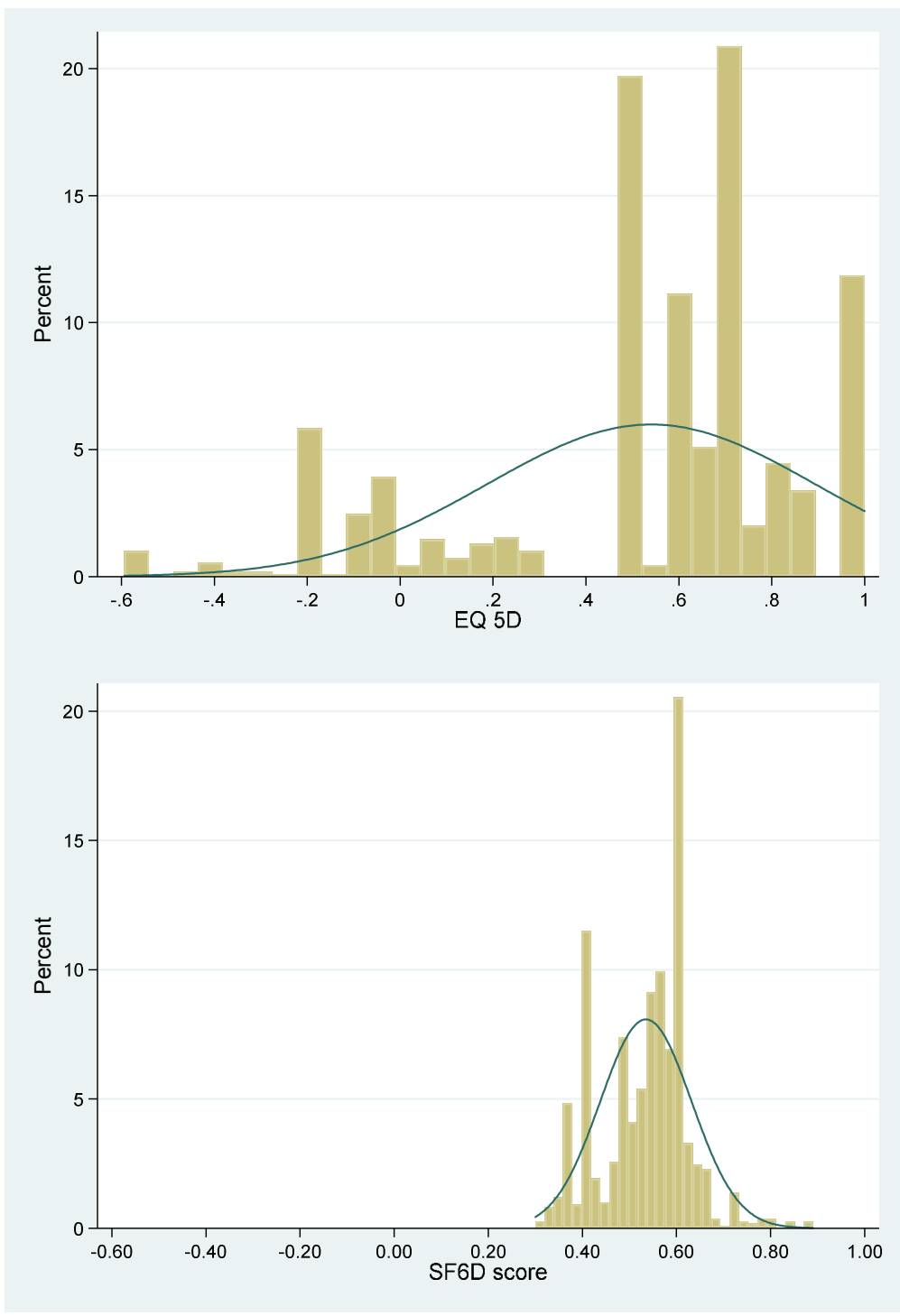


Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

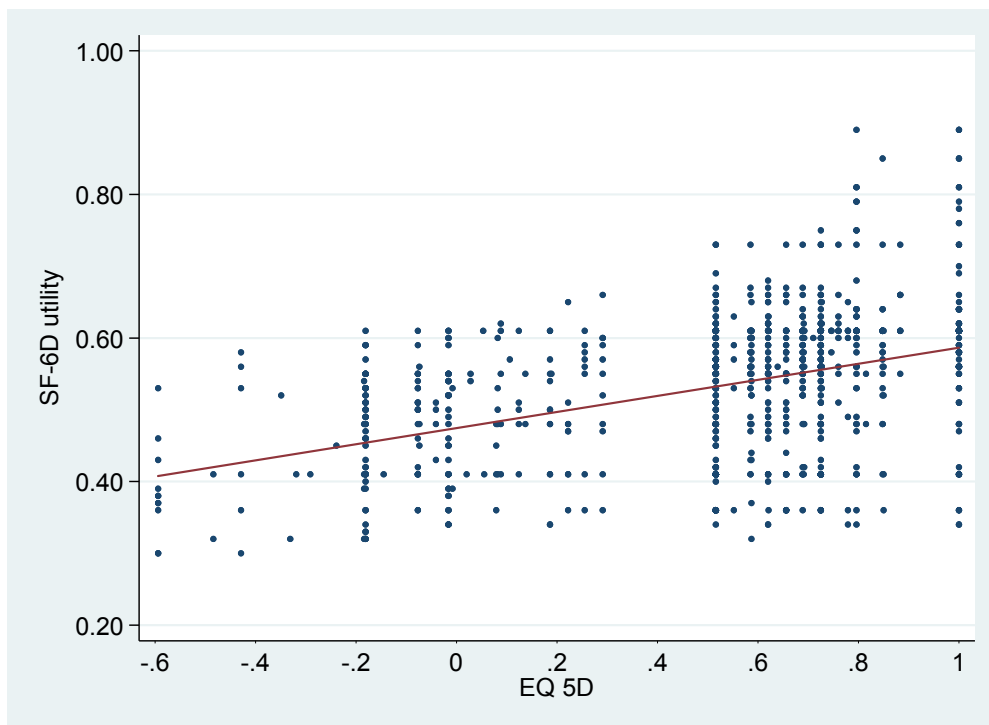


Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

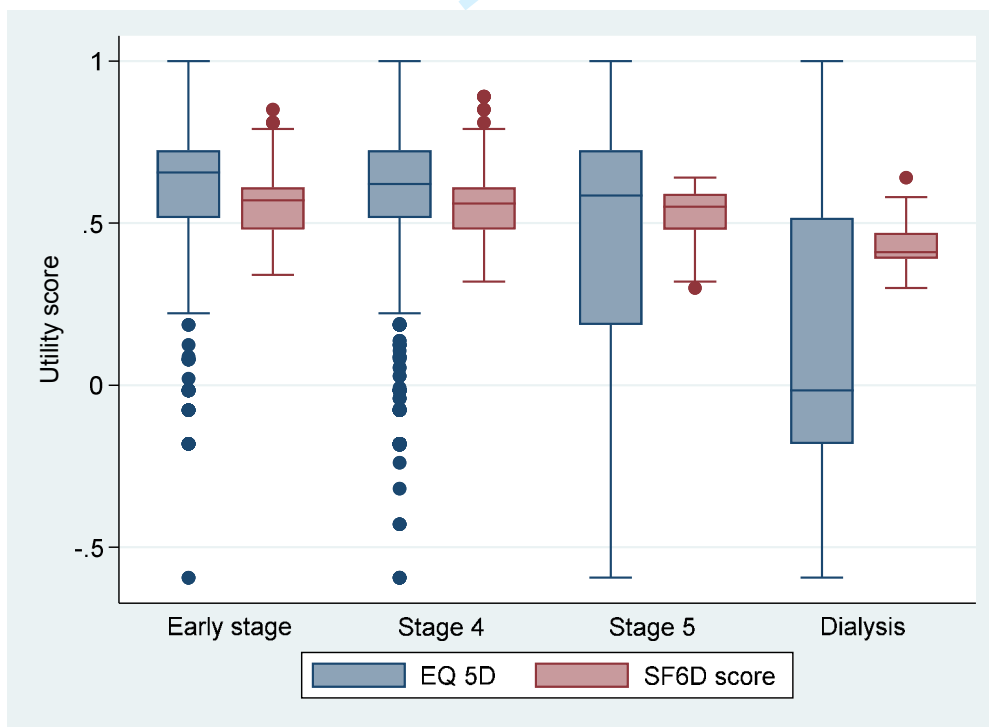


Fig 3: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD stage

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

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2018-024854.R1
Article Type:	Research
Date Submitted by the Author:	22-Sep-2018
Complete List of Authors:	Kularatna, Sanjeewa; Australian Centre For Health Services Innovation, Queensland University of Technology Senanayake, Sameera; Australian Centre For Health Services Innovation, Queensland University of Technology Gunawardena, Nalika; World Health Organization Graves, Nicholas; Australian Centre For Health Services Innovation, Queensland University of Technology
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health economics, Health services research, Renal medicine
Keywords:	Cost utility analysis, Quality-Adjusted Life Years (QALYs), utility, Chronic Kidney Disease (CKD)

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

**1 Comparison of the EQ-5D-3L and the SF-6D (SF-36) contemporaneous utility scores in
2 patients with Chronic Kidney Disease in Sri Lanka; a cross sectional survey**

3
4
5 Sanjeewa Kularatna¹, Sameera Senanayake¹, Nalika Gunawardena², Nicholas Graves¹

6
7 ¹Australian Centre for Health Services Innovation, Queensland University of Technology

8 ² World Health Organization, Sri Lanka

9
10 **Corresponding author;**

11 Sanjeewa Kularatna - *sanjeewa.kularatna@qut.edu.au*

12
13 **Word count : 3277**

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66

Strengths and limitations of this study

- The response rate of the study is very high.
- Both tools used in the study (EQ 5D 3L and SF 36) have been previously validated to the Sri Lankan setting before.
- Data collectors were experienced for many local and international studies done among CKD patients 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 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 minimize this issue.

67 **Introduction**

68 Chronic kidney disease (CKD) is a substantial public health problem with adverse
69 psychological, physical and economic outcomes. The burden of CKD is increasing
70 globally(1). World Health Report (2002) and Global Burden of Disease (GBD) project stated
71 that the diseases of the kidney contribute much to the global disease burden with
72 approximately 850,000 deaths every year globally (2). Furthermore, according to GBD study
73 conducted in 2010, of the top causes of DALY, CKD is ranked 29th globally, 23rd in South
74 East Asia and 14th in Sri Lanka (3). Due to the progressive and disabling nature of CKD, it
75 poses a substantial impact on the quality of life (QOL) of individuals. It is important to
76 measure QOL indicators for the management of chronic kidney disease patients. Several
77 studies have demonstrated a relationship between reduced QOL and increased morbidity and
78 mortality (4-7).

79
80 World over, the importance of including QOL indicators in the clinical management of
81 patients has been highlighted. This has come to the limelight after several studies
82 demonstrated the strong relationship between reduced QOL and increased morbidity and
83 mortality (5, 8). Meantime, economic evaluation has become increasingly popular among
84 researchers and policy makers during resource allocation in recent years. Due to the
85 relationship between QOL and clinical outcome, during the recent years, QOL has become an
86 important health outcome in economic evaluations. In cost utility analysis (CUA), a method
87 of economic evaluation, outcomes are usually measured in Quality-Adjusted Life Years
88 (QALYs), which is a measure of QOL.

89
90 The concept of QALYs was developed in the 1970s. It can measure the changes of an
91 individual's quality and quantity of life and can also aggregate these improvements across
92 individual (9, 10). The change in the quality of life in QALY is measured using a set of
93 weights, called utilities, which reflect different health states. For all possible health states,
94 utilities should be measured on a scale where 1 refers to best imaginable health and 0 refers
95 to death (11). Measuring utilities for different health states is complex and time-consuming.
96 Thus, Multi Attribute Utility Instruments (MAUI) such as EuroQol-5D (EQ-5D-3L) (12),
97 Short Form-6D (SF-6D) (13) or the Health Utility Index (HUI) (14, 15) are used to define
98 different health states. The utility scores for different health states in different instruments are
99 derived from methods such as Standard Gambling method (16), Discrete Choice Experiments
100 (17) and Time Trade-Off experiments (18). EQ-5D-3L is the most widely used utility
101 instrument at present (19). EQ-5D-5L, which is the newer version of EQ-5D, has also been
102 developed and tested recently (20).

