Depressive and anxiety symptoms in Dutch chronic kidney disease patients

Loosman, W.L.

2016

Chapter 3

Short Form 12 or Short Form 36 to measure quality of life changes in dialysis patients?

Wim L. Loosman, Tiny Hoekstra, Sandra van Dijk, Caroline B. Terwee, Adriaan Honig, Carl E.H. Siegert, and Friedo W. Dekker

Nephrology Dialysis Transplantation. 2015; 30: 1170-1176

ABSTRACT

Background Short Form 36 (SF-36) is a self-report health related quality of life (HRQOL) questionnaire, widely used in dialysis patients. It consists of physical and mental component scores (PCS/MCS), ranging from 0-100. To improve efficiency, Short Form 12 (SF-12) was developed to reproduce PCS and MCS. We assessed the ability of SF-12 versus SF-36 to detect change over time, and the association of SF12 versus SF36 with short-term and long-term mortality in dialysis patients.

Methods Patients were selected from NECOSAD (N=1,379), a prospective follow-up study among incident dialysis patients (62.1% haemodialysis) who completed SF-36 measurements every six months. Change in scores of SF-12 versus SF-36 were compared with intra class correlation coefficients (ICC). Subsequently, Bland-Altman plots were used to assess limits of agreement. Relationship with mortality was assessed with Cox models with and without a time-dependent variable, adjusted for age, sex, ethnicity, comorbidity and dialysis modality at baseline.

Results ICC for change in scores was 0.90 for MCS and 0.84 for PCS. Mean difference was -0.1 and 0.2 respectively, and limits of agreement were -8.3 to 8.4 for MCS change in scores and -8.8 to 9.2 for PCS. Adjusted hazard ratio's for mortality per 5 units increase were 0.87 (95% CI: 0.84-0.91) for MCS12, 0.87 (95% CI: 0.84-0.90) for MCS36, 0.79 (95% CI: 0.76-0.83) for PCS12 and 0.75 (95% CI: 0.71-0.78) for PCS36.

Conclusions SF-12 can be used to detect change in HRQOL in cohort studies on dialysis patients. SF-12 and SF-36 were similarly associated with short-term and long-term mortality. However, the wide limits of agreement indicate that SF-12 and SF-36 can give different scores on the individual level, suggesting that for individual purposes SF-36 instead of SF-12 should be used.

INTRODUCTION

Short Form 36 (SF-36) is a self-report questionnaire that is frequently being used to assess Health Related Quality Of Life (HRQOL). SF-36 has been validated in the general population as well as in various sub populations[1] including patients with end stage renal disease (ESRD).[2] From this questionnaire a Physical Component Score (PCS) and a Mental Component Score (MCS) can be calculated. The questionnaire can also be used to measure eight domains to give more specific information about a patient's HRQOL. The 36 items bring a burden on the patient, especially when patients need to fill out concomitant questionnaires in research settings. Therefore a shorter questionnaire, Short Form 12 (SF-12), has been developed.[3] This questionnaire consists of 12 items selected from SF-36 to calculate and reproduce PCS and MCS. With this short questionnaire it is not possible to accurately reproduce the eight domains.

SF-12 has been validated in the general population[4] and in various subgroups of patients such as patients with myocardial infarction[5] and ischemic stroke.[6] Three studies evaluated the use of SF-12 in patients with ESRD. First, an Iranian study of 140 dialysis patients showed high internal consistency and good test-retest reliability.[7] Second, an American study of 44,395 prevalent ESRD patients showed a correlation of 0.94 for both the physical as well as the mental component scores between both questionnaires.[8] These results were confirmed by a Norwegian study of 301 prevalent dialysis patients, which also showed correlations above 0.9.[9]

However, so far the validity of the SF-12 has only been determined in a cross-sectional manner. This poses a problem for using the SF-12 in clinical practice, as the validity is not evaluated for using the SF-12 as a screening instrument to detect changes over time. It is thus unknown whether the SF-12 is a good instrument to detect HRQOL changes over time, for example to study the effect of a new treatment. Furthermore, lower HRQOL has been associated with increased mortality in dialysis patients.[10-15] Previous studies showed comparable associations with mortality for the SF-12 as the SF-36.[8;9] Follow up time in these studies was fixed, ranging between one and 4.5 years. As a consequence, we do not know whether the SF-12 can also be used as a marker for short-term mortality in clinical care. There have been no studies that compared the associations of SF-12 and SF-36 with short-term mortality using updated component scores.

