BMJ Open

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

Journal:	*BMJ Open*
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 Technoloy 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)

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

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

[1]*Australian Centre for Health Services Innovation, Queensland University of Technology*

[2] *World Health Organization, Sri Lanka*

Corresponding author;

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

Word count : 3277

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.

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).

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

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

Methods

Patients selection

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

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

The data collectors assessed the eligibility of patients by reviewing their clinical records. Informed consent was obtained from those who were eligible for participation in the study.

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

Calculation of utility scores

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

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

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

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

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

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

Data analysis

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

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

Agreement and differences

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

Correlation

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

(EQ-5D-3L) and mental health (SF-6D)/anxiety and depression (EQ-5D-3L). The vitality dimension of the SF-6D did not have any related dimension with the EQ-5D-3L. The magnitude of the correlation coefficients 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

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

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

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

Agreement, differences and correlation between the two utility scores

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

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

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

Discrimination

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

Test-retest reliability

The test re-test Spearman's correlations was more than 0.9 for both the instruments indicating good test re-test reliability.

Discussion

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

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

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

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

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

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

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

Limitations

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

Conclusions

The correlation between the scores was moderate. Both tools were able to discriminate advancement of CKD stages. Effect size, which denoted the discriminating ability of the different CKD stages, is highest when disease condition is advanced and the highest effect size was seen in SF-6D. Findings indicate that both 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.

Ethics approval and consent to participate

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

Funding

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

Competing interests: None declared.

Patient consent : Obtained.

Author Contributions

SK and SS : Research idea, study design, statistical analysis and drafting of the manuscript. NG[2]: study design, data analysis/interpretation. NG[1]: participated in study design, data interpretation and supervision. All authors read and approved the final manuscript.

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

Data sharing statement

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

1. Senanayake S, Gunawardena N, Palihawadana 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: Institue 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, Tangtiphaiboontana 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.
22. Kwakkenbos L, Fransen J, Vonk MC, Becker ES, Jeurissen M, van den Hoogen FH, et al. A comparison of the measurement properties and estimation of minimal important differences of the EQ-5D and SF-6D utility measures in patients with systemic sclerosis. Clinical and experimental rheumatology. 2013;31(2 Suppl 76):50-6.
23. Rowen D, Young T, Brazier J, Gaugris S. Comparison of generic, condition-specific, and mapped health state utility values for multiple myeloma cancer. Value in Health. 2012;15(8):1059-68.
24. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D and SF-6D utilities in mental health patients. Health economics. 2006;15(11):1229-36.
25. Barton GR, Bankart J, Davis AC. A comparison of the quality of life of hearing-impaired people as estimated by three different utility measures Un comparación de la calidad de vida de personas con trastornos auditivos estimada por tres diferentes medidas de utilidad. International Journal of Audiology. 2005;44(3):157-63.
26. Barton GR, Bankart J, Davis AC, Summerfield QA. Comparing utility scores before and after hearing-aid provision. Applied health economics and health policy. 2004;3(2):103-5.
27. van Stel HF, Buskens E. Comparison of the SF-6D and the EQ-5D in patients with coronary heart disease. Health and Quality of Life Outcomes. 2006;4(1):20.

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

45. Cunillera O, Tresserras R, Rajmil L, Vilagut G, Brugulat P, Herdman M, et al. Discriminative capacity of the EQ-5D, SF-6D, and SF-12 as measures of health status in population health survey. Quality of Life Research. 2010;19(6):853-64.

46. Bharmal M, Thomas J. Comparing the EQ□5D and the SF□6D descriptive systems to assess their ceiling effects in the US general population. Value in Health. 2006;9(4):262-71.

47. Janssen M, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Quality of Life Research. 2013;22(7):1717-27.

48. de Boer IH, Olan GA, Patel UD. Screening for chronic kidney disease. Ann Intern Med. 2013;158(5 Pt 1):362-3.

49. Hummel A. Chronic renal failure, a frequent and asymptomatic pathology. Soins; la revue de reference infirmiere. 2010(745):24-6.

50. Jesky MD, Dutton M, Dasgupta I, Yadav P, Ng KP, Fenton A, et al. Health-related quality of life impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic kidney disease: a prospective observational study. PloS one. 2016;11(11):e0165675.

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)