1 103
2
3 104 Since all the MAUIs aim at measuring the health state of individuals, all the instruments
4 105 should generate the same utility value for a particular state of health. However, the evidence
5 106 indicates that there is an essential difference in the utility scores for a particular health state
6 107 between different instruments (19, 21-29). This, in turn, indicates that the choice of the
7 108 MAUI used may adversely influence the results of CUA and thereby the decision-making
8 109 process (30). Furthermore, for incremental analyses, use of different MAUIs may lead to
9 110 different results regarding the magnitude, direction or significance of any change in health-
10 111 related quality of life measure.
11
12
13
14
15
16

17 112
18 113 Though the differences between MAUIs have been evaluated in many disease conditions (19,
19 114 21, 22, 24), there are no evidence in the literature comparing MAUIs using patients with
20 115 CKD. The aim of our study is to compare contemporaneous EQ-5D-3L and SF-6D utility
21 116 scores in patients with CKD. Results may be useful for researchers selecting a generic MAUI
22 117 to estimate utilities for use in economic modelling of treatments for CKD.
23
24
25
26

27 118

27 119 **Methods**

28 120 *Participant selection*

29 121 A population-based descriptive cross-sectional study was conducted in the district of
30 122 Anuradhapura in the North Central Province (NCP) of Sri Lanka between Septembers to
31 123 December 2015. The study population consisted of 1162 confirmed CKD patients, calculated
32 124 using the appropriate formula (31), who were over 18 years old with documented evidence of
33 125 CKD living in the Anuradhapura district. The diagnosis of CKD was made if the Glomerular
34 126 Filtration Rate (GFR) was less than 60 ml/min per 1.73m² of body surface area in two
35 127 measurements made three months apart.
36
37
38
39
40
41

42 128

43 129 The inclusion criteria were patients above 18 years of age and those who were diagnosed as
44 130 having CKD by a specialist nephrologist or a consultant physician. Presence of evidence of
45 131 such diagnosis was made by way of diagnosis cards, clinic records or any other record issued
46 132 by a specialist nephrologist, a consultant physician or a government hospital. Patients who
47 133 had previous renal transplantation, who were unable to provide rational information due to
48 134 any cause (e.g. mental retardation) and who were critically ill were excluded from the study.
49
50
51
52

53 135

54 136 The study instrument was an interviewer-administered questionnaire to gather information on the
55 137 socio-demographic information, CKD related information, EQ-5D-3L and SF 36.
56
57

58 138

1 139 Five Public Health Inspectors working in the CKD unit in the North Central Province were used
2 140 for the data collection and all have been working in the unit for more than 5 years and they had
3 141 experience in functioning as data collectors for many local and international studies done among
4 142 CKD patients in the NCP. Data collection was mostly done on weekdays considering the fact
5 143 that most of the study units were expected to be at home, since most are employed in the informal
6 144 sector. The data collectors assessed the eligibility of patients by reviewing their clinical
7 145 records. Informed consent was obtained from those who were eligible for participation in the
8 146 study.

9 147
10 148 The study was conducted in all nineteen Medical Officer of Health (MOH) areas of the
11 149 Anuradhapura district. The number of participants to be included from each MOH area was
12 150 based on probability proportionate to the size of CKD patients registered in each of the MOH
13 151 areas. The required number of participants from each MOH area was selected using simple
14 152 random sampling method. The population-based CKD register – which records the patients
15 153 with a confirmed diagnosis of CKD from renal clinics in hospitals of the NCP since 2003 –
16 154 was used as the sampling frame. The register was obtained from the office of the Provincial
17 155 Director of Health Services (32).

18 156

19 157 ***Calculation of utility scores***

20 158 Currently, there is no algorithm based on preferences of the Sri Lankan public to score the
21 159 SF-6D on a utility scale. Therefore, the UK algorithm was used for this purpose (13).

22 160 Though Sri Lankan EQ-5D-3L utility scores are available (18), the UK utility scores were
23 161 used for the EQ-5D-3L (33) because of the unavailability of comparable Sri Lankan SF-6D
24 162 utility scores as mentioned earlier. This allowed the comparison of utility scores from the
25 163 same country.

26 164

27 165 The EQ-5D-3L instrument contains five domains; mobility, self-care, usual activities, pain /
28 166 discomfort and anxiety / depression. Each domain has one item and each item has three
29 167 levels: one denoting no problems and three denoting severe problems (12). Thus, EQ-5D-3L
30 168 has mutually exclusive 243 different health states.

31 169

32 170 SF-6D is derived from SF-36, SF-12 Version 1 and SF-12 Version 2. The current study
33 171 utilised SF-36 for the data collection. SF-36 includes 36 items that measure eight domains;
34 172 role limitations caused by physical problems (4 items), physical function (10 items), role
35 173 limitations caused by emotional problems (3 items), pain (2 items), social function (2 items),
36 174 general health perceptions (5 items), emotional well-being (5 items) and energy / fatigue (4

175 items). Questions have different answer options which range from two to six. While scoring,
176 each question is scored in a scale ranging from 0 (worst health) to 100 (best health). All items
177 in a domain are summed up and averaged to give an average score for each domain which
178 ranges from 0 (worst health) to 100 (best health). To calculate the utility scores of the SF-6D,
179 11 items are used covering six domains; physical functioning, role limitation, social
180 functioning, pain, mental health and vitality (13).

181

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

186

187 The EQ-5D utility scores range from -0.59, 0=being dead; negative values represent health
188 status considered worse than “dead”, to 1.00 which indicate best imaginable health. The SF-
189 6D utility scores ranged from 0.296 which indicate severely impaired levels in all dimensions
190 to 1.0 which indicates no difficulty in any dimensions.

191

192 ***Data analysis***

193 STATA 15.1 software was used for the analysis. Distribution of the socio-demographic
194 characteristics of the study population was compared with their mean utility scores.
195 Normality of the two distributions were assessed using Kolmogorov-Smirnov and Shapiro-
196 Wilk tests. Wilcoxon signed-rank test was used to assess the difference between the two
197 instruments in each socio-demographic class (35). Histograms were plotted for the two
198 utility values distribution. Floor effects and ceiling (proportion of patients with the highest
199 and lowest possible scores respectively) were calculated for the EQ-5D and SF-6D. Ceiling
200 and floor effects were considered small if $\leq 15\%$ of patients occupy the best or worst health
201 states, but they were considered serious if $> 15\%$ of patients occupy these states (36).

202

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

208

209 ***Agreement and differences***

1 210 The Wilcoxon Signed-Ranks Test was used to assess the overall difference between the EQ-5D-
2 211 3L and SF-6D utility scores and the difference of the utility scores according to different
3 212 socio-demographic and disease related features. Furthermore, the distribution of the
4 213 responses to the different domains of the two instruments was tabulated to present the
5 214 agreement and the differences between the two instruments. Bland-Altman plot was also
6 215 used to assess the proportional error and the limit of agreement (37).
7
8
9
10

11 216

12 217 ***Intra Class Correlation (ICC)***

13 218 The dimensions of the two instruments were compared using ICC. The related dimensions
14 219 between the two MAUIs are role limitation (SF-6D)/usual activities (EQ-5D-3L), physical
15 220 functioning (SF-6D)/mobility and self-care (EQ-5D-3L), pain (SF-6D)/pain and discomfort
16 221 (EQ-5D-3L), social functioning (SF-6D)/usual activities (EQ-5D-3L) and mental health (SF-
17 222 6D)/anxiety and depression (EQ-5D-3L). The vitality dimension of the SF-6D did not have
18 223 any related dimension with the EQ-5D-3L. The magnitude of the correlation coefficients
19 224 were interpreted according to Guilford's criteria (38).
20
21
22
23
24
25

26 225

27 226 ***Discrimination***

28 227 It is important that MAUIs can discriminate correctly among groups of different severity as
29 228 MAUIs are meant to measure change in QOL due to health improvement in the condition of
30 229 interest.
31
32
33

34 230

35 231 Glomerular Filtration Rate (GFR) is the most important indicator of kidney function of
36 232 patients with CKD (39). Studies have shown that decreased GFR is associated with infection,
37 233 impaired cognitive and physical function as well as threats to patient safety (40). Though
38 234 classifications exist to classify stages of CKD, it is evident that at present most of the clinical
39 235 decision making in CKD is solely based on GFR base classification (41, 42). Depending on
40 236 the GFR value, CKD is categorised into five stages; stage I to stage V. For analytical
41 237 purposes, the CKD stages I to III were categorised as "early stage" in the present study. It is
42 238 expected that with advanced stages of the disease, the utility scores should be lower than the
43 239 early stages.
44
45
46
47
48
49

50 240

51 241 Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using the
52 242 non-parametric test, Kruskal–Wallis, and effect size. The instrument's ability to discriminate
53 243 between two adjacent stages was estimated by calculating the effect size. The effect size was
54 244 calculated by dividing the mean difference of two adjacent CKD stages by the standard
55 245 deviation of the milder of the two CKD stages (23, 43). Large effect size indicates better
56
57
58
59
60

1 246 discriminating ability of the instrument. The effect size was categorised into small (0.2–0.5),
2 247 medium (0.5–0.8) and large (more than 0.8) (44).

3 248

4 249 ***Test-retest reliability***

5 250 To assess the test-retest reliability of the study instrument, within a period of one week, 30
6 251 randomly selected study participants were visited at their households by the data collectors.
7 252 Test re-test reliability of the utility scores of the two instruments was assessed using ICC and
8 253 a value of 0.70 or greater was considered as satisfactory reliability (45).

9 254

10 255 ***Patient and Public Involvement***

11 256 The main stakeholders in the provision of care for the CKD patients such as consultants,
12 257 medical officers working in nephrology units, community leaders and the patients living in
13 258 this area were involved in planning the study. Their concerns were always entertained and
14 259 where feasible their concerns were incorporated into the study. During the data collection,
15 260 stage permission was obtained from the respective local officers. The results of the study was
16 261 communicated to the local level officials such as Medical Officer of Health, Divisional
17 262 Secretariat, Regional Director of Health Services and Provincial Director of Health Services.

18 263

19 264 **Results**

20 265 ***Sample characteristics***

21 266 Out of 1162 participants selected to be included in the study, 66 (5.6%) did not participate in
22 267 the study giving a response rate of 94.4%. The mean age of the study population was 58.4
23 268 years (Standard Deviation (SD) 10.8). There was a preponderance of males among the study
24 269 population (62.6%, N=686). The mean eGFR of the population was 31.8 (SD 20.2)
25 270 ml/min/1.73 m². The mean number of years since diagnosed with CKD was found to be 4.1
26 271 (SD 3.2) years. The majority of participants was in the later stages, stage 4 or beyond, of
27 272 CKD (n=803; 73.2%). 38 participants (3.6%), with stage 5 of the disease and undergoing
28 273 dialysis, were on haemodialysis (Table 1). Chronic Kidney Disease of Unknown origin
29 274 (CKDu) was the cause of the CKD in most of the study population (n=489; 43.7%).

30 275

31 276 ***Distribution of EQ-5D-3L and SF-6D utility scores***

32 277 The mean EQ-5D-3L utility score at baseline was 0.540 compared with 0.534 for the SF-6D
33 278 as summarised in Table 1. The EQ-5D-3L utility score ranged from -0.594 to 1, while SF-6D
34 279 ranged from 0.3 to 0.89. The median baseline values have different locations in their
35 280 respective scoring ranges (Fig 1). The EQ-5D-3L showed 1.0% floor effect and 11.8% ceiling
36 281 effect, while SF-6D had 0.0% floor and ceiling effects.

37 282

283 **Agreement, differences and correlation between the two utility scores**

284 Analyses revealed non- normal distribution of the utility scores of both the instruments, thus
285 Wilcoxon Signed-Ranks test was used to compare the two utility scores. There was
286 significant difference ($p < 0.001$) between overall scores of the two utility instruments.
287 Further the two utility scores were significantly different among males (< 0.001), age more
288 than 40 years groups, those who were educated from grade 5 to General Certificate of
289 Education (GCE) - Ordinary Level, those who were employed, among both who had and
290 didn't have comorbidities, those who didn't have comorbidities, up to stage IV of CKD and
291 among dialysis patients (Table 1). The standard deviation of the EQ-5D-3L was considerably
292 larger than that of the SF-6D among all sub groups.

293
294 Significant proportion of the patients reported “no problem” in any of the EQ-5D-3L
295 dimensions than the SF-6D. However, fewer patients reported “extreme problems” in the EQ-
296 5D-3L than in the SF-6D (Tables 2 and 3). Patients reported different results for the related
297 dimensions of the two MAUIs (Tables 2 and 3). Nearly half of the patients reported “no
298 problem” in Mobility domain of the EQ-5D-3L, while only 0.7% reported “no problems”
299 with the physical functioning of the SF-6D. Nearly quarter (23.8%) of patients reported “no
300 problems” for the anxiety / depression dimension in the EQ-5D-3L, whereas only 0.6%
301 reported the same for the mental health dimension of the SF-6D.

302
303 The correlation between EQ-5D-3L and SF-6D was 0.408, which was statistically significant
304 at $p < 0.001$ level (Figure 2). Regarding the ICC between different domains of the two
305 instruments, according to the Guilford's criteria, low level of correlation (0.2-0.4) was seen
306 between Mobility and Physical functioning (0.381), Mobility and Mental health (0.293),
307 Mobility and Vitality (0.322), Self-care and Physical functioning (0.326), Self-care and Pain
308 (0.330), Self-care and Mental Health (0.323), Usual activities and Physical functioning
309 (0.296), Usual activities and Pain (0.355), Usual activities and Mental health (0.295), Pain/
310 discomfort and Physical functioning (0.382), Pain/ discomfort and Mental health (0.240),
311 Anxiety/ depression and Physical functioning (0.381). Also, moderate correlation (0.4-0.6)
312 was evident between Social functioning and Mobility (0.517), Social functioning and Self-
313 care (0.424), Social functioning and Usual activities (0.464), Social functioning and
314 Pain/discomfort (0.566), Social functioning and Anxiety/depression (0.528), Pain and
315 Mobility (0.475), Pain and Pain/ discomfort (0.482), Pain and Anxiety/depression (0.484),
316 Vitality and Pain/ discomfort (0.475) and Vitality and Anxiety/depression (0.453). All other
317 domains were poorly correlated between the two instruments (Table 4). The Bland-Altman
318 plot showed proportional error and wide limits of agreement (Figure 3).