The objective of this study is (1) to compare the ability of SF-12 with SF-36 to detect change in HRQOL in dialysis patients, and (2) to compare the relationship between SF-12 and SF-36 scores with short-term and long-term mortality.

SUBJECTS AND METHODS

Study population

The study population was selected from the Netherlands Cooperative Study on the Adequacy of Dialysis (NECOSAD), a nationwide multi centre prospective follow up study among ESRD patients. From 1997 until 2007 incident patients were enrolled from 38 dialysis centres. Inclusion criteria were age above 18 years and starting renal replacement therapy for the first time. The study was approved by the Medical Ethics Committees of all participating hospitals. All patients gave written informed consent. To allow for stabilisation of laboratory data and treatment modality the baseline measures were taken at three months after the start of dialysis. The following patients demographics and clinical characteristics were collected at three months: gender, age, ethnicity, marital status, having children, employment status, dialysis modality, body mass index (BMI), the Davies comorbidity index indicating the level of comorbidity (no, intermediate, severe)[16], and primary causes of ESRD. Patients were followed until time of death, transplantation, 25 May 2009 or for this current analysis a maximum follow up of 4.5 years, whichever was earliest. For the current analysis patients were selected with available data on SF-36 at three months after start of dialysis.

Health related quality of life scores

The participants filled out the SF-36 questionnaire at three and six months after start of dialysis and thereafter every six months. SF-36 component scores (PCS36 and MCS36) were calculated as standardised scores towards the United States (US) population as a reference group. This reference population has a mean score of 50 with a standard deviation (SD) of 10.[17] Thus, a score below or above 50 indicates a respectively worse or better quality of life than the reference group. Furthermore, a higher score indicates a better HRQOL. SF-12 scores were obtained by using 12 questions from SF-36. PCS12 and MCS12 scores were calculated with orthogonal regression weights, and are also presented towards the US population as a reference group.[4;18] Of both SF-36 and SF-12 questionnaires version 1 was used.

The EuroQol is a questionnaire to assess HRQOL and has been used in dialysis patients to assess HRQOL for economic evaluations.[19-22] The questionnaire consists of a classification system and a health thermometer. The classification system consists of five questions which covers five domains of health (mobility, self-care, usual activities, pain/discomfort, anxiety/depression). The health thermometer is a visual analogue scale on which patients had to indicate their health status on a thermometer scale of 0-100 after answering the following question: 'what is your health status today?'. A score of 100 indicates the best health status. The EuroQol thermometer was assessed at the same measurement points as the SF-36. For the current analysis the change of the health thermometer score was used as an external criterion for changes in health related quality of life.

Statistical analysis

The time course of the SF-12 and SF-36 component scores were described using linear mixed models. Changes in scores were calculated between each consecutive measure point for the SF-12, SF-36 and the EuroQol thermometer scores. Intra class correlation coefficients (ICC) were calculated between the component scores derived from the SF-12 and SF-36 at the different measurement points as well as for the change in scores. In addition, ICCs were calculated between both SF-12 and SF-36 change in scores and the EuroQol thermometer change in scores.