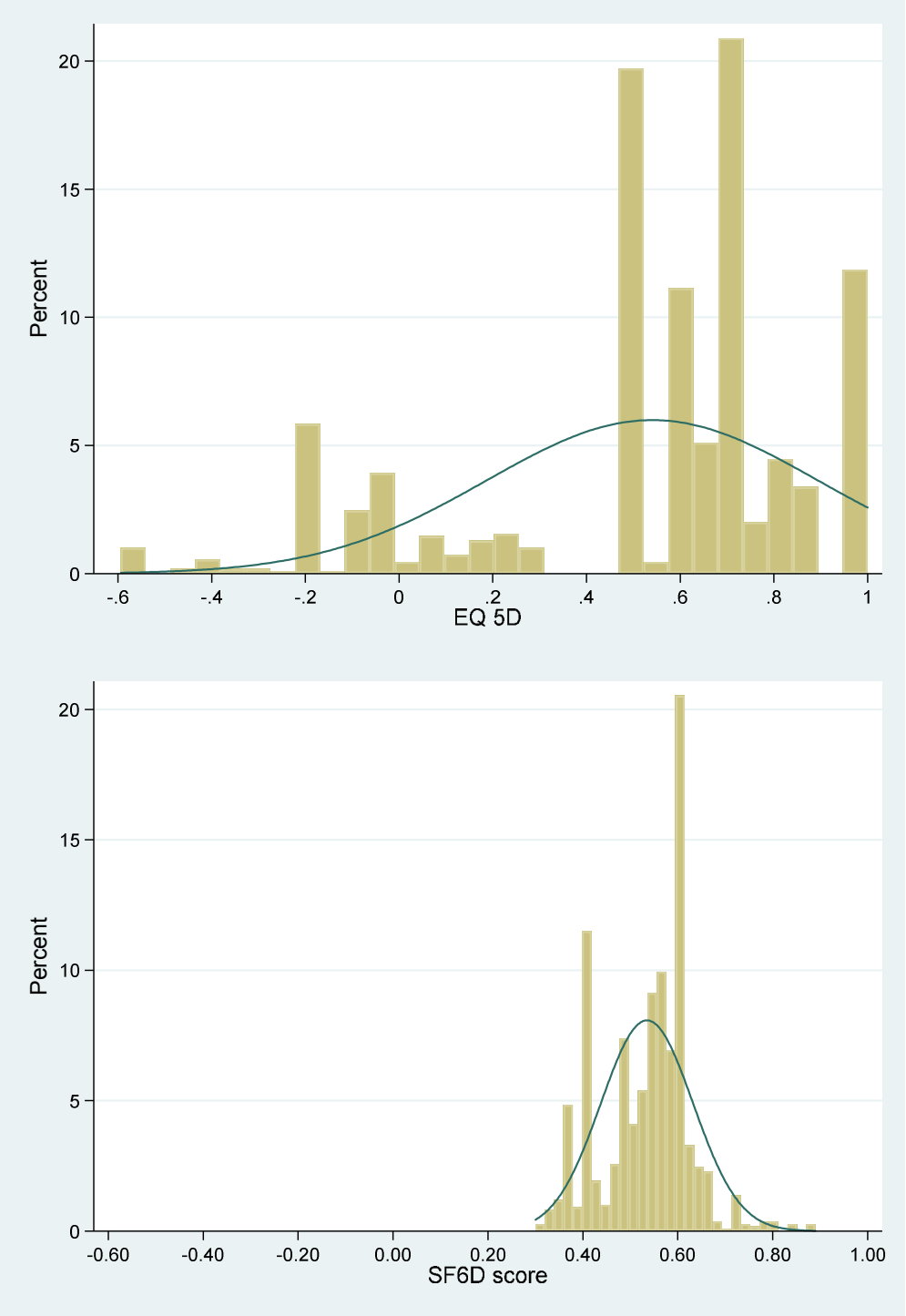

Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

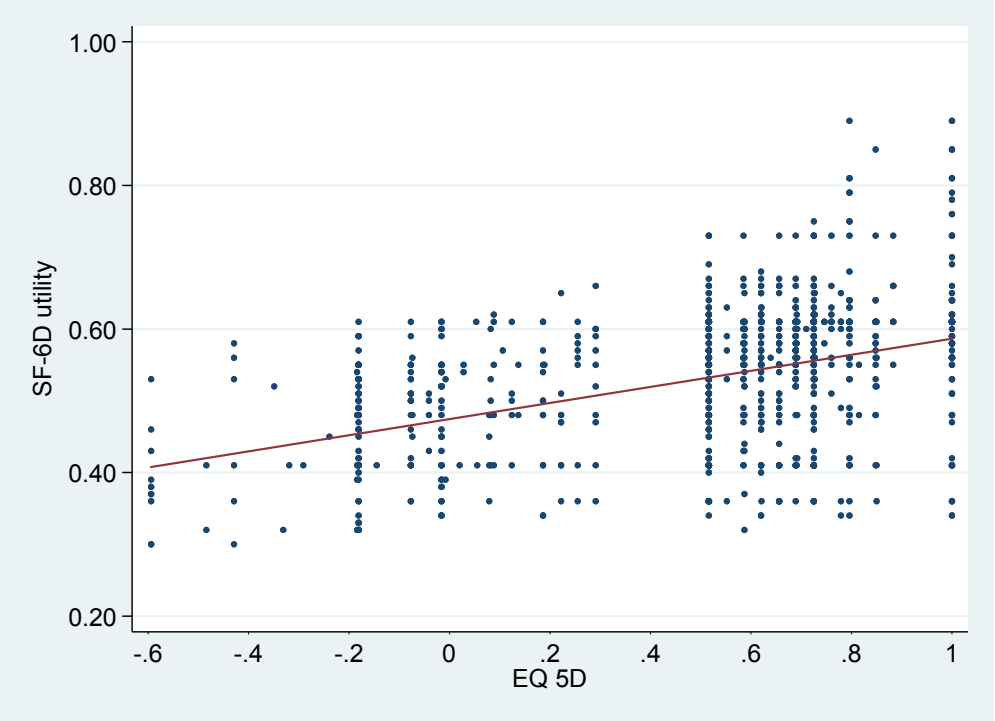

Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

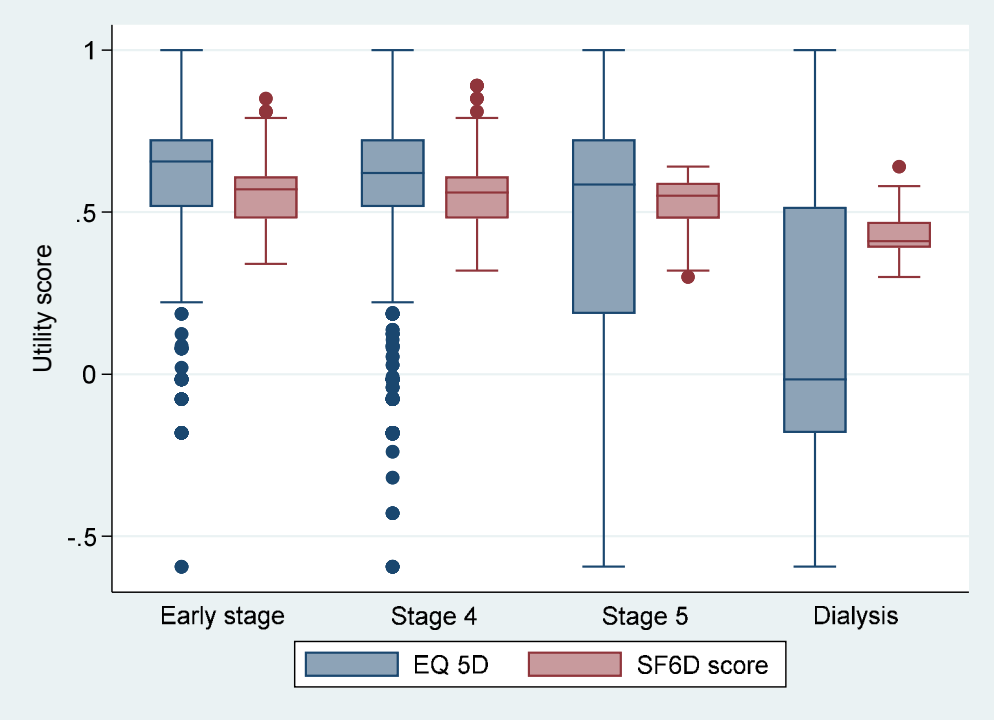

Fig 3: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD stage

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:	*BMJ Open*
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 Technoloy 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)

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

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

[1]Australian Centre for Health Services Innovation, Queensland University of Technology

[2] World Health Organization, Sri Lanka

Corresponding author;

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

Word count : 3277

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 may 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. SF-6D was constructed from the SF-36 according to the published algorithm. The study assessed discrimination, correlation and differences across the two instruments.

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

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

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

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.

Introduction

Chronic kidney disease (CKD) is a substantial public health problem with adverse psychological, physical and economic outcomes. The burden of CKD is increasing globally(1). World Health Report (2002) and Global Burden of Disease (GBD) project stated that the diseases of the kidney contribute much to the global disease burden with approximately 850,000 deaths every year globally (2). Furthermore, according to GBD study conducted in 2010, of the top causes of 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 poses a 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 have demonstrated 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 best imaginable health and 0 refers to death (11). Measuring utilities for different health states is complex and time-consuming. Thus, Multi Attribute Utility Instruments (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).

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

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

Methods

Participant selection

A population-based descriptive cross-sectional study was conducted in the district of Anuradhapura in the North Central Province (NCP) of Sri Lanka between Septembers to December 2015. The study population consisted of 1162 confirmed CKD patients, calculated using the appropriate formula (31), who were over 18 years old with documented evidence of CKD living in the Anuradhapura district. The diagnosis of CKD was made if the Glomerular Filtration Rate (GFR) was less than 60 ml/min per 1.73m^2 of body surface area in two measurements made three months apart.

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

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

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

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

Calculation of utility scores

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

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

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

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

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

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

Data analysis

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

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

Agreement and differences

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

Intra Class Correlation (ICC)

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

Discrimination

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

Glomerular Filtration Rate (GFR) is the most important indicator of kidney function of patients with CKD (39). Studies have shown that decreased GFR is associated with infection, impaired cognitive and physical function as well as threats to patient safety (40). 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 (41, 42). Depending on the GFR value, CKD is categorised into five stages; stage I to stage V. For analytical purposes, the CKD stages I to III were categorised as "early stage" in the present study. It is expected that with advanced stages of the disease, the utility scores should be lower than the early stages.

Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using the non-parametric test, Kruskal–Wallis, and effect size. The instrument's ability to discriminate between two adjacent stages was estimated by calculating the effect size. The effect size was calculated by dividing the mean difference of two adjacent CKD stages by the standard deviation of the milder of the two CKD stages (23, 43). 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) (44).

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 ICC and a value of 0.70 or greater was considered as satisfactory reliability (45).

Patient and Public Involvement

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

Results

Sample characteristics

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

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

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

Agreement, differences and correlation between the two utility scores

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

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

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

Discrimination

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

Test-retest reliability

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

Discussion

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

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

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

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

The present study did not find any difference (p=0.588) between the overall mean scores of the two utility instruments. This was similar to a study conducted among a group of HIV/AIDS patients (28), but different from other studies available in the literature where different results have been reported. Significantly higher utility values for EQ-5D-3L were found among general population (29, 48), cardio-vascular disease patients (19), rheumatoid 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 obtained for SF-6D instrument (24). These varying results could be due to different recall periods of the two instruments. EQ-5D-3L assessed the health status of the day of instrument administration while SF-6D, which was derived from SF-36, assessed the health status of the past 30 days.

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

Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using ANOVA and effect size. In both MAUIs, the utility differences across CKD stages were statistically significant ($p<0.05$; ANOVA) indicating good discrimination. However, the effect size was small for both the tools until the dialysis stage. At the dialysis stage, the effect

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

Limitations

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

Conclusions

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

Ethics approval and consent to participate

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

Acknowledgement:

Authors acknowledge Asanga Ranasinghe, Priyantha Kumara and Anura Ranasinghe for the support rendered during the study.

Funding

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

Competing interests: None declared.

Patient consent : Obtained.

Author Contributions

SK and SS : Research idea, study design, statistical analysis and drafting of the manuscript. NG^2: study design, data analysis/interpretation. NG^1: participated in study design, data interpretation and supervision. All authors read and approved the final manuscript.

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

Data sharing statement

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Figure legend

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

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

References

1. Senanayake S, Gunawardena N, Palihawadana 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: Institue 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, Tangtiphaiboontana 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.

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

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

Tables

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

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

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

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 : 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*

* Significant at p<0.05 level

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			254 (24.0)	0.551 (0.10)	0.570		
IV	614 (58.1)	0.566 (0.42)	0.620	<0.001	0.071	614 (58.1)	0.536 (0.09)	0.560	<0.001	0.141
V	151 (14.3)	0.467 (0.42)	0.585		0.305	151 (14.3)	0.523 (0.08)	0.550		0.138
Dialysis	38 (3.6)	0.126 (0.39)	-0.016		0.807	38 (3.6)	0.432 (0.07)	0.410		1.098