319

320 Discrimination

321 With both MAUIs, utility scores decreased with increasing severity (as measured by CKD
322 stage) (Table 5). In both MAUIs, the utility differences across CKD stages were statistically
323 significant ($p < 0.05$) indicating good discrimination. Figure 3 indicated the box-plots present
324 the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD
325 stage. Furthermore, the calculated effect size between CKD early stage and stage IV was
326 0.071 and 0.141 for EQ-5D-3L and SF-6D respectively. The highest effect size was observed
327 between CKD stage V and dialysis group, which was 0.807 for EQ-5D-3L and 1.098 for SF-
328 6D.

329

330 Test-retest reliability

331 The test re-test ICC was more than 0.943 in EQ-5D-3L while it was 0.921 in SF 6D,
332 indicating good test re-test reliability in both the instruments.

333

334 Discussion

335 This is the first study to compare the utility scores arising from the EQ-5D-3L and SF-6D in
336 CKD patients. According to the current study, the correlation between the scores was
337 moderate. Both tools were able to discriminate advancement of CKD stages. Effect size,
338 which denoted the discriminating ability of different CKD stages, is highest when disease
339 condition is advanced and the highest effect size was seen in SF-6D. Further, the lowest
340 ceiling effect and the floor effect were seen in SF 6D.

341

342 Evidence indicate that the choice of MAUI (e.g.; EQ 5D or SF6D) has an impact on the
343 results of the cost-utility analysis (46, 47). Sack et al. (2009) compared the results of cost-
344 utility estimates using both EQ 5D and SF 6D. Results indicated contrasting results for the
345 two instruments and authors concluded that the choice of the instrument does matter in cost-
346 utility analysis (46). Thus, from an economic perspective it is important to know the most
347 suitable MAUI to be used among CKD patients.

348

349 At present, there is no consensus on the methodology to compare the utility scores of
350 different MAUIs (19, 35). The present study adopted the methodologies used by Kularatna et
351 at. (2017) and Lamers et al. (2006) (19, 35). Only one time assessment of the utilities was
352 done in the present study. Thus, the responsiveness of the two instruments was not assessed.
353 Though Sri Lankan EQ-5D-3L utility scores are available (18), yet we used the UK utility
354 scores for the EQ-5D-3L (33) because of the unavailability of comparable Sri Lankan SF-6D

1 355 utility scores. This is an accepted method of calculating the utility scores in the absence of
2 356 country specific utilities. Two studies conducted in Netherlands (24) and Italy (21),
3 357 comparing the utility scores of the two instruments, had used the UK derived EQ-5D-3L and
4 358 SF-6D utility scores.
5
6

7 359
8
9 360 The present study did not find any difference ($p=0.588$) between the overall mean scores of
10 361 the two utility instruments. This was similar to a study conducted among a group of
11 362 HIV/AIDS patients (28), but different from other studies available in the literature where
12 363 different results have been reported. Significantly higher utility values for EQ-5D-3L were
13 364 found among general population (29, 48), cardio-vascular disease patients (19), rheumatoid
14 365 arthritis patients (21) and patients with stable angina (16). However, in a study conducted
15 366 among a group of patients with psychiatric disorders, significantly higher utility values were
16 367 obtained for SF-6D instrument (24). These varying results could be due to different recall
17 368 periods of the two instruments. EQ-5D-3L assessed the health status of the day of instrument
18 369 administration while SF-6D, which was derived from SF-36, assessed the health status of the
19 370 past 30 days.
20
21
22
23
24
25
26

27 371
28 372 Though overall ceiling and floor effects of both instruments were small, relatively higher
29 373 ceiling effect was evident in the EQ-5D-3L. This was consistent with several other studies
30 374 conducted elsewhere, where EQ 5D 3L reported a relatively higher ceiling effect compared to
31 375 SF 6D (16, 19, 49-51). This is mainly due to the fact that the EQ-5D-3L has limited response
32 376 levels and the five level newer version of EQ-5D expected to improve the properties of the
33 377 three-level in terms of reduced ceiling effects, increased reliability and improved ability to
34 378 discriminate between different levels of health (52). Further, the current study reported
35 379 relatively lower ceiling effect, for the EQ 5D, compared to results obtained among
36 380 Parkinson's disease (13.5%) and stable angina (15.5%) patients. However, our result was
37 381 higher compared to the ceiling effect observed among patients with systemic sclerosis
38 382 (7.0%). Among many other factors that could contribute to these differences, the level of
39 383 morbidity of a disease is said to be one of the factors which could influence the ceiling effect
40 384 observed in EQ 5D (53). Thus the diseases with lower morbidity are expected to have higher
41 385 ceiling effects.
42
43
44
45
46
47
48
49
50

51 386
52
53 387 Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using
54 388 ANOVA and effect size. In both MAUIs, the utility differences across CKD stages were
55 389 statistically significant ($p<0.05$; ANOVA) indicating good discrimination. However, the
56 390 effect size was small for both the tools until the dialysis stage. At the dialysis stage, the effect
57
58
59
60

1 391 size is large and this was highest in the SF-6D instrument. It could be because CKD is
2 392 considered asymptomatic until the later stages of the disease (54, 55), not allowing the
3 393 instruments to discriminate the different stages. According to a recent study conducted by
4 394 Jesky et al. (2016), the EQ-5D-3L utility scores of the adjacent pre-dialysis CKD stages were
5 395 not found to be statistically significant (56).
6
7
8

9 396

10 397 **Limitations**

11 398 Some of the information related to QOL in SF-36 is considered to be sensitive in nature and
12 399 the fact that this information was obtained utilising an interviewer-administered questionnaire
13 400 could have led to some under-reporting in the assessment of QOL though many measures
14 401 were taken to minimize this issue. Our study was a cross-sectional study, thus we could not
15 402 assess how utility scores of the two instruments change over time.
16
17
18
19

20 403

21 404 **Conclusions**

22 405 The correlation between the scores was moderate. Both tools were able to discriminate
23 406 advancement of CKD stages. Effect size, which denoted the discriminating ability of the
24 407 different CKD stages, is highest when disease condition is advanced and the highest effect
25 408 size was seen in SF-6D. Findings indicate that both tools cover different aspects of health.
26 409 Thus, although there was a moderate correlation between the measures, both scores cannot be
27 410 used interchangeably while assessing QALY during cost utility analysis. Finally, SF 6D had
28 411 the lowest floor and ceiling effect, and was better at detecting different stages of the disease.
29 412 Thus based on the evidence presented in this study, SF 6D appears to be more appropriate to
30 413 be used among CKD patients.
31
32
33
34
35
36
37

38 414

39 415 **Ethics approval and consent to participate**

40 416 The study is in accordance with Helsinki Declaration. The study protocol has been approved
41 417 by the Ethics Committee of Colombo Medical Faculty. Permission was obtained from the
42 418 Provincial Director of Health Service, to assess the CKD register available at his office.
43 419 Participants gave their informed consent.
44
45
46
47

48 420

49 421 **Acknowledgement:**

50 422 Authors acknowledge Asanga Ranasinghe, Priyantha Kumara and Anura Ranasinghe for the
51 423 support rendered during the study.
52
53
54

55 424

56 425 **Funding**

57 426 This study was funded by the Ministry of Health, Sri Lanka
58
59

1 427

2 428 **Competing interests:** None declared.

3 429

4 430 **Patient consent :** Obtained.

5 431

6 432 **Author Contributions**

7 433 SK and SS : Research idea, study design, statistical analysis and drafting of the manuscript.

8 434 NG²: study design, data analysis/interpretation. NG¹: participated in study design, data

9 435 interpretation and supervision. All authors read and approved the final manuscript.

10 436 *Provenance and peer review :* Not commissioned; externally peer reviewed.

11 437

12 438 **Data sharing statement**

13 439 The datasets used and/or analysed during the current study available from the corresponding

14 440 author on reasonable request.

15 441

16 442 **Figure legend**

17 443 Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

18 444 Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

19 445 Fig. 3 : Bland and Altman plot of differences between EQ-5D and SF-6D for patients with

20 446 CKD

21 447 Fig 4: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-

22 448 6D utility scores for CKD stage

23 449

References

1. Senanayake S, Gunawardena N, Paliyawadana P, Kularatna S, Peiris T. Validity and reliability of the Sri Lankan version of the kidney disease quality of life questionnaire (KDQOL-SF™). *Health and quality of life outcomes*. 2017;15(1):119.
2. Schieppati A, Remuzzi G. Chronic renal diseases as a public health problem: epidemiology, social, and economic implications. *Kidney International*. 2005;68:S7-S10.
3. Evaluation IfHMa. Global Burden of Disease 2010 Arrow Diagram: Institute for Health Metrics and Evaluation; 2010 [Available from: <http://vizhub.healthdata.org/irank/arrow.php>].
4. Rebollo-Rubio A, Morales-Asencio JM, Pons-Raventos ME, Mansilla-Francisco JJ. Review of studies on health related quality of life in patients with advanced chronic kidney disease in Spain. *Nefrologia*. 2015;35(1):92-109.
5. Mujais SK, Story K, Brouillette J, Takano T, Soroka S, Franek C, et al. Health-related quality of life in CKD patients: correlates and evolution over time. *Clinical Journal of the American Society of Nephrology*. 2009;4(8):1293-301.
6. Tsai Y-C, Hung C-C, Hwang S-J, Wang S-L, Hsiao S-M, Lin M-Y, et al. Quality of life predicts risks of end-stage renal disease and mortality in patients with chronic kidney disease. *Nephrology dialysis transplantation*. 2009;25(5):1621-6.
7. DeOreo PB. Hemodialysis patient-assessed functional health status predicts continued survival, hospitalization, and dialysis-attendance compliance. *American Journal of Kidney Diseases*. 1997;30(2):204-12.
8. Gussous Y, Theologis AA, Demb JB, Tangtiphaibontana J, Berven S. Correlation Between Lumbopelvic and Sagittal Parameters and Health-Related Quality of Life in Adults With Lumbosacral Spondylolisthesis. *Global spine journal*. 2018;8(1):17-24.
9. Howren MB. Quality-Adjusted Life Years (QALYs). *Encyclopedia of Behavioral Medicine*: Springer; 2013. p. 1605-6.
10. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *International journal of technology assessment in health care*. 1989;5(4):559-75.
11. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *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(1):53-72.
13. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *Journal of health economics*. 2002;21(2):271-92.
14. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical care*. 2002;40(2):113-28.
15. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. *Pharmacoeconomics*. 1995;7(6):490-502.
16. Wu J, Han Y, Zhao F-L, Zhou J, Chen Z, Sun H. Validation and comparison of EuroQoL-5 dimension (EQ-5D) and Short Form-6 dimension (SF-6D) among stable angina patients. *Health and quality of life outcomes*. 2014;12(1):156.
17. Stolk EA, Oppe M, Scalone L, Krabbe PF. Discrete choice modeling for the quantification of health states: the case of the EQ-5D. *Value in Health*. 2010;13(8):1005-13.
18. Kularatna S, Whitty JA, Johnson NW, Jayasinghe R, Scuffham PA. Valuing EQ-5D health states for Sri Lanka. *Quality of Life Research*. 2015;24(7):1785-93.
19. Kularatna S, Byrnes J, Chan YK, Ski CF, Carrington M, Thompson D, 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(12):3399-408.
20. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research*. 2011;20(10):1727-36.
21. Salaffi F, Carotti M, Ciapetti A, Gasparini S, Grassi W. A comparison of utility measurement using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis. *Clinical and experimental rheumatology*. 2011;29(4):661-71.