To assess agreement between the scores of the SF-12 and SF-36 we used the Bland-Altman plot. [23] In this scatter plot individual differences between SF-12 and SF-36 scores are plotted against the mean of both measurements. The graph shows the population average of all individual differences between SF-12 and SF-36 scores and the 95% limits of agreement. These limits of agreement were calculated as the mean difference between SF-12 and SF-36 plus or minus 1.96 standard deviations of the differences. To explore whether the degree of disagreement on an individual level are driven by subgroups of patients we compared characteristics of patients with large individual differences between the two scores (> 5 points on both PCS and MCS) with the group of patients with good agreement. Baseline characteristics were compared between both groups. Survival time was calculated as the difference between the first measurement (i.e. 3 months on dialysis) and date of death or the last follow-up recorded. Follow-up was assessed by employing the reverse Kaplan-Meier method.[24] Patients were censored after death, transplantation, and loss of follow-up. Cox proportional hazard models were used to determine the associations between the different component scores and all-cause mortality. The baseline scores were used to determine long-term mortality. The proportional hazard assumption was assessed graphically by creating cumulative hazard plots of SF-36 and SF-12 scores, and by adding interaction terms with time. Short-term mortality risk was determined using a time-dependent Cox proportional hazard model using updated information on the component scores. In case of missing data at a specific time point the component score of the preceding measurement point were used instead (last carried forward). Hazard ratios were calculated per five units increase of the component scores. Analyses were performed with and without adjustment for age, sex, comorbidity (davies score), ethnicity, and dialysis modality. All analyses were performed using the Statistical Package for Social Sciences (SPSS version 20).

RESULTS

A total of 2,051 incident patients were enrolled in NECOSAD of whom 1,956 were still in the study after 3 months. For 1,379 participants SF-12 and SF-36 component scores were available. Baseline characteristics of patients with and without SF-36/SF-12 scores at baseline are summarised in Table 1. Patients with complete scores were younger ($P < 0.05$), more often white ($P < 0.01$), and had less often cardiovascular disease ($P < 0.01$). At baseline mean PCS36 and PCS12 score were 38.4 (95% CI 37.9 – 38.9) and 37.8 (95% CI 37.3 – 38.3), respectively. Median and interquartile range (IQR) for PCS36 and PCS12 were 37.4 (30.6 – 44.8) and 38.1 (30.9 – 45.2), respectively. Furthermore, mean MCS36 and MCS12 score were 44.8 (95% CI 44.3 – 45.5) and 46.4 (95% CI 45.9 – 47.0), respectively. Finally, median and IQR for MCS36 and MCS12 were 47.6 (38.0 – 55.2) and 46.3 (35.7 – 53.7), respectively.

Cross-sectional analyses

Mean SF-12 and SF-36 scores were relatively stable over time (Figure 1). Both at baseline and during follow-up population mean PCS12 scores were equivalent to the PCS scores derived from the SF-36. For the MCS slightly higher scores were observed using the SF-12 than the SF-36 which was consistent during follow-up. At baseline, correlations between SF-12 and SF-36 scores were very good with an ICC of 0.92 for PCS and 0.94 for MCS. Comparable ICCs were observed at the other measurement points (data not shown). Individual differences between the component scores of the SF-12 and the SF-36 at baseline are shown in the Bland-Altman plots (Figure 2). Limits of agreement ranged from -6.8 to 8.0 for PCS scores and from -8.7 to 5.5 points for MCS scores.

Four hundred and fifteen of 1379 patients had a difference of more than five points on both PCS and MCS scores. These patients received more frequently haemodialysis ($p < 0.01$), had a lower educational level ($p < 0.01$), had a higher comorbidity score ($p < 0.01$), and were older ($p < 0.01$). The differences in patient characteristics remained the same after comparing patients above and below five points difference on both PCS and MCS change in scores.

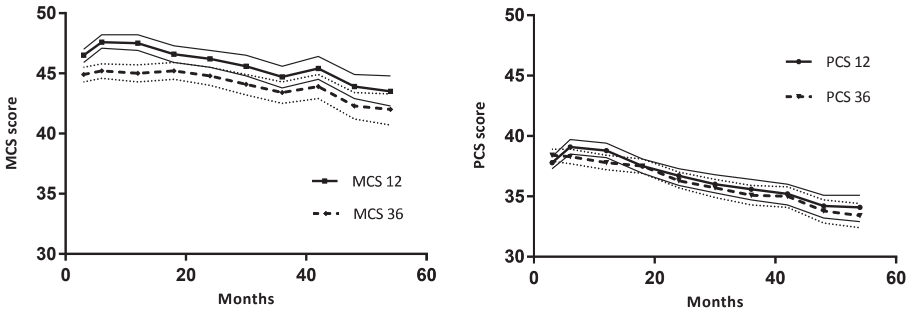

Figure 1. Differences over time of SF-36 and SF-12 scores using linear mixed models.