[#] Kruskal–Wallis test

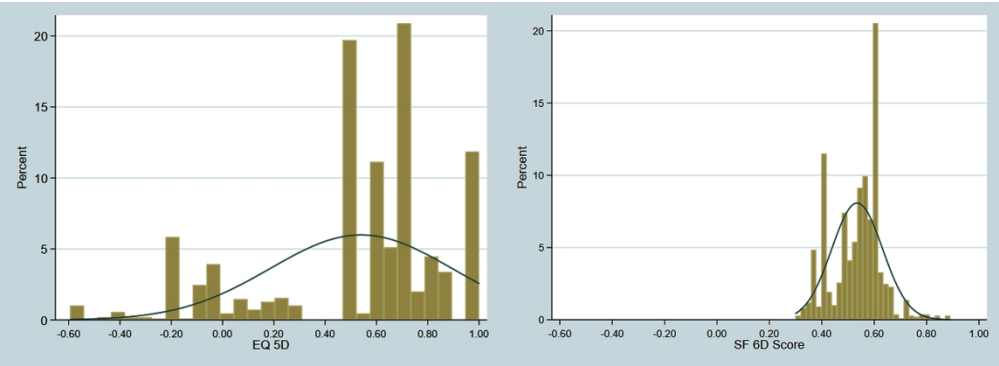

Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

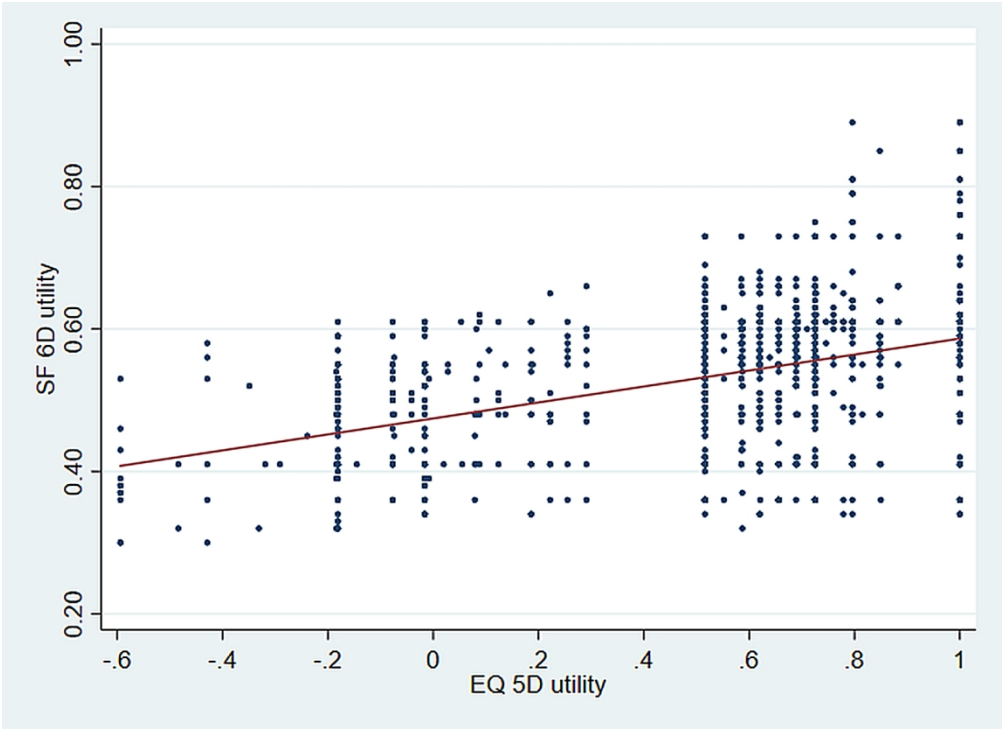

Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

137x99mm (300 x 300 DPI)

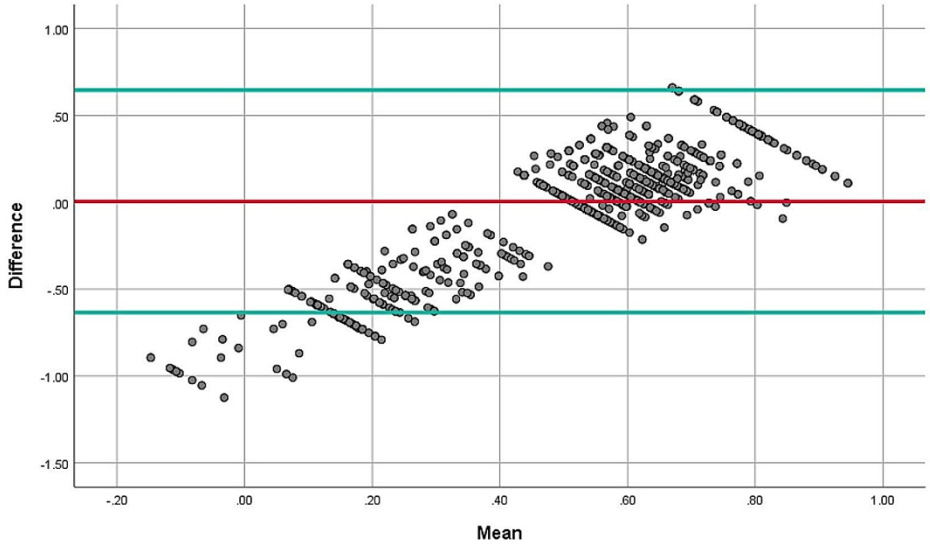

Fig. 3 : Bland and Altman plot of differences between EQ-5D and SF-6D for patients with CKD

166x90mm (300 x 300 DPI)

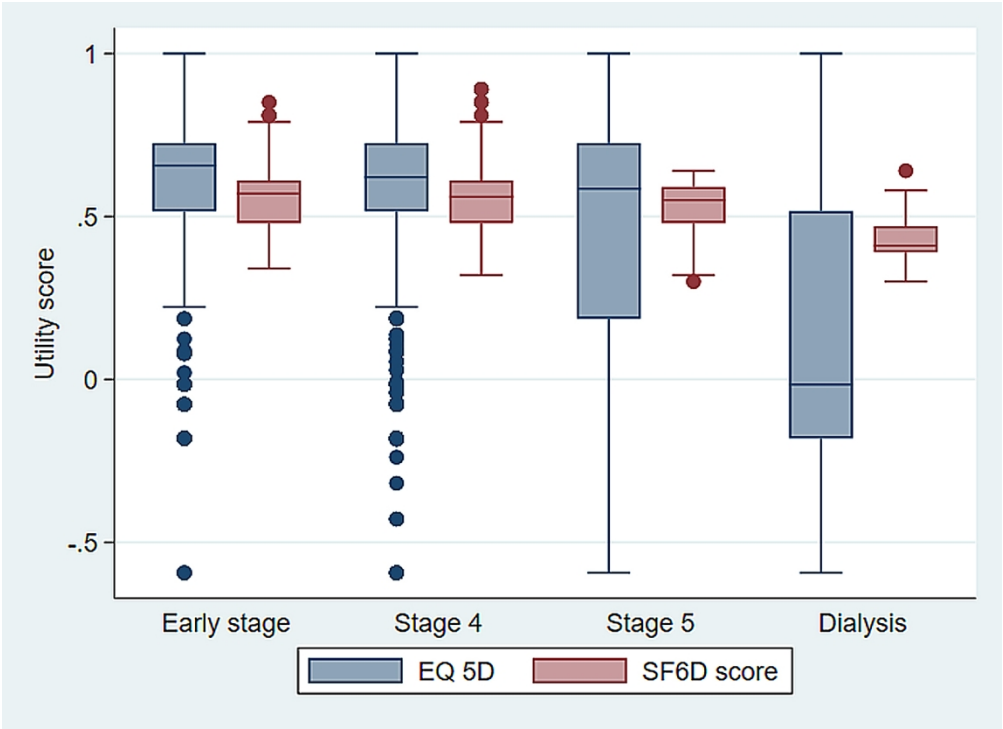

Fig 4: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD stage

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:	*BMJ Open*
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)