- 1 503 22. Kwakkenbos L, Fransen J, Vonk MC, Becker ES, Jeurissen M, van den Hoogen FH, et al. A
2 504 comparison of the measurement properties and estimation of minimal important differences of the
3 505 EQ-5D and SF-6D utility measures in patients with systemic sclerosis. *Clinical and experimental*
4 506 *rheumatology*. 2013;31(2 Suppl 76):50-6.
- 5 507 23. Rowen D, Young T, Brazier J, Gaugris S. Comparison of generic, condition-specific, and
6 508 mapped health state utility values for multiple myeloma cancer. *Value in Health*. 2012;15(8):1059-
7 509 68.
- 8 510 24. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D and
9 511 SF-6D utilities in mental health patients. *Health economics*. 2006;15(11):1229-36.
- 10 512 25. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired
11 513 people as estimated by three different utility measures Un comparación de la calidad de vida de
12 514 personas con trastornos auditivos estimada por tres diferentes medidas de utilidad. *International*
13 515 *Journal of Audiology*. 2005;44(3):157-63.
- 14 516 26. Barton GR, Bankart J, Davis AC, Summerfield QA. Comparing utility scores before and after
15 517 hearing-aid provision. *Applied health economics and health policy*. 2004;3(2):103-5.
- 16 518 27. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary
17 519 heart disease. *Health and Quality of Life Outcomes*. 2006;4(1):20.
- 18 520 28. Stavem K, Frøland SS, Hellum KB. Comparison of preference-based utilities of the 15D, EQ-
19 521 5D and SF-6D in patients with HIV/AIDS. *Quality of Life Research*. 2005;14(4):971-80.
- 20 522 29. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D
21 523 based on hypothetical preferences in a general population. *Health economics*. 2005;14(11):1169-89.
- 22 524 30. McDonough CM, Tosteson AN. Measuring preferences for cost-utility analysis.
23 525 *Pharmacoeconomics*. 2007;25(2):93-106.
- 24 526 31. Lwanga SK, Lemeshow S, Organization WH. Sample size determination in health studies: a
25 527 practical manual. 1991.
- 26 528 32. Senanayake S, Gunawardena N, Paliwardana P, Bandara P, Haniffa R, Karunarathna R, et al.
27 529 Symptom burden in chronic kidney disease; a population based cross sectional study. *BMC*
28 530 *nephrology*. 2017;18(1):228.
- 29 531 33. Dolan P. Modeling valuations for EuroQol health states. *Medical care*. 1997;35(11):1095-
30 532 108.
- 31 533 34. Ramos-Goni JM, Rivero-Arias O. eq5d: A command to calculate index values for the EQ-5D
32 534 quality-of-life instrument. *Stata Journal*. 2011;11(1):120.
- 33 535 35. Lamers L, Bouwmans C, van Straten A, Donker M, Hakkaart L. Comparison of EQ-5D and SF-
34 536 6D utilities in mental health patients. *Health economics*. 2006;15(11):1229-36.
- 35 537 36. Harrison M, Davies L, Bansback N, McCoy M, Verstappen S, Watson K, et al. The comparative
36 538 responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. *Quality of*
37 539 *Life Research*. 2009;18(9):1195.
- 38 540 37. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of
39 541 clinical measurement. *The lancet*. 1986;327(8476):307-10.
- 40 542 38. Tredoux C, Durrheim K. Numbers, hypotheses and conclusions: A course in statistics for the
41 543 social sciences. Cape Town: University of Cape Town Press[Links]. 2002.
- 42 544 39. Levey AS, Stevens LA, Coresh J. Conceptual model of CKD: applications and implications.
43 545 *American journal of kidney diseases*. 2009;53(3):S4-S16.
- 44 546 40. Seliger SL, Zhan M, Walker LD, Fink JC. Chronic kidney disease adversely influences patient
45 547 safety. *Journal of the American Society of Nephrology*. 2008;19(12):2414-9.
- 46 548 41. Stevens PE, Levin A. Evaluation and management of chronic kidney disease: synopsis of the
47 549 kidney disease: improving global outcomes 2012 clinical practice guideline. *Annals of internal*
48 550 *medicine*. 2013;158(11):825-30.
- 49 551 42. Tseng C-L, Lafrance J-P, Lu S-E, Soroka O, Miller DR, Maney M, et al. Variability in estimated
50 552 glomerular filtration rate values is a risk factor in chronic kidney disease progression among patients
51 553 with diabetes. *BMC nephrology*. 2015;16(1):34.
- 52 554 43. Kularatna S, Byrnes J, Chan YK, Ski CF, Carrington M, Thompson D, et al. Comparison of the
53 555 EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with cardiovascular
54 556 disease. *Quality of life research : an international journal of quality of life aspects of treatment, care*
55 557 *and rehabilitation*. 2017;26(12):3399-408.

- 1 558 44. Cohen J. A power primer. *Psychological bulletin*. 1992;112(1):155.
- 2 559 45. Litwin MS, Fink A. *How to measure survey reliability and validity*: Sage; 1995.
- 3 560 46. Sack TH, Barton GR, Jenkinson C, Doherty M, Avery AJ, Muir KR. Comparing cost-utility
4 561 estimates: does the choice of EQ-5D or SF-6D matter? *Med Care*. 2009;889-95.
- 5 562 47. Grieve R, Grishchenko M, Cairns J. SF-6D versus EQ-5D: reasons for differences in utility
6 563 scores and impact on reported cost-utility. *The European journal of health economics*.
7 564 2009;10(1):15-23.
- 8 565 48. Stewart ST, Cutler DM, Rosen AB. Comparison of Trends in US Health-Related Quality of Life
9 566 over the 2000's Using the SF-6D, HALex, EQ-5D, and EQ-5D Visual Analog Scale versus a Broader Set
10 567 of Symptoms and Impairments. *Medical care*. 2014;52(12):1010.
- 11 568 49. Zhao FI, Yue M, Yang H, Wang T, Wu Jh, Li SC. Validation and comparison of EuroQol and
12 569 short form 6D in chronic prostatitis patients. *Value in Health*. 2010;13(5):649-56.
- 13 570 50. Cunillera O, Tresserras R, Rajmil L, Vilagut G, Brugulat P, Herdman M, et al. Discriminative
14 571 capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey.
15 572 *Quality of Life Research*. 2010;19(6):853-64.
- 16 573 51. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess
17 574 their ceiling effects in the US general population. *Value in Health*. 2006;9(4):262-71.
- 18 575 52. Janssen M, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement
19 576 properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country
20 577 study. *Quality of Life Research*. 2013;22(7):1717-27.
- 21 578 53. Bharmal M, Thomas III J. Comparing the EQ-5D and the SF-6D descriptive systems to assess
22 579 their ceiling effects in the US general population. *Value Health*. 2006;9(4):262-71.
- 23 580 54. de Boer IH, Olan GA, Patel UD. Screening for chronic kidney disease. *Ann Intern Med*.
24 581 2013;158(5 Pt 1):362-3.
- 25 582 55. Hummel A. Chronic renal failure, a frequent and asymptomatic pathology. *Soins; la revue de*
26 583 *reference infirmiere*. 2010(745):24-6.
- 27 584 56. Jesky MD, Dutton M, Dasgupta I, Yadav P, Ng KP, Fenton A, et al. Health-related quality of
28 585 life impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic kidney
29 586 disease: a prospective observational study. *PloS one*. 2016;11(11):e0165675.
- 30 587
- 31
- 32
- 33 588
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

589 **Tables**

590

591 **Table 1 : Demographic distribution of the sample by the EQ-5D-3L and the SF-6D utility**
592 **scores**

Variable	N (%)	EQ-5D-3L utility mean (SD)	Sf-6D utility mean (SD)	p value [#]
All sample	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 [*]

593 [#] Wilcoxon signed-rank test ; ^{*} Significant at p<0.05

594

595 **Table 2 : Distribution of the sample by the EQ-5D-3L dimensions**

	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)

596

597

598 **Table 3 : Distribution of the sample by the SF-6D dimensions**

	Physical functioning (%)	Role limitation (%)	Social functioning (%)	Pain (%)	Mental health (%)	Vitality (%)
1 ^a	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 ^b	319 (29.1)	NA	NA	240 (21.9)	NA	NA

599 ^a No problems600 ^b Severe problems

601

602

603

604 **Table 4 : Intra Class Correlation 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*

605 * Significant at $p < 0.05$ level

606

607

608 **Table 5 : Discrimination across clinical severity groups**

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

609 [#] Kruskal–Wallis test

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

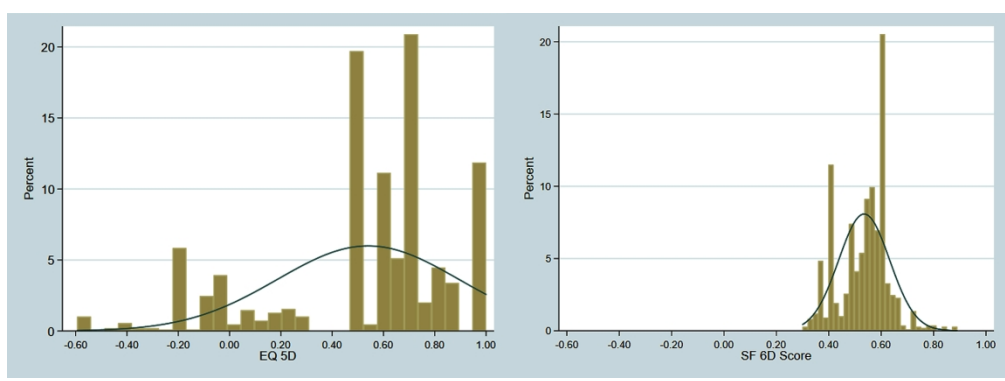


Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

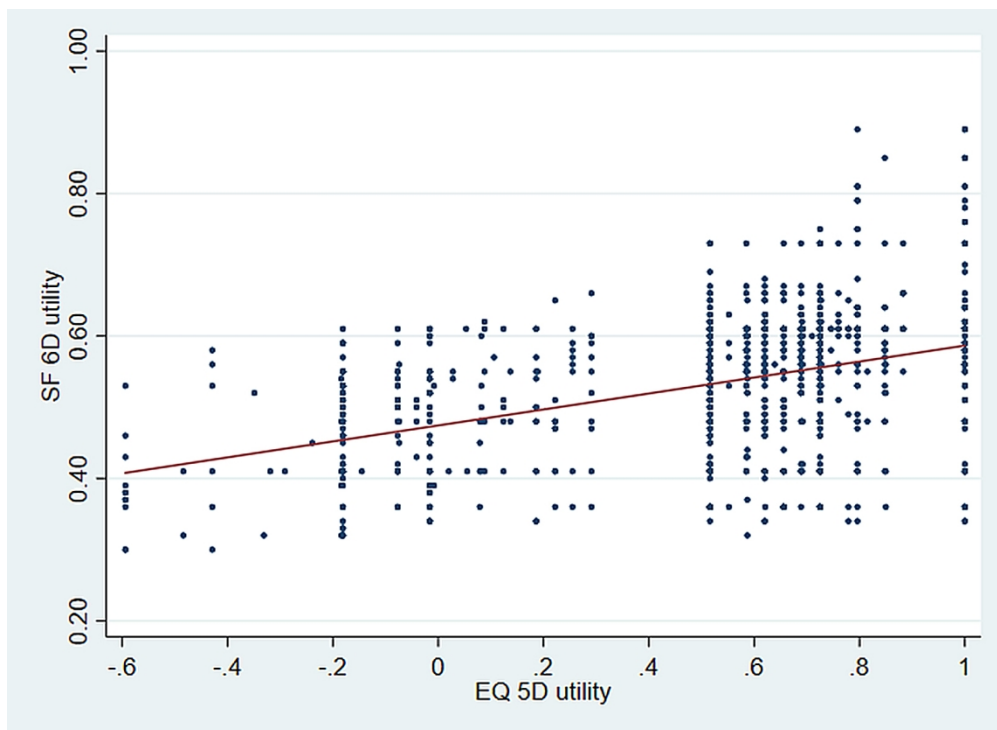


Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

137x99mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

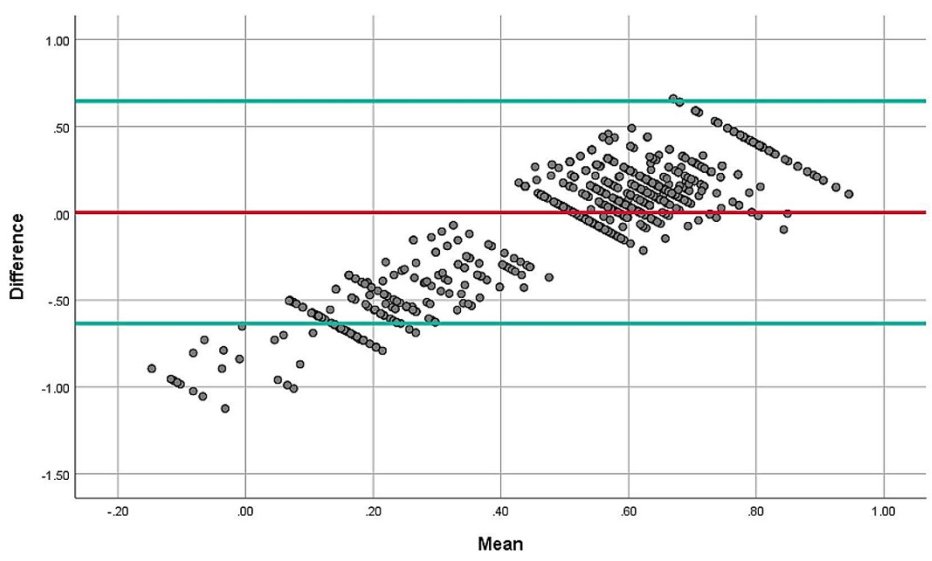


Fig. 3 : Bland and Altman plot of differences between EQ-5D and SF-6D for patients with CKD

166x90mm (300 x 300 DPI)

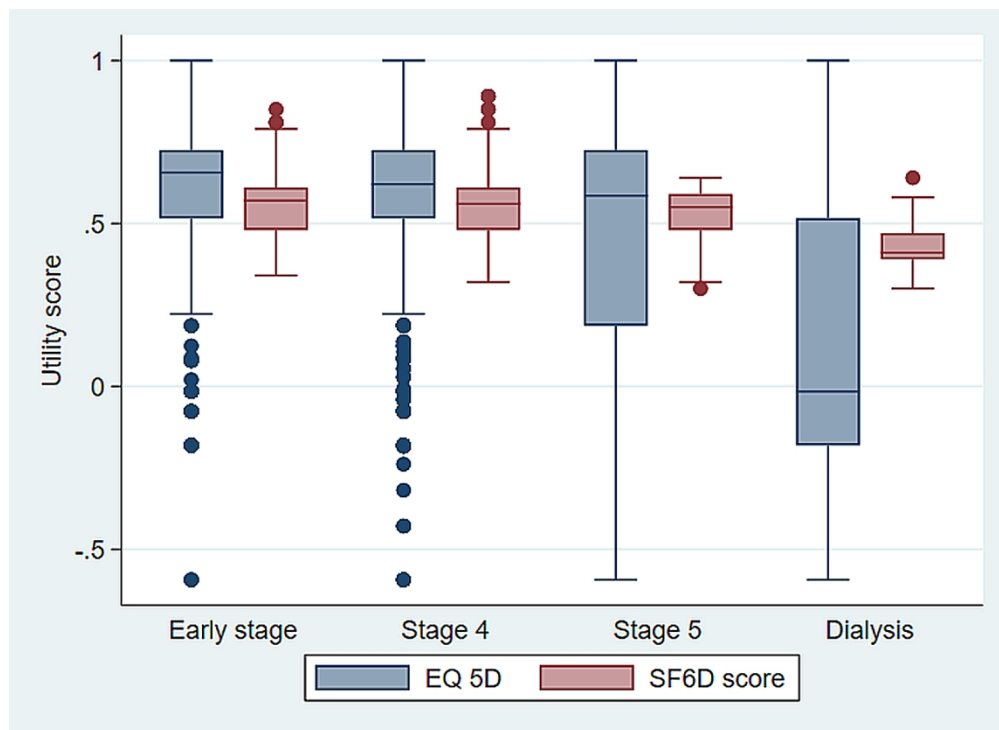


Fig 4: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD stage

137x99mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of *cross-sectional studies*

	Item No	Recommendation	Line number
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	16
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	67
Objectives	3	State specific objectives, including any prespecified hypotheses	115
Methods			
Study design	4	Present key elements of study design early in the paper	121
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	121-126
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	129-134
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	157-190
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	157-190
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	123
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	157-190
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	193-207
		(b) Describe any methods used to examine subgroups and interactions	NA
		(c) Explain how missing data were addressed	NA
		(d) If applicable, describe analytical methods taking account of sampling strategy	NA
		(e) Describe any sensitivity analyses	NA
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	266-267
		(b) Give reasons for non-participation at each stage	267
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	266 - 274
		(b) Indicate number of participants with missing data for each variable of interest	NA

Outcome data	15*	Report numbers of outcome events or summary measures	NA
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	NA
		(b) Report category boundaries when continuous variables were categorized	Table 1
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	335 - 340
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	397
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	405 - 413
Generalisability	21	Discuss the generalisability (external validity) of the study results	410 - 413
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	425

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.