Table 1. Baseline characteristics of patients whit and without complete SF-36 /SF-12 questionnaires.

Characteristics	Patient with complete SF-scores (N=1379)	Patient without complete SF-scores (N=577)	
Sociodemographic			
Age, y	59.7 (14.9)	61.2 (15.4)	P < 0.05
Sex,% men	61.3	64.6	P = 0.17
Education, % low	55.7	57.1	P = 0.59
Married/living together, % yes	72.4	66.3	P = 0.01
Having children, % yes	78.4	80.5	P = 0.32
Employed, % yes	21.0	18.5	P = 0.26
Ethnicity, % white	93.5	88.0	P < 0.01
Clinical			
Modality, % haemodialysis	62.1	71.4	P < 0.01
BMI, kg/m2	24.8 (4.2)	24.5 (4.1)	P = 0.18
Causes of ESRD, %			P = 0.01
Diabetes Mellitus	14.6	13.9	
Glomerulonephritis	13.9	8.7	
Renal vascular disease	16.2	18.8	
Other	55.4	58.6	
Comorbidity			
Davies comorbidity, %			P = 0.26
No	48.5	50.1	
Intermediate	43.1	39.8	
Severe	8.4	10.2	
Diabetes Mellitus, % yes	22.1	22.7	P = 0.77
Cardiovascular disease, % yes	33.6	40.9	P < 0.01
Laboratory			
Residual GFR, creat ml/min Per 1.73 m2	5.3 (4.5)	4.7 (4.1)	P < 0.01
Albumin, g/l	36.2 (5.2)	35.2 (5.4)	P < 0.01
Haemoglobin, g/l	6.9 (1.0)	6.8 (1.0)	P < 0.01
iPTH, pmol/l	22.9 (30.3)	24.1 (27.1)	P = 0.51

Continuous variables are presented as means (standard deviation) and statistical test include 2-tailed t-tests and X^2 tests.
ESRD: End-Stage Renal Disease, BMI: Body Mass Index, iPTH: intact parathyroid hormone

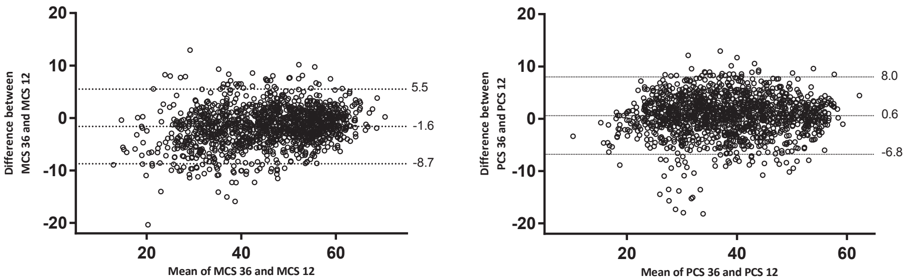

Figure 2. Mean difference and limits of agreement between SF-36 and SF-12 scores.

Change in scores

Changes in SF-12 and SF-36 component scores between the baseline measurement, i.e. at three months after start of dialysis, and at 6 months after start of dialysis could be calculated for 1,117 patients. The mean changes in scores were comparable for SF-12 and SF-36. The mean difference between PCS36 and PCS12 change in scores was 0.2 and between the MCS36 and MCS12 change in scores was -0.1 (Figure 3). Furthermore, the changes in scores were also strongly correlated with an ICC of 0.83 for PCS and 0.90 for MCS respectively. Limits of agreement between the PCS change in scores ranged from -8.8 to 9.2 points and between the MCS change in scores from -8.3 to 8.4 points (Figure 3).