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

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

[1]Australian Centre for Health Services Innovation, Queensland University of Technology

[2] World Health Organization, Sri Lanka

Corresponding author;

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

Word count : 3277

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

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

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

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

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.

Introduction

Chronic kidney disease (CKD) is a substantial public health problem with adverse psychological, physical and economic outcomes. The burden of CKD is increasing globally(1). World Health Report (2002) and Global Burden of Disease (GBD) project stated that the diseases of the kidney contribute much to the global disease burden with approximately 850,000 deaths every year globally (2). Furthermore, according to GBD study conducted in 2010, of the top causes of 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 poses a 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 have demonstrated 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 best imaginable health and 0 refers to death (11). Measuring utilities for different health states is complex and time-consuming. Thus, Multi Attribute Utility Instruments (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).

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

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

Methods

Participant selection

A population-based descriptive cross-sectional study was conducted in the district of Anuradhapura in the North Central Province (NCP) of Sri Lanka between Septembers to December 2015. The study population consisted of 1162 confirmed CKD patients, calculated using the appropriate formula (31), who were over 18 years old with documented evidence of CKD living in the Anuradhapura district. The diagnosis of CKD was made if the Glomerular Filtration Rate (GFR) was less than 60 ml/min per 1.73m^2 of body surface area in two measurements made three months apart.

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

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

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

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

Calculation of utility scores

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

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

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

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

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

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

Data analysis

STATA 15.1 software was used for the analysis. Mean utility scores on each instrument were compared by socio-demographic characteristics. Normality of the two distributions were assessed using Kolmogorov-Smirnov and Shapiro-Wilk tests. Wilcoxon signed-rank test was used to assess the difference between the two instruments in each socio-demographic class (35). Histograms were plotted for the two utility values distribution. Floor effects and ceiling (proportion of patients with the highest and lowest possible scores respectively) were calculated for the EQ-5D-3L and SF-6D. Ceiling and floor effects were considered small if ≤15% of patients occupy the best or worst health states, but they were considered serious if >15% of patients occupy these states (36).

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

Agreement and differences

The Wilcoxon Signed-Ranks Test was used to assess the overall difference between the EQ-5D-3L and SF-6D utility scores and the difference of the utility scores according to different socio-demographic and disease related features. Furthermore, the distribution of the

responses to the different domains of the two instruments was tabulated to present the agreement and the differences between the two instruments. Bland-Altman plot was also used to assess the proportional error and the limit of agreement (37).

Intra Class Correlation (ICC)

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

Discrimination

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

Glomerular Filtration Rate (GFR) is the most important indicator of kidney function of patients with CKD (39). Studies have shown that decreased GFR is associated with infection, impaired cognitive and physical function as well as threats to patient safety (40). 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 (41, 42). Depending on the GFR value, CKD is categorised into five stages; stage I to stage V. For analytical purposes, the CKD stages I to III were categorised as "early stage" in the present study. It is expected that with advanced stages of the disease, the utility scores should be lower than the early stages.

Discrimination of EQ-5D-3L and SF-6D for different CKD stages was examined using the non-parametric test, Kruskal–Wallis, and effect size. The instrument's ability to discriminate between two adjacent stages was estimated by calculating the effect size. The effect size was calculated by dividing the mean difference of two adjacent CKD stages by the standard deviation of the milder of the two CKD stages (23, 43). 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) (44).

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 ICC and a value of 0.70 or greater was considered as satisfactory reliability (45).

Patient and Public Involvement

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

Results

Sample characteristics

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

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

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

Agreement, differences and correlation between the two utility scores

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

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

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

The correlation between EQ-5D-3L and SF-6D was 0.408, which was statistically significant at p<0.001 level (Figure 2). Regarding the ICC between different domains of the two instruments, according to the Guilford's criteria, moderate correlation (0.4-0.6) was evident between Social functioning and Mobility (0.517), Social functioning and Self-care (0.424), Social functioning and Usual activities (0.464), Social functioning and Pain/discomfort (0.566), Social functioning and Anxiety/depression (0.528), Pain and Mobility (0.475), Pain and Pain/ discomfort (0.482), Pain and Anxiety/depression (0.484), Vitality and Pain/ discomfort (0.475) and Vitality and Anxiety/depression (0.453) (Table 4). The Bland-Altman plot showed proportional error and wide limits of agreement (Figure 3).