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

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2018-024854.R2
Article Type:	Research
Date Submitted by the Author:	25-Nov-2018
Complete List of Authors:	Kularatna, Sanjeewa; Australian Centre For Health Services Innovation, Queensland University of technology Senanayake, Sameera; Australian Centre For Health Services Innovation, Queensland University of Technology Gunawardena, Nalika; World Health Organization Graves, Nicholas; Australian Centre For Health Services Innovation, Queensland University of Technology
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health economics, Health services research, Renal medicine
Keywords:	Cost utility analysis, Quality-Adjusted Life Years (QALYs), utility, Chronic Kidney Disease (CKD)

SCHOLARONE™
Manuscripts

1 1 **Comparison of the EQ-5D-3L and the SF-6D (SF-36) contemporaneous utility scores in**
2
3 2 **patients with Chronic Kidney Disease in Sri Lanka; a cross sectional survey**
4

5 3
6 4
7
8 5 Sanjeewa Kularatna¹, Sameera Senanayake¹, Nalika Gunawardena², Nicholas Graves¹
9

10 6
11 7 ¹Australian Centre for Health Services Innovation, Queensland University of Technology
12

13 8 ² World Health Organization, Sri Lanka
14

15 9
16
17 10 **Corresponding author;**

18 11 Sanjeewa Kularatna - *sanjeewa.kularatna@qut.edu.au*
19

20 12
21 13 **Word count : 3277**
22

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

16 Abstract

17
18 **Objectives** : The aim of this study was to compare utility weights of EQ-5D-3L and SF-6D in
19 a representative cohort of patients with Chronic Kidney Disease (CKD). A cost-utility
20 analysis is designed to report the change to costs required to achieve an estimated change to
21 Quality-Adjusted Life Years (QALYs). The quality component of a QALY is measured by
22 utility. Utility represents the preference of general population for a given health state.
23 Classification systems of the multi-attribute utility instruments (MAUI) are used to define
24 these health states. Utility weights developed from different classification systems can vary
25 and may affect the conclusions from cost-utility analyses.

26
27 **Design**: A community based cross sectional study

28
29 **Setting** : Anuradhapura a rural district in Sri Lanka.

30
31 **Participants**: A representative sample of 1096 CKD patients, selected using the population-
32 based CKD register, completed the EQ-5D-3L and SF-36. SF-6D was constructed from the
33 SF-36 according to the published algorithm. The study assessed discrimination, correlation
34 and differences across the two instruments.

35
36 **Results**: Study participants were predominantly male (62.6%). Mean EQ-5D-3L utility score
37 was 0.540 (SD 0.35) compared with 0.534 (SD 0.09) for the SF-6D ($p=0.588$). The
38 correlation (r) between the scores was 0.40 ($p<0.001$). Utility scores were significantly
39 different in both males and females between the two tools, but there was no difference in age
40 and educational categories. Both MAUI scores were significantly lower ($p<0.001$) among
41 those who were in more advanced stages of the disease and the corresponding utility scores
42 of the two instruments in different CKD stages were also significantly different ($p<0.05$).
43 The largest effect size was seen among the dialysis patients.

44
45 **Conclusions**: The correlation between the scores was moderate. SF 6D had the lowest floor
46 and ceiling effect, and was better at detecting different stages of the disease. Thus based on the
47 evidence presented in this study, SF 6D appears to be more appropriate to be used among CKD
48 patients

49
50 **Key words**: Cost utility analysis, Quality-Adjusted Life Years (QALYs), utility, Chronic
51 Kidney Disease (CKD)

1 52 **Strengths and limitations of this study**

- 2
- 3 53 • The response rate of the study is very high.
- 4
- 5 54 • Both tools used in the study (EQ 5D 3L and SF 36) have been previously validated to
- 6 the Sri Lankan setting before.
- 7 55
- 8 56 • Data collectors were experienced for many local and international studies done among
- 9 CKD patients in Sri Lanka and further they were trained by the principal investigator to
- 10 57 ensure the quality of the data collected.
- 11 58
- 12 59 • Our study was a cross-sectional study, thus we could not assess how utility scores of
- 13 the two instruments change over time.
- 14 60
- 15 61 • Some of the information related to QOL in SF-36 is considered to be sensitive in
- 16 nature and the fact that this information was obtained utilising an interviewer-
- 17 62 administered questionnaire could have led to some under-reporting in the assessment
- 18 of QOL though many measures were taken to minimize this issue.
- 19 63
- 20 64
- 21 65
- 22 66
- 23 67
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

68 **Introduction**

69 Chronic kidney disease (CKD) is a substantial public health problem with adverse
70 psychological, physical and economic outcomes. The burden of CKD is increasing
71 globally(1). World Health Report (2002) and Global Burden of Disease (GBD) project stated
72 that the diseases of the kidney contribute much to the global disease burden with
73 approximately 850,000 deaths every year globally (2). Furthermore, according to GBD study
74 conducted in 2010, of the top causes of DALY, CKD is ranked 29th globally, 23rd in South
75 East Asia and 14th in Sri Lanka (3). Due to the progressive and disabling nature of CKD, it
76 poses a substantial impact on the quality of life (QOL) of individuals. It is important to
77 measure QOL indicators for the management of chronic kidney disease patients. Several
78 studies have demonstrated a relationship between reduced QOL and increased morbidity and
79 mortality (4-7).

81 World over, the importance of including QOL indicators in the clinical management of
82 patients has been highlighted. This has come to the limelight after several studies
83 demonstrated the strong relationship between reduced QOL and increased morbidity and
84 mortality (5, 8). Meantime, economic evaluation has become increasingly popular among
85 researchers and policy makers during resource allocation in recent years. Due to the
86 relationship between QOL and clinical outcome, during the recent years, QOL has become an
87 important health outcome in economic evaluations. In cost utility analysis (CUA), a method
88 of economic evaluation, outcomes are usually measured in Quality-Adjusted Life Years
89 (QALYs), which is a measure of QOL.

91 The concept of QALYs was developed in the 1970s. It can measure the changes of an
92 individual's quality and quantity of life and can also aggregate these improvements across
93 individual (9, 10). The change in the quality of life in QALY is measured using a set of
94 weights, called utilities, which reflect different health states. For all possible health states,
95 utilities should be measured on a scale where 1 refers to best imaginable health and 0 refers
96 to death (11). Measuring utilities for different health states is complex and time-consuming.
97 Thus, Multi Attribute Utility Instruments (MAUI) such as EuroQol-5D (EQ-5D-3L) (12),
98 Short Form-6D (SF-6D) (13) or the Health Utility Index (HUI) (14, 15) are used to define
99 different health states. The utility scores for different health states in different instruments are
100 derived from methods such as Standard Gambling method (16), Discrete Choice Experiments
101 (17) and Time Trade-Off experiments (18). EQ-5D-3L is the most widely used utility
102 instrument at present (19). EQ-5D-5L, which is the newer version of EQ-5D, has also been
103 developed and tested recently (20).

1 104
2
3 105 Since all the MAUIs aim at measuring the health state of individuals, all the instruments
4
5 106 should generate the same utility value for a particular state of health. However, the evidence
6
7 107 indicates that there is an essential difference in the utility scores for a particular health state
8
9 108 between different instruments (19, 21-29). This, in turn, indicates that the choice of the
10
11 109 MAUI used may adversely influence the results of CUA and thereby the decision-making
12
13 110 process (30). Furthermore, for incremental analyses, use of different MAUIs may lead to
14
15 111 different results regarding the magnitude, direction or significance of any change in health-
16
17 112 related quality of life measure.

18
19 114 Though the differences between MAUIs have been evaluated in many disease conditions (19,
20
21 115 21, 22, 24), there are no evidence in the literature comparing MAUIs using patients with
22
23 116 CKD. The aim of our study is to compare contemporaneous EQ-5D-3L and SF-6D utility
24
25 117 scores in patients with CKD. Results may be useful for researchers selecting a generic MAUI
26
27 118 to estimate utilities for use in economic modelling of treatments for CKD.

28 119

29 120 **Methods**

30 121 *Participant selection*

31 122 A population-based descriptive cross-sectional study was conducted in the district of
32
33 123 Anuradhapura in the North Central Province (NCP) of Sri Lanka between Septembers to
34
35 124 December 2015. The study population consisted of 1162 confirmed CKD patients, calculated
36
37 125 using the appropriate formula (31), who were over 18 years old with documented evidence of
38
39 126 CKD living in the Anuradhapura district. The diagnosis of CKD was made if the Glomerular
40
41 127 Filtration Rate (GFR) was less than 60 ml/min per 1.73m² of body surface area in two
42
43 128 measurements made three months apart.

44 129

45
46 130 The inclusion criteria were patients above 18 years of age and those who were diagnosed as
47
48 131 having CKD by a specialist nephrologist or a consultant physician. Presence of evidence of
49
50 132 such diagnosis was made by way of diagnosis cards, clinic records or any other record issued
51
52 133 by a specialist nephrologist, a consultant physician or a government hospital. Patients who
53
54 134 had previous renal transplantation, who were unable to provide rational information due to
55
56 135 any cause (e.g. mental retardation) and who were critically ill were excluded from the study.

57 136

58 137 The study instrument was an interviewer-administered questionnaire to gather information on the
59
60 138 socio-demographic information, CKD related information, EQ-5D-3L and SF 36.

139

1 140 Five Public Health Inspectors working in the CKD unit in the North Central Province were used
2 141 for the data collection and all have been working in the unit for more than 5 years and they had
3 142 experience in functioning as data collectors for many local and international studies done among
4 143 CKD patients in the NCP. The data collectors assessed the eligibility of patients by reviewing
5 144 their clinical records. Informed consent was obtained from those who were eligible for
6 145 participation in the study before doing the face-to-face interview.
7
8
9
10
11
12

13 147 The study was conducted in all nineteen Medical Officer of Health (MOH) areas of the
14 148 Anuradhapura district. The number of participants to be included from each MOH area was
15 149 based on probability proportionate to the size of CKD patients registered in each of the MOH
16 150 areas. The required number of participants from each MOH area was selected using simple
17 151 random sampling method. The population-based CKD register – which records the patients
18 152 with a confirmed diagnosis of CKD from renal clinics in hospitals of the NCP since 2003 –
19 153 was used as the sampling frame. The register was obtained from the office of the Provincial
20 154 Director of Health Services (32).
21
22
23
24
25
26
27
28

29 156 ***Calculation of utility scores***

30 157 Currently, there is no algorithm based on preferences of the Sri Lankan public to score the
31 158 SF-6D on a utility scale. Therefore, the UK algorithm was used for this purpose (13).

32 159 Though Sri Lankan EQ-5D-3L utility scores are available (18), the UK utility scores were
33 160 used for the EQ-5D-3L (33) because of the unavailability of comparable Sri Lankan SF-6D
34 161 utility scores as mentioned earlier. This allowed the comparison of utility scores from the
35 162 same country.
36
37
38
39
40
41

42 164 The EQ-5D-3L instrument contains five domains; mobility, self-care, usual activities, pain /
43 165 discomfort and anxiety / depression. Each domain has one item and each item has three
44 166 levels: one denoting no problems and three denoting severe problems (12). Thus, EQ-5D-3L
45 167 has mutually exclusive 243 different health states.
46
47
48
49

50 168
51 169 SF-6D is derived from either SF-36 or SF-12 (Version 1 and Version 2). The current study
52 170 utilised SF-36 for the data collection. SF-36 includes 36 items that measure eight domains;
53 171 role limitations caused by physical problems (4 items), physical function (10 items), role
54 172 limitations caused by emotional problems (3 items), pain (2 items), social function (2 items),
55 173 general health perceptions (5 items), emotional well-being (5 items) and energy / fatigue (4
56 174 items). Questions have different answer options which range from two to six. While scoring,
57 175 each question is scored in a scale ranging from 0 (worst health) to 100 (best health). All items

1 176 in a domain are summed up and averaged to give an average score for each domain which
2
3 177 ranges from 0 (worst health) to 100 (best health). To calculate the utility scores of the SF-6D,
4
5 178 11 items are used covering six domains; physical functioning, role limitation, social
6
7 179 functioning, pain, mental health and vitality (13).