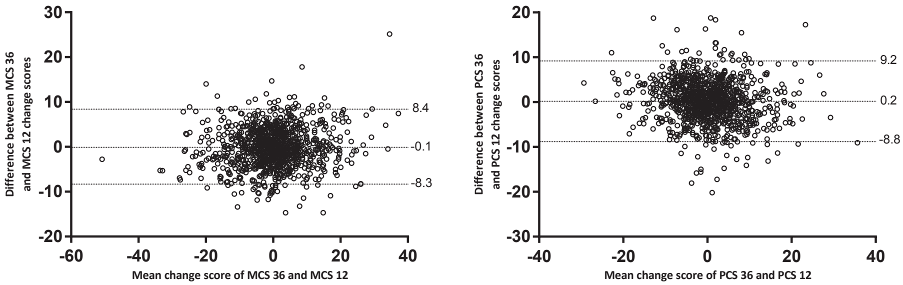

Figure 3. Mean difference and limits of agreement between SF-36 and SF-12 change in scores. Differences are between 3 and 6 months on dialysis.

Mean change in score of EuroQol thermometer between the first two measurement (i.e. at 3 and 6 months) was 0.24 (SD 13.42). Mean change in score of PCS36 and PCS12 was 0.21 (SD 7.41) and -0.04 (SD 8.01) respectively. Mean change in score of MCS36 and MCS12 was 0.01 (SD 9.85) and 0.13 (SD 9.50), respectively. Table 2 shows ICCs for SF-12 and SF-36 change in scores with EuroQol thermometer change in scores for the first three measurements (i.e. at 3, 6 and 12 months). These correlations were moderate, but not different for the SF-12 and SF-36 component scores. Both questionnaires had slightly higher Pearson correlations than ICCs. The analyses with the changes in scores were repeated for the subsequent measurement points. Comparable results were found (data not shown).

Table 2. Intra class correlation coefficients between SF-36/SF-12 change in scores, and EuroQol thermometer change in scores. Differences between 3,6 and 12 months after start of dialysis.

	3-6 months N = 1,026	**95% CI**	**3-12 months N = 844**	**95% CI**	**6-12 months N = 947**	**95% CI**
	EuroQol thermometer		**EuroQol thermometer**		**EuroQol thermometer**	
MCS12	0.22	0.16 – 0.28	0.31	0.25 – 0.37	0.28	0.22 – 0.34
MCS36	0.25	0.20 – 0.31	0.35	0.29 – 0.40	0.27	0.21 – 0.31
PCS12	0.26	0.20 – 0.31	0.31	0.25 – 0.37	0.31	0.25 – 0.36
PCS36	0.27	0.21 – 0.32	0.31	0.25 – 0.37	0.27	0.25 – 0.36

MCS: Mental Component Score, PCS: Physical Component Score, CI: Confidence Interval

Mortality

Median survival time was 4.4 years (95% CI: 4.1 – 4.8) and the median follow-up time was 3.8 years with a minimum of 0 years and a maximum of 9.7 years. Higher MCS and PCS scores were associated with lower short-term mortality in the time-dependent Cox proportional hazard model (Table 3). Associations for PCS scores were somewhat more pronounced than for the MCS scores. The observed associations were not different for the SF-12 and the SF-36 scores. The analysis using baseline scores, reflecting long-term mortality, gave similar results for both questionnaires with adjusted hazard ratios of 0.81 (95% CI: 0.77 - 0.85) for PCS12 and 0.85 (95% CI: 0.81 - 0.90) for MCS12. By adding interaction terms with time differences between SF-36 and SF-12 remained the same (data not shown).

Table 3. Time-dependent hazard ratio's per 5 unit increase. Adjusted for age, sex, comorbidity (davies index), ethnicity and dialysis modality.

unadjusted	**HR**	**95% CI**	**unadjusted**	**HR**	**95% CI**
MCS12	0.84	0.80 – 0.87	PCS12	0.74	0.71 – 0.78
MCS36	0.84	0.81 – 0.87	PCS36	0.71	0.67 – 0.74
Adjusted*	HR	95% CI	Adjusted*	HR	95% CI
MCS12	0.87	0.84 – 0.91	PCS12	0.79	0.76 – 0.83
MCS36	0.87	0.84 – 0.90	PCS36	0.75	0.71 – 0.78

MCS: Mental Component Score, PCS: Physical Component Score, CI: Confidence Interval
*Adjusted for age, sex, comorbidity (davies score), ethnicity, and dialysis modality

DISCUSSION

The component scores derived from the SF-12 showed good agreement with the SF-36 scores and the patterns over time were very similar. Responsiveness was also comparable between the short and long version of the questionnaire. In addition, SF-12 scores, as well as SF-36 scores, were associated with mortality with comparable risk estimates.