Discrimination

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

Test-retest reliability

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

Discussion

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

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

At present, there is no consensus on the methodology to compare the utility scores of different MAUIs (19, 35). The present study adopted the methodologies used by Kularatna et at. (2017) and Lamers et al. (2006) (19, 35). Only one time assessment of the utilities was done in the present study. Thus, the responsiveness of the two instruments to changes in kidney function over time was not assessed. Though Sri Lankan EQ-5D-3L utility scores are available (18), yet we used the UK utility scores for the EQ-5D-3L (33) because of the unavailability of comparable Sri Lankan SF-6D utility scores values. This is an accepted method of calculating the utility scores in the absence of country specific utilities. Two studies conducted in Netherlands (24) and Italy (21), comparing the utility scores of the two instruments, had used the UK derived EQ-5D-3L and SF-6D utility scores.

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

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

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

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

Limitations

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

Conclusions

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

Ethics approval and consent to participate

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

Acknowledgement:

Authors acknowledge Asanga Ranasinghe, Priyantha Kumara and Anura Ranasinghe for the support rendered during the study.

Funding

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

Competing interests: None declared.

Patient consent : Obtained.

Author Contributions

SK and SS : Research idea, study design, statistical analysis and drafting of the manuscript. NG[2]: study design, data analysis/interpretation. NG[1]: participated in study design, data interpretation and supervision. All authors read and approved the final manuscript.

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

Data sharing statement

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Figure legend

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

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

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

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

References

1. Senanayake S, Gunawardena N, Palihawadana 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: Institue 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, Tangtiphaiboontana 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.

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

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

Tables

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

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

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

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 : 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*

* Significant at $p<0.05$ level

Table 5 : Discrimination across clinical severity groups

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

Kruskal–Wallis test

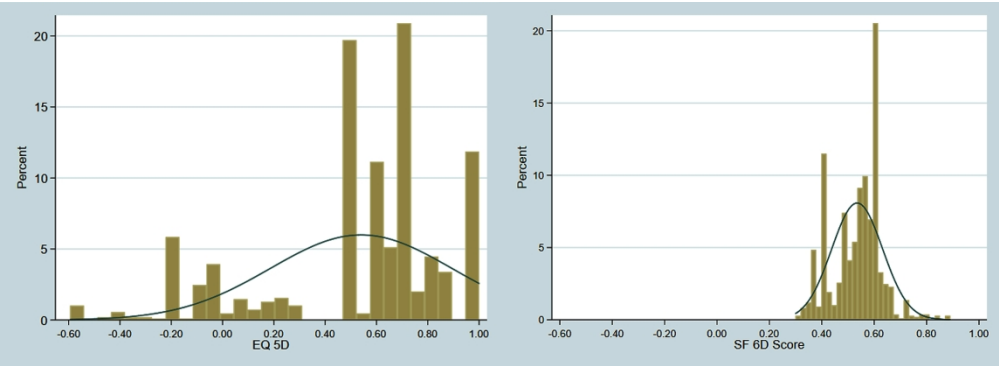

Fig. 1. Distribution of EQ-5D (A) and SF-6D (B)

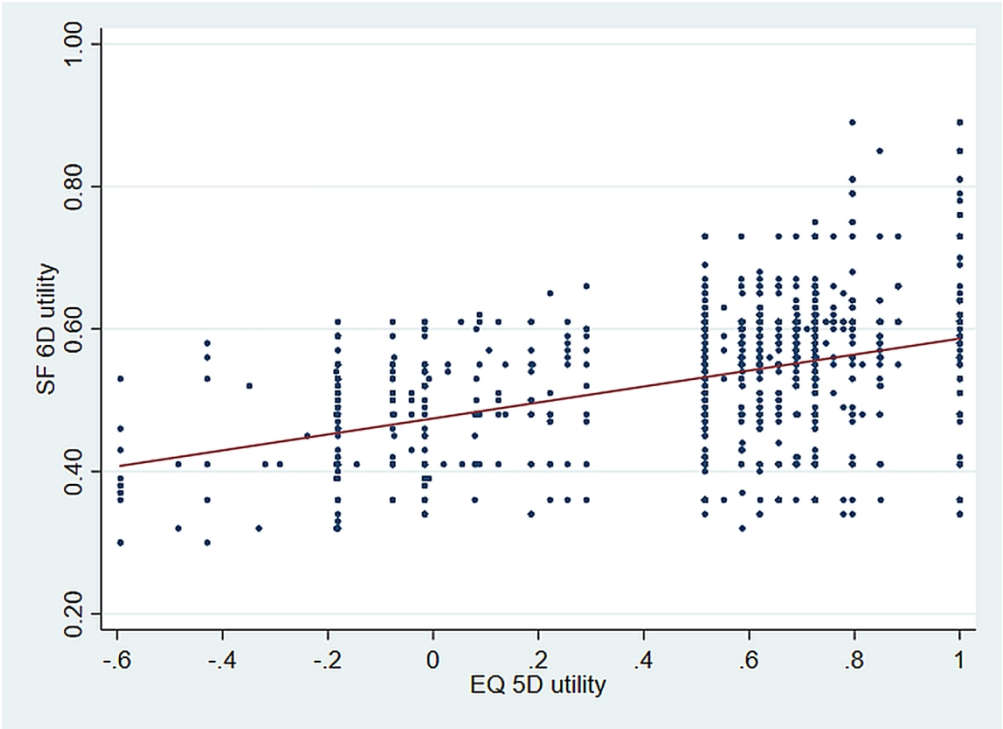

Fig. 2 Correlation between the EQ-5D-3L and the SF-6D utilities

137x99mm (300 x 300 DPI)