8 180
9
10 181 The EQ-5D-3L utility calculation was undertaken using the STATA syntax developed by
11
12 182 Ramos-Goni et al. (34). The SF-6D scores were computed based on published algorithms
13
14 183 (13). Patients for whom one of the two measurements was missing were excluded from the
15
16 184 analysis.

17 185
18
19 186 The EQ-5D-3L utility scores range from -0.59, 0=being dead; negative values represent
20
21 187 health status considered worse than “dead”, to 1.00 which indicate best imaginable health.
22
23 188 The SF-6D utility scores ranged from 0.296 which indicate severely impaired levels in all
24
25 189 dimensions to 1.0 which indicates no difficulty in any dimensions.

26 190
27 191 **Data analysis**
28
29 192 STATA 15.1 software was used for the analysis. Mean utility scores on each instrument were
30
31 193 compared by socio-demographic characteristics. Normality of the two distributions were
32
33 194 assessed using Kolmogorov-Smirnov and Shapiro-Wilk tests. Wilcoxon signed-rank test
34
35 195 was used to assess the difference between the two instruments in each socio-demographic
36
37 196 class (35). Histograms were plotted for the two utility values distribution. Floor effects and
38
39 197 ceiling (proportion of patients with the highest and lowest possible scores respectively) were
40
41 198 calculated for the EQ-5D-3L and SF-6D. Ceiling and floor effects were considered small if
42
43 199 $\leq 15\%$ of patients occupy the best or worst health states, but they were considered serious if
44
45 200 $> 15\%$ of patients occupy these states (36).

46 201
47 202 Currently, an established methodology to compare different MAUIs is not available. Thus,
48
49 203 recently published methodologies, which compared different MAUIs, were followed in the
50
51 204 current study (19, 23, 35). This included a combination of statistical and psychometric
52
53 205 analyses to examine discrimination, agreement, differences and correlation between the two
54
55 206 instruments.

56 207
57 208 **Agreement and differences**
58
59 209 The Wilcoxon Signed-Ranks Test was used to assess the overall difference between the EQ-5D-
60
210 3L and SF-6D utility scores and the difference of the utility scores according to different
211
socio-demographic and disease related features. Furthermore, the distribution of the

1 212 responses to the different domains of the two instruments was tabulated to present the
2 213 agreement and the differences between the two instruments. Bland-Altman plot was also
3 214 used to assess the proportional error and the limit of agreement (37).
4
5
6
7

215

216 ***Intra Class Correlation (ICC)***

217 The dimensions of the two instruments were compared using ICC. The related dimensions
218 between the two MAUIs are role limitation (SF-6D)/usual activities (EQ-5D-3L), physical
219 functioning (SF-6D)/mobility and self-care (EQ-5D-3L), pain (SF-6D)/pain and discomfort
220 (EQ-5D-3L), social functioning (SF-6D)/usual activities (EQ-5D-3L) and mental health (SF-
221 6D)/anxiety and depression (EQ-5D-3L). The vitality dimension of the SF-6D did not have
222 any related dimension with the EQ-5D-3L. The magnitude of the correlation coefficients
223 were interpreted according to Guilford's criteria (38).
224

225

225 ***Discrimination***

226 It is important that MAUIs can discriminate correctly among groups of different severity as
227 MAUIs are meant to measure change in QOL due to improvement or worsening of the health,
228 in the condition of interest.
229

230

230 Glomerular Filtration Rate (GFR) is the most important indicator of kidney function of
231 patients with CKD (39). Studies have shown that decreased GFR is associated with infection,
232 impaired cognitive and physical function as well as threats to patient safety (40). Though
233 classifications exist to classify stages of CKD, it is evident that at present most of the clinical
234 decision making in CKD is solely based on GFR base classification (41, 42). Depending on
235 the GFR value, CKD is categorised into five stages; stage I to stage V. For analytical
236 purposes, the CKD stages I to III were categorised as "early stage" in the present study. It is
237 expected that with advanced stages of the disease, the utility scores should be lower than the
238 early stages.
239

240

240 Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using the
241 non-parametric test, Kruskal-Wallis, and effect size. The instrument's ability to discriminate
242 between two adjacent stages was estimated by calculating the effect size. The effect size was
243 calculated by dividing the mean difference of two adjacent CKD stages by the standard
244 deviation of the milder of the two CKD stages (23, 43). Large effect size indicates better
245 discriminating ability of the instrument. The effect size was categorised into small (0.2-0.5),
246 medium (0.5-0.8) and large (more than 0.8) (44).
247

248 ***Test-retest reliability***

249 To assess the test-retest reliability of the study instrument, within a period of one week, 30
250 randomly selected study participants were visited at their households by the data collectors.
251 Test re-test reliability of the utility scores of the two instruments was assessed using ICC and
252 a value of 0.70 or greater was considered as satisfactory reliability (45).

254 ***Patient and Public Involvement***

255 The main stakeholders in the provision of care for the CKD patients such as consultants,
256 medical officers working in nephrology units, community leaders and the patients living in
257 this area were involved in planning the study. Their concerns were always entertained and
258 where feasible their concerns were incorporated into the study. During the data collection,
259 stage permission was obtained from the respective local officers. The results of the study was
260 communicated to the local level officials such as Medical Officer of Health, Divisional
261 Secretariat, Regional Director of Health Services and Provincial Director of Health Services.

263 **Results**

264 ***Sample characteristics***

265 Out of 1162 participants selected to be included in the study, 66 (5.6%) did not participate in
266 the study giving a response rate of 94.4%. The mean age of the study population was 58.4
267 years (Standard Deviation (SD) 10.8). There was a preponderance of males among the study
268 population (62.6%, N=686). The mean eGFR of the population was 31.8 (SD 20.2)
269 ml/min/1.73 m². The mean number of years since diagnosed with CKD was found to be 4.1
270 (SD 3.2) years. The majority of participants was in the later stages, stage 4 or beyond, of
271 CKD (n=803; 73.2%). 38 participants (3.6%), with stage 5 of the disease and undergoing
272 dialysis, were on haemodialysis (Table 1). Chronic Kidney Disease of Unknown origin
273 (CKDu) was the cause of the CKD in most of the study population (n=489; 43.7%).

275 ***Distribution of EQ-5D-3L and SF-6D utility scores***

276 The mean EQ-5D-3L utility score at baseline was 0.540 compared with 0.534 for the SF-6D
277 as summarised in Table 1. The EQ-5D-3L utility score ranged from -0.594 to 1, while SF-6D
278 ranged from 0.3 to 0.89. The median baseline values have different locations in their
279 respective scoring ranges (Fig 1). The EQ-5D-3L showed 1.0% floor effect and 11.8% ceiling
280 effect, while SF-6D had 0.0% floor and ceiling effects.

282 ***Agreement, differences and correlation between the two utility scores***

283 Analyses revealed non- normal distribution of the utility scores of both the instruments, thus
284 Wilcoxon Signed-Ranks test was used to compare the two utility scores. There was significant

1 285 difference ($p < 0.001$) between overall scores of the two utility instruments. Further the two
2 286 utility scores were significantly different among males (< 0.001), age more than 40 years
3 287 groups, those who were employed, among both who had and didn't have comorbidities, up to
4 288 stage IV of CKD and among dialysis patients (Table 1). The standard deviation of the EQ-
5 289 5D-3L was considerably larger than that of the SF-6D among all sub groups.

6 290
7
8 291 Significant proportion of the patients reported "no problem" in any of the EQ-5D-3L
9 292 dimensions than the SF-6D. However, fewer patients reported "extreme problems" in the EQ-
10 293 5D-3L than in the SF-6D (Tables 2 and 3). Patients reported different results for the related
11 294 dimensions of the two MAUIs (Tables 2 and 3). Nearly half of the patients reported "no
12 295 problem" in Mobility domain of the EQ-5D-3L, while only 0.7% reported "no problems"
13 296 with the physical functioning of the SF-6D. Nearly quarter (23.8%) of patients reported "no
14 297 problems" for the anxiety / depression dimension in the EQ-5D-3L, whereas only 0.6%
15 298 reported the same for the mental health dimension of the SF-6D.

16 299
17 300 The correlation between EQ-5D-3L and SF-6D was 0.408, which was statistically significant
18 301 at $p < 0.001$ level (Figure 2). Regarding the ICC between different domains of the two
19 302 instruments, according to the Guilford's criteria, moderate correlation (0.4-0.6) was evident
20 303 between Social functioning and Mobility (0.517), Social functioning and Self-care (0.424),
21 304 Social functioning and Usual activities (0.464), Social functioning and Pain/discomfort
22 305 (0.566), Social functioning and Anxiety/depression (0.528), Pain and Mobility (0.475), Pain
23 306 and Pain/ discomfort (0.482), Pain and Anxiety/depression (0.484), Vitality and Pain/
24 307 discomfort (0.475) and Vitality and Anxiety/depression (0.453) (Table 4). The Bland-Altman
25 308 plot showed proportional error and wide limits of agreement (Figure 3).

26 309
27 310 **Discrimination**
28 311 With both MAUIs, utility scores decreased with increasing severity (as measured by CKD
29 312 stage) (Table 5). In both MAUIs, the utility differences across CKD stages were statistically
30 313 significant ($p < 0.05$) indicating good discrimination. Figure 4 indicated the box-plots present
31 314 the median, quartiles and extreme values for the EQ-5D-3L and SF-6D utility scores for CKD
32 315 stage. Furthermore, the calculated effect size between CKD early stage and stage IV was
33 316 0.071 and 0.141 for EQ-5D-3L and SF-6D respectively. The highest effect size was observed
34 317 between CKD stage V and dialysis group, which was 0.807 for EQ-5D-3L and 1.098 for SF-
35 318 6D.

36 319
37 320 **Test-retest reliability**

1 321 The test re-test ICC was 0.943 in EQ-5D-3L while it was 0.921 in SF 6D, indicating good test
2
3 322 re-test reliability in both the instruments.
4
5 323

6 324 **Discussion**

8 325 This is the first study to compare the utility scores arising from the EQ-5D-3L and SF-6D in
9
10 326 CKD patients. According to the current study, the correlation between the scores was
11
12 327 moderate. Both tools were able to discriminate advancement of CKD stages. Effect size,
13
14 328 which denoted the discriminating ability of different CKD stages, is highest when disease
15
16 329 condition is advanced and the highest effect size was seen in SF-6D. Further, the lowest
17
18 330 ceiling effect and the floor effect were seen in SF 6D.
19

20 331
21 332 Evidence indicate that the choice of MAUI (e.g.; EQ 5D or SF6D) has an impact on the
22
23 333 results of the cost-utility analysis (46, 47). Sack et al. (2009) compared the results of cost-
24
25 334 utility estimates using both EQ 5D and SF 6D. Results indicated contrasting results for the
26
27 335 two instruments and authors concluded that the choice of the instrument does matter in cost-
28
29 336 utility analysis (46). Thus, from an economic perspective it is important to know the most
30
31 337 suitable MAUI to be used among CKD patients.
32

33 338
34 339 At present, there is no consensus on the methodology to compare the utility scores of
35
36 340 different MAUIs (19, 35). The present study adopted the methodologies used by Kularatna et
37
38 341 al. (2017) and Lamers et al. (2006) (19, 35). Only one time assessment of the utilities was
39
40 342 done in the present study. Thus, the responsiveness of the two instruments to changes in
41
42 343 kidney function over time was not assessed. Though Sri Lankan EQ-5D-3L utility scores are
43
44 344 available (18), yet we used the UK utility scores for the EQ-5D-3L (33) because of the
45
46 345 unavailability of comparable Sri Lankan SF-6D utility scores values. This is an accepted
47
48 346 method of calculating the utility scores in the absence of country specific utilities. Two
49
50 347 studies conducted in Netherlands (24) and Italy (21), comparing the utility scores of the two
51
52 348 instruments, had used the UK derived EQ-5D-3L and SF-6D utility scores.
53

54 349
55 350 The present study did not find any difference ($p=0.588$) between the overall mean scores of
56
57 351 the two utility instruments. This was similar to a study conducted among a group of
58
59 352 HIV/AIDS patients (28), but different from other studies available in the literature where
60
353 different results have been reported. Significantly higher utility values for EQ-5D-3L were
354
355 found among general population (29, 48), cardio-vascular disease patients (19), rheumatoid
356
arthritis patients (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

1 357 obtained for SF-6D instrument (24). These varying results could be due to different recall
2 358 periods of the two instruments. EQ-5D-3L assessed the health status of the day of instrument
3 359 administration while SF-6D, which was derived from SF-36, assessed the health status of the
4 360 past 30 days.

5 361
6 362 Though overall ceiling and floor effects of both instruments were small, relatively higher
7 363 ceiling effect was evident in the EQ-5D-3L. This was consistent with several other studies
8 364 conducted elsewhere, where EQ 5D 3L reported a relatively higher ceiling effect compared to
9 365 SF 6D (16, 19, 49-51). This is mainly due to the fact that the EQ-5D-3L has limited response
10 366 levels and the five level newer version of EQ-5D-3L expected to improve the properties of
11 367 the three-level in terms of reduced ceiling effects, increased reliability and improved ability
12 368 to discriminate between different levels of health (52). Further, the current study reported
13 369 relatively lower ceiling effect, for the EQ 5D, compared to results obtained among
14 370 Parkinson's disease (13.5%) and stable angina (15.5%) patients. However, our result was
15 371 higher compared to the ceiling effect observed among patients with systemic sclerosis
16 372 (7.0%). Among many other factors that could contribute to these differences, the level of
17 373 morbidity of a disease is said to be one of the factors which could influence the ceiling effect
18 374 observed in EQ 5D (53). Thus the diseases with lower morbidity are expected to have higher
19 375 ceiling effects.

20 376
21 377 Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using
22 378 ANOVA and effect size. In both MAUIs, the utility differences across CKD stages were
23 379 statistically significant ($p < 0.05$; ANOVA) indicating good discrimination. However, the
24 380 effect size was small for both the tools until the dialysis stage. At the dialysis stage, the effect
25 381 size is large and this was highest in the SF-6D instrument. It could be because CKD is
26 382 considered asymptomatic until the later stages of the disease (54, 55), not allowing the
27 383 instruments to discriminate the different stages. According to a recent study conducted by
28 384 Jesky et al. (2016), the EQ-5D-3L utility scores of the adjacent pre-dialysis CKD stages were
29 385 not found to be statistically significant (56).

30 386

31 387 **Limitations**

32 388 Some of the information related to QOL in SF-36 is considered to be sensitive in nature and
33 389 the fact that this information was obtained utilising an interviewer-administered questionnaire
34 390 could have led to some under-reporting in the assessment of QOL though many measures
35 391 were taken to minimize this issue. Our study was a cross-sectional study, thus we could not
36 392 assess how utility scores of the two instruments change over time.

1 393

2
3 394 **Conclusions**

4
5 395 The correlation between the scores was moderate. Both tools were able to discriminate
6 396 advancement of CKD stages. Effect size, which denoted the discriminating ability of the
7
8 397 different CKD stages, is highest when disease condition is advanced. Findings indicate that
9
10 398 both tools cover different aspects of health. Thus, although there was a moderate correlation
11
12 399 between the measures, both scores cannot be used interchangeably while assessing QALY
13
14 400 during cost utility analysis. Finally, SF 6D had the lowest floor and ceiling effect, and was
15
16 401 better at detecting different stages of the disease. Thus based on the evidence presented in
17
18 402 this study, SF 6D appears to be more appropriate to be used among CKD patients.

19 403

20 404 **Ethics approval and consent to participate**

21
22 405 The study is in accordance with Helsinki Declaration. The study protocol has been approved
23
24 406 by the Ethics Committee of Colombo Medical Faculty. Permission was obtained from the
25
26 407 Provincial Director of Health Service, to assess the CKD register available at his office.
27
28 408 Participants gave their informed consent.

29 409

30 410 **Acknowledgement:**

31
32 411 Authors acknowledge Asanga Ranasinghe, Priyantha Kumara and Anura Ranasinghe for the
33
34 412 support rendered during the study.

35 413

36 414 **Funding**

37
38 415 This study was funded by the Ministry of Health, Sri Lanka

39 416

40
41 417 **Competing interests:** None declared.

42 418

43
44 419 **Patient consent :** Obtained.

45 420

46 421 **Author Contributions**

47
48 422 SK and SS : Research idea, study design, statistical analysis and drafting of the manuscript.
49
50 423 NG²: study design, data analysis/interpretation. NG¹: participated in study design, data
51
52 424 interpretation and supervision. All authors read and approved the final manuscript.

53
54 425 *Provenance and peer review :* Not commissioned; externally peer reviewed.

55 426

56 427 **Data sharing statement**

1 428 The datasets used and/or analysed during the current study available from the corresponding
2
3 429 author on reasonable request.
4

5 430

6 431 **Figure legend**

7
8 432 Fig. 1. Distribution of EQ-5D-3L (A) and SF-6D (B)

9
10 433 Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

11
12 434 Fig. 3 : Bland and Altman plot of differences between EQ-5D-3L and SF-6D for patients with
13 435 CKD

14
15 436 Fig 4: The box-plots present the median, quartiles and extreme values for the EQ-5D-3L and
16 437 SF-6D utility scores for CKD stage

17
18 438
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

439 **References**

- 440 1. Senanayake S, Gunawardena N, Paliyawadana P, Kularatna S, Peiris T. Validity and reliability
441 of the Sri Lankan version of the kidney disease quality of life questionnaire (KDQOL-SF™). Health and
442 quality of life outcomes. 2017;15(1):119.
- 443 2. Schieppati A, Remuzzi G. Chronic renal diseases as a public health problem: epidemiology,
444 social, and economic implications. *Kidney International*. 2005;68:S7-S10.
- 445 3. Evaluation of the Global Burden of Disease 2010 Arrow Diagram: Institute for Health Metrics
446 and Evaluation; 2010 [Available from: <http://vizhub.healthdata.org/irank/arrow.php>].
- 447 4. Rebollo-Rubio A, Morales-Asencio JM, Pons-Raventos ME, Mansilla-Francisco JJ. Review of
448 studies on health related quality of life in patients with advanced chronic kidney disease in Spain.
449 *Nefrologia*. 2015;35(1):92-109.
- 450 5. Mujais SK, Story K, Brouillette J, Takano T, Soroka S, Franek C, et al. Health-related quality of
451 life in CKD patients: correlates and evolution over time. *Clinical Journal of the American Society of*
452 *Nephrology*. 2009;4(8):1293-301.
- 453 6. Tsai Y-C, Hung C-C, Hwang S-J, Wang S-L, Hsiao S-M, Lin M-Y, et al. Quality of life predicts risks
454 of end-stage renal disease and mortality in patients with chronic kidney disease. *Nephrology dialysis*
455 *transplantation*. 2009;25(5):1621-6.
- 456 7. DeOreo PB. Hemodialysis patient-assessed functional health status predicts continued
457 survival, hospitalization, and dialysis-attendance compliance. *American Journal of Kidney Diseases*.
458 1997;30(2):204-12.
- 459 8. Gussous Y, Theologis AA, Demb JB, Tangtiphaiboontana J, Berven S. Correlation Between
460 Lumbopelvic and Sagittal Parameters and Health-Related Quality of Life in Adults With Lumbosacral
461 Spondylolisthesis. *Global spine journal*. 2018;8(1):17-24.
- 462 9. Howren MB. Quality-Adjusted Life Years (QALYs). *Encyclopedia of Behavioral Medicine*:
463 Springer; 2013. p. 1605-6.
- 464 10. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *International journal of*
465 *technology assessment in health care*. 1989;5(4):559-75.
- 466 11. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic*
467 *evaluation of health care programmes*: Oxford university press; 2015.
- 468 12. Brooks R. EuroQol: the current state of play. *Health policy*. 1996;37(1):53-72.
- 469 13. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from
470 the SF-36. *Journal of health economics*. 2002;21(2):271-92.
- 471 14. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and
472 single-attribute utility functions for the health utilities index mark 3 system. *Medical care*.
473 2002;40(2):113-28.
- 474 15. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification
475 systems. *Pharmacoeconomics*. 1995;7(6):490-502.
- 476 16. Wu J, Han Y, Zhao F-L, Zhou J, Chen Z, Sun H. Validation and comparison of EuroQoL-5
477 dimension (EQ-5D) and Short Form-6 dimension (SF-6D) among stable angina patients. *Health and*
478 *quality of life outcomes*. 2014;12(1):156.
- 479 17. Stolk EA, Oppe M, Scalone L, Krabbe PF. Discrete choice modeling for the quantification of
480 health states: the case of the EQ-5D. *Value in Health*. 2010;13(8):1005-13.
- 481 18. Kularatna S, Whitty JA, Johnson NW, Jayasinghe R, Scuffham PA. Valuing EQ-5D health states
482 for Sri Lanka. *Quality of Life Research*. 2015;24(7):1785-93.
- 483 19. Kularatna S, Byrnes J, Chan YK, Ski CF, Carrington M, Thompson D, et al. Comparison of the
484 EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with cardiovascular
485 disease. *Quality of Life Research*. 2017;26(12):3399-408.
- 486 20. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary
487 testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research*. 2011;20(10):1727-
488 36.
- 489 21. Salaffi F, Carotti M, Ciapetti A, Gasparini S, Grassi W. A comparison of utility measurement
490 using EQ-5D and SF-6D preference-based generic instruments in patients with rheumatoid arthritis.
491 *Clinical and experimental rheumatology*. 2011;29(4):661-71.

- 1 492 22. Kwakkenbos L, Fransen J, Vonk MC, Becker ES, Jeurissen M, van den Hoogen FH, et al. A
2 493 comparison of the measurement properties and estimation of minimal important differences of the
3 494 EQ-5D and SF-6D utility measures in patients with systemic sclerosis. *Clinical and experimental*
4 495 *rheumatology*. 2013;31(2 Suppl 76):50-6.
- 5 496 23. Rowen D, Young T, Brazier J, Gaugris S. Comparison of generic, condition-specific, and mapped
6 497 health state utility values for multiple myeloma cancer. *Value in Health*. 2012;15(8):1059-68.
- 7 498 24. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D and
8 499 SF-6D utilities in mental health patients. *Health economics*. 2006;15(11):1229-36.
- 9 500 25. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired people
10 501 as estimated by three different utility measures Un comparación de la calidad de vida de personas con
11 502 trastornos auditivos estimada por tres diferentes medidas de utilidad. *International Journal of*
12 503 *Audiology*. 2005;44(3):157-63.
- 13 504 26. Barton GR, Bankart J, Davis AC, Summerfield QA. Comparing utility scores before and after
14 505 hearing-aid provision. *Applied health economics and health policy*. 2004;3(2):103-5.
- 15 506 27. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart
16 507 disease. *Health and Quality of Life Outcomes*. 2006;4(1):20.
- 17 508 28. Stavem K, Frøland SS, Hellum KB. Comparison of preference-based utilities of the 15D, EQ-5D
18 509 and SF-6D in patients with HIV/AIDS. *Quality of Life Research*. 2005;14(4):971-80.
- 19 510 29. Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based
20 511 on hypothetical preferences in a general population. *Health economics*. 2005;14(11):1169-89.
- 21 512 30. McDonough CM, Tosteson AN. Measuring preferences for cost-utility analysis.
22 513 *Pharmacoeconomics*. 2007;25(2):93-106.
- 23 514 31. Lwanga SK, Lemeshow S, Organization WH. Sample size determination in health studies: a
24 515 practical manual. 1991.
- 25 516 32. Senanayake S, Gunawardena N, Paliyawadana P, Bandara P, Haniffa R, Karunaratna R, et al.
26 517 Symptom burden in chronic kidney disease; a population based cross sectional study. *BMC*
27 518 *nephrology*. 2017;18(1):228.
- 28 519 33. Dolan P. Modeling valuations for EuroQol health states. *Medical care*. 1997;35(11):1095-108.
- 29 520 34. Ramos-Goni JM, Rivero-Arias O. eq5d: A command to calculate index values for the EQ-5D
30 521 quality-of-life instrument. *Stata Journal*. 2011;11(1):120.
- 31 522 35. Lamers L, Bouwmans C, van Straten A, Donker M, Hakkaart L. Comparison of EQ-5D and SF-6D
32 523 utilities in mental health patients. *Health economics*. 2006;15(11):1229-36.
- 33 524 36. Harrison M, Davies L, Bansback N, McCoy M, Verstappen S, Watson K, et al. The comparative
34 525 responsiveness of the EQ-5D and SF-6D to change in patients with inflammatory arthritis. *Quality of*
35 526 *Life Research*. 2009;18(9):1195.
- 36 527 37. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of
37 528 clinical measurement. *The lancet*. 1986;327(8476):307-10.
- 38 529 38. Tredoux C, Durrheim K. Numbers, hypotheses and conclusions: A course in statistics for the
39 530 social sciences. Cape Town: University of Cape Town Press[Links]. 2002.
- 40 531 39. Levey AS, Stevens LA, Coresh J. Conceptual model of CKD: applications and implications.
41 532 *American journal of kidney diseases*. 2009;53(3):S4-S16.
- 42 533 40. Seliger SL, Zhan M, Walker LD, Fink JC. Chronic kidney disease adversely influences patient
43 534 safety. *Journal of the American Society of Nephrology*. 2008;19(12):2414-9.
- 44 535 41. Stevens PE, Levin A. Evaluation and management of chronic kidney disease: synopsis of the
45 536 kidney disease: improving global outcomes 2012 clinical practice guideline. *Annals of internal*
46 537 *medicine*. 2013;158(11):825-30.
- 47 538 42. Tseng C-L, Lafrance J-P, Lu S-E, Soroka O, Miller DR, Maney M, et al. Variability in estimated
48 539 glomerular filtration rate values is a risk factor in chronic kidney disease progression among patients
49 540 with diabetes. *BMC nephrology*. 2015;16(1):34.
- 50 541 43. Kularatna S, Byrnes J, Chan YK, Ski CF, Carrington M, Thompson D, et al. Comparison of the
51 542 EQ-5D-3L and the SF-6D (SF-12) contemporaneous utility scores in patients with cardiovascular
52 543 disease. *Quality of life research : an international journal of quality of life aspects of treatment, care*
53 544 *and rehabilitation*. 2017;26(12):3399-408.
- 54 545 44. Cohen J. A power primer. *Psychological bulletin*. 1992;112(1):155.
- 55 546 45. Litwin MS, Fink A. How to measure survey reliability and validity: Sage; 1995.