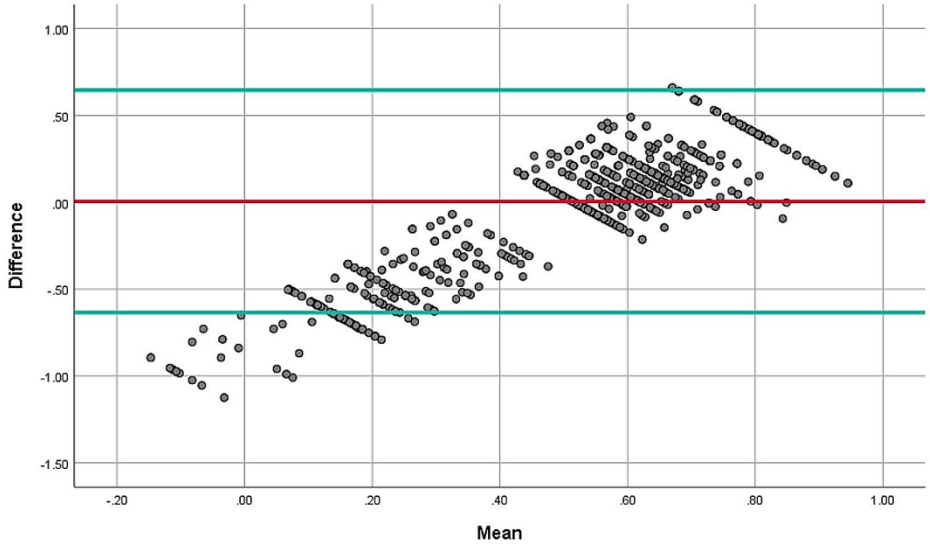

Fig. 3 : Bland and Altman plot of differences between EQ-5D and SF-6D for patients with CKD

166x90mm (300 x 300 DPI)

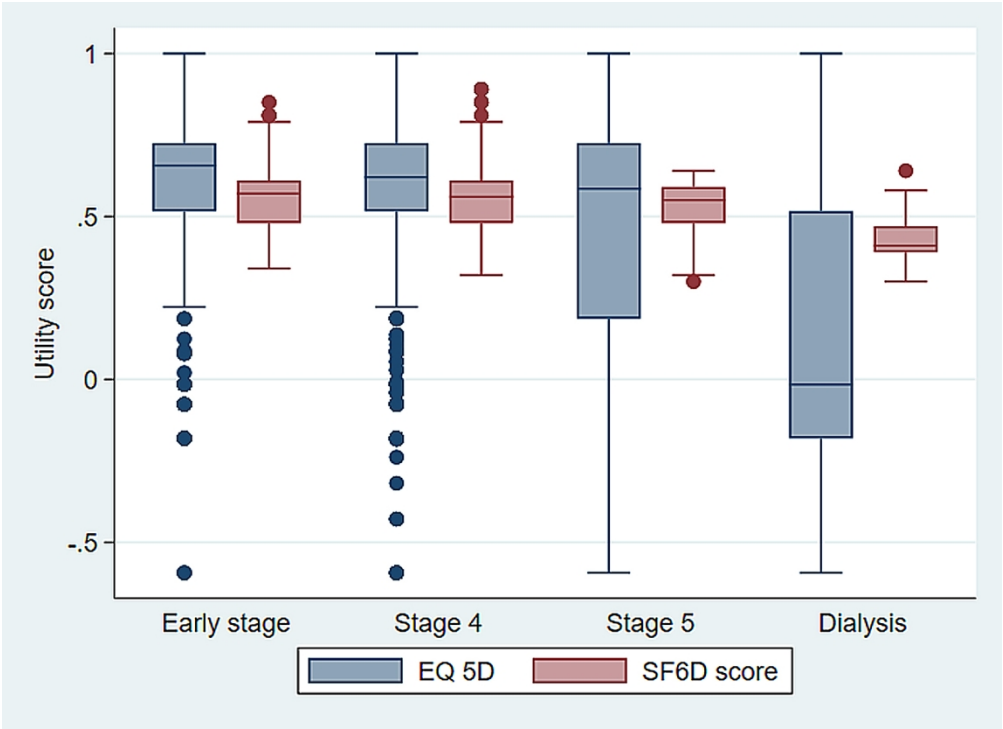

Fig 4: The box-plots present the median, quartiles and extreme values for the EQ-5D and SF-6D utility scores for CKD stage

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.