- 1 547 46. Sack TH, Barton GR, Jenkinson C, Doherty M, Avery AJ, Muir KR. Comparing cost-utility
2 548 estimates: does the choice of EQ-5D or SF-6D matter? *Med Care*. 2009;889-95.
- 3 549 47. Grieve R, Grishchenko M, Cairns J. SF-6D versus EQ-5D: reasons for differences in utility scores
4 550 and impact on reported cost-utility. *The European journal of health economics*. 2009;10(1):15-23.
- 5 551 48. Stewart ST, Cutler DM, Rosen AB. Comparison of Trends in US Health-Related Quality of Life
6 552 over the 2000's Using the SF-6D, HALex, EQ-5D, and EQ-5D Visual Analog Scale versus a Broader Set
7 553 of Symptoms and Impairments. *Medical care*. 2014;52(12):1010.
- 8 554 49. Zhao FJ, Yue M, Yang H, Wang T, Wu Jh, Li SC. Validation and comparison of EuroQol and short
9 555 form 6D in chronic prostatitis patients. *Value in Health*. 2010;13(5):649-56.
- 10 556 50. Cunillera O, Tresserras R, Rajmil L, Vilagut G, Brugulat P, Herdman M, et al. Discriminative
11 557 capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey.
12 558 *Quality of Life Research*. 2010;19(6):853-64.
- 13 559 51. Bharmal M, Thomas J. Comparing the EQ-5D and the SF-6D descriptive systems to assess their
14 560 ceiling effects in the US general population. *Value in Health*. 2006;9(4):262-71.
- 15 561 52. Janssen M, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement
16 562 properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country
17 563 study. *Quality of Life Research*. 2013;22(7):1717-27.
- 18 564 53. Bharmal M, Thomas III J. Comparing the EQ-5D and the SF-6D descriptive systems to assess
19 565 their ceiling effects in the US general population. *Value Health*. 2006;9(4):262-71.
- 20 566 54. de Boer IH, Olan GA, Patel UD. Screening for chronic kidney disease. *Ann Intern Med*.
21 567 2013;158(5 Pt 1):362-3.
- 22 568 55. Hummel A. Chronic renal failure, a frequent and asymptomatic pathology. *Soins; la revue de*
23 569 *reference infirmiere*. 2010(745):24-6.
- 24 570 56. Jesky MD, Dutton M, Dasgupta I, Yadav P, Ng KP, Fenton A, et al. Health-related quality of life
25 571 impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic kidney disease:
26 572 a prospective observational study. *PloS one*. 2016;11(11):e0165675.
- 27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

575 **Tables**

576

577 **Table 1 : Demographic distribution of the sample by the EQ-5D-3L and the SF-6D utility**
578 **scores**

Variable	N (%)	EQ-5D-3L utility mean (SD)	Sf-6D utility mean (SD)	p value [#]
All sample	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*

579 [#] Wilcoxon signed-rank test ; * Significant at p<0.05

580

581 **Table 2 : Distribution of the sample by the EQ-5D-3L dimensions**

	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)

582

583

584 **Table 3 : Distribution of the sample by the SF-6D dimensions**

	Physical functioning (%)	Role limitation (%)	Social functioning (%)	Pain (%)	Mental health (%)	Vitality (%)
1 ^a	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 ^b	319 (29.1)	NA	NA	240 (21.9)	NA	NA

585 ^a No problems586 ^b Severe problems

587

588

1 589

2 3 4 590 **Table 4 : Intra Class Correlation 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*

22 591 * Significant at p<0.05 level

23 592

24 593

25 26 27 594 **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	0.071	254 (24.0)	0.551 (0.10)	0.570	<0.001	0.141
IV	614 (58.1)	0.566 (0.42)	0.620			614 (58.1)	0.536 (0.09)	0.560		
V	151 (14.3)	0.467 (0.42)	0.585			151 (14.3)	0.523 (0.08)	0.550		
Dialysis	38 (3.6)	0.126 (0.39)	-0.016			38 (3.6)	0.432 (0.07)	0.410		

28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 595 # Kruskal–Wallis test

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

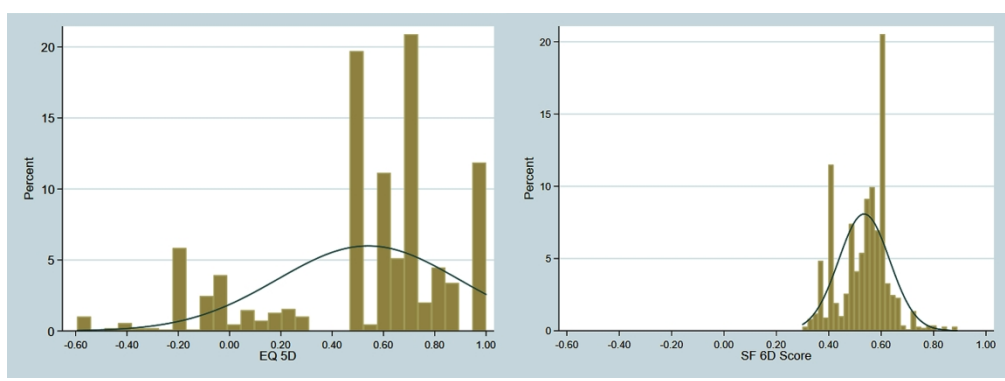


Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

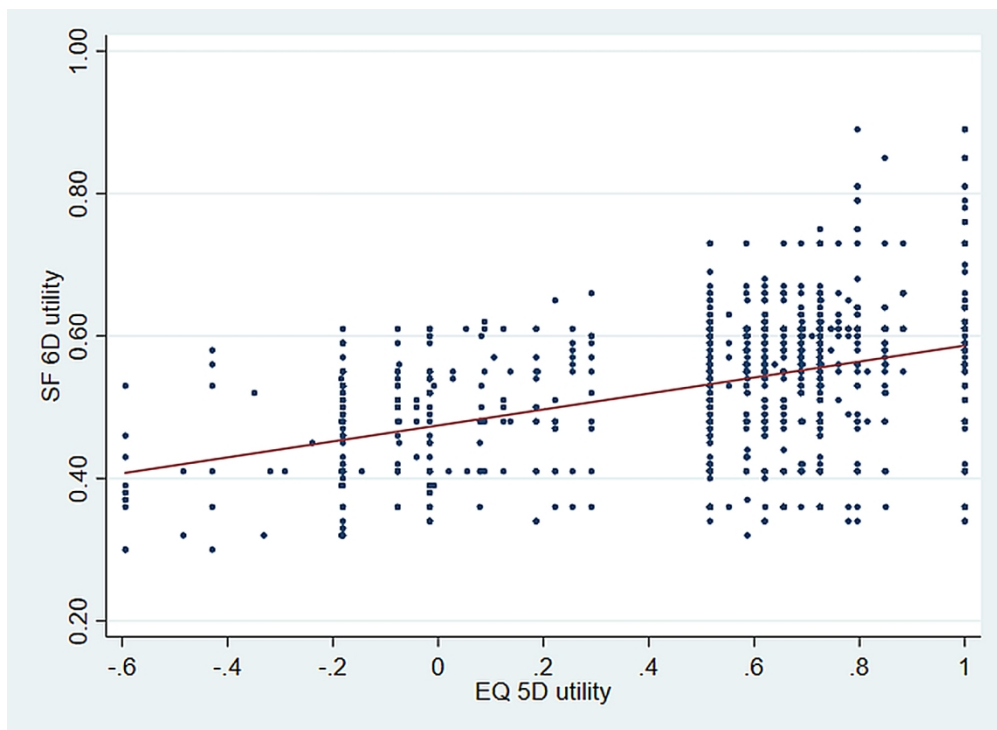


Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

137x99mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

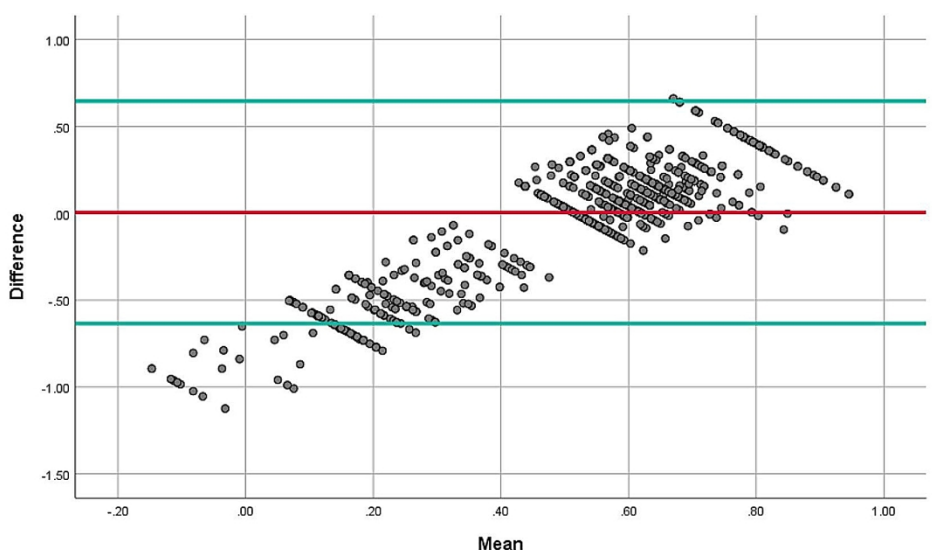


Fig. 3 : Bland and Altman plot of differences between EQ-5D and SF-6D for patients with CKD
166x90mm (300 x 300 DPI)

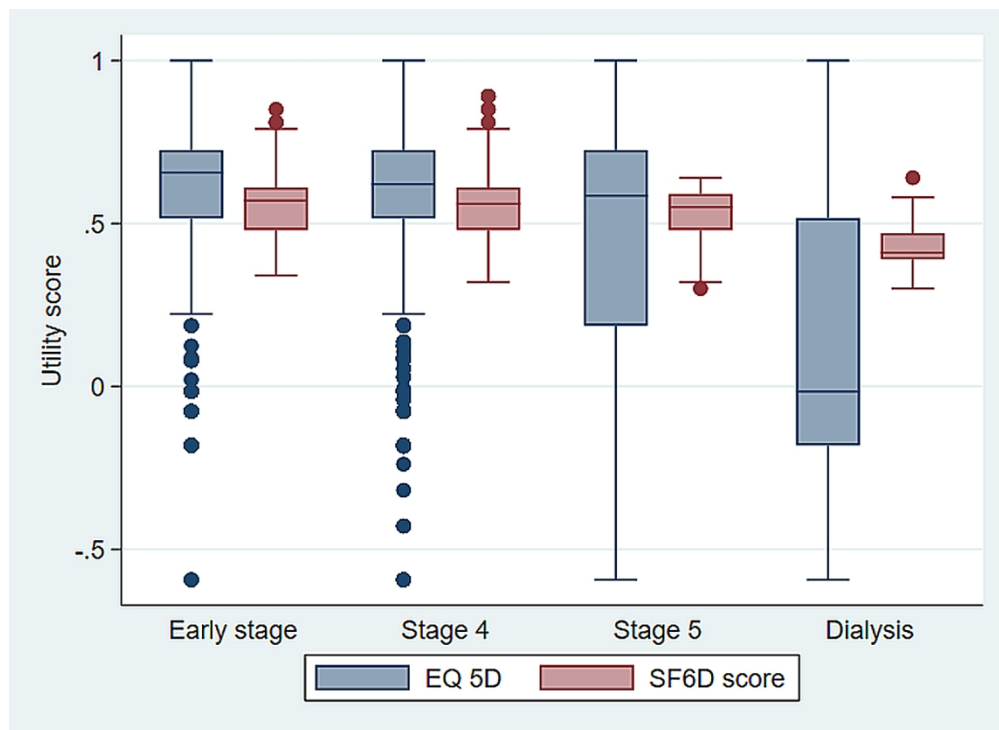


Fig 4: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD stage

137x99mm (300 x 300 DPI)

STROBE Statement—Checklist of items that should be included in reports of *cross-sectional studies*

	Item No	Recommendation	Line number
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	16
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	67
Objectives	3	State specific objectives, including any prespecified hypotheses	115
Methods			
Study design	4	Present key elements of study design early in the paper	121
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	121-126
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants	129-134
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	157-190
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	157-190
Bias	9	Describe any efforts to address potential sources of bias	NA
Study size	10	Explain how the study size was arrived at	123
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	157-190
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	193-207
		(b) Describe any methods used to examine subgroups and interactions	NA
		(c) Explain how missing data were addressed	NA
		(d) If applicable, describe analytical methods taking account of sampling strategy	NA
		(e) Describe any sensitivity analyses	NA
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	266-267
		(b) Give reasons for non-participation at each stage	267
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	266 - 274
		(b) Indicate number of participants with missing data for each variable of interest	NA

Outcome data	15*	Report numbers of outcome events or summary measures	NA
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	NA
		(b) Report category boundaries when continuous variables were categorized	Table 1
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	335 - 340
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	397
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	405 - 413
Generalisability	21	Discuss the generalisability (external validity) of the study results	410 - 413
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	425

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.