Our finding that SF-12 and SF-36 give comparable results is in line with other studies in dialysis patients.[8;9] However, these studies only investigated agreement between SF-12 and SF-36 at one moment in time, and did not compare changes over time. Responsiveness to change of SF-12 has been assessed in various other populations such as patients with stroke, myocardial infarction or a trauma population.[25-29] These studies also showed that SF-12 is comparable to SF-36 with respect to their ability to detect change in HRQOL over time. However, the majority of these studies used effect sizes to determine responsiveness to change. An effect size is calculated as a mean difference over time divided by the standard deviation at baseline or the standard deviation of that difference (standardized response mean). A major problem with this type of measurement is that it reflects the magnitude of the change in scores rather than the validity of the change in scores.[30] Thus, large changes over time result in large effect sizes and small changes in small effect sizes. This however does not give any information on whether the observed change is valid. Therefore, we chose not to use effect sizes but to assess responsiveness of SF-12 and SF-36 by calculating ICCs and create Bland-Altman plots. Our results showed that MCS12 and MCS36 change in scores were comparable. The same holds true for PCS12 and PCS36. However, for the PCS change in scores a somewhat lower ICC was observed. Thus, MCS12 change in scores seem to be a good substitute for MCS36 change in scores. PCS12 change in scores should be interpreted with more caution than MCS12. In addition, we used the EuroQol thermometer change in scores as an external criterion. The ICCs between the SF questionnaires and the EuroQol thermometer were moderate, but the same for SF-36 and SF-12. The moderate ICC is probably due to the difference in type of measurement which could result into a greater variation of the EuroQol thermometer change in scores. The thermometer indicates today's HRQOL with one question and the SF questionnaire consists of 36 or 12 questions which indicates the HRQOL of the last four weeks. However, the ICCs of SF-36 and SF-12 with EuroQol thermometer were comparable which supports the ability of SF-12 to reproduce SF-36 scores.

The Bland-Altman plots showed a small average difference between SF-12 and SF-36 scores. This indicates that SF-12 is comparable to SF-36 when an average of multiple SF-12 scores is calculated. However, the 95% limits of agreement showed that the difference between individual SF-12 and SF-36 scores can vary up to 8 points. In other words, the SF-12 score of an individual patient, represented by 1 dot on the Bland-Altman plot, can vary up to 8 points in comparison to the SF-36 score. A clinical meaningful change is often set to 5 points.[31] Therefore, it is important to realize that by using SF-12 clinical meaningful changes in HRQOL might be missed in individual patients. Consequently, it might be preferable to use SF-36 instead of SF-12 to detect changes in HRQOL over time in individual patients. To find an explanation for these individual differences we compared two groups of patients. One group with large differences between SF-12/ SF-36 and the other group with small differences between SF-12/SF-36. Results showed that more vulnerable patients (i.e. higher co-morbidity and older) were less consistent in their answers. Future research could address the differences between subgroups of patients. Furthermore, the eight specific domains calculated by SF-36 were lost when SF-12 was calculated. Thus, when specific information about HRQOL is needed SF-12 is not useful and SF-36 should be used instead.

A lower HRQOL is related to a higher mortality rate as shown in studies with dialysis patients.[10-15] Results found in our study showed that MCS12/MCS36 and PCS12/PCS36 were equally associated with both short-term and long-term mortality. This is in agreement with the paper by Lacson et al.[8] The paper written by Osthus et al.[9] showed less consistent results but this is probably due to a small sample size.

A few limitations should be noted for this study. First, SF-12 and SF-36 were not administered separately, but SF-12 was extracted from SF-36 items. Therefore, it is possible that the ability of SF-12 to reproduce PCS and MCS was somewhat overestimated. Second, SF-36 and SF12 scores can differ between countries. Generalizability of study findings could therefore be difficult, but associations between SF-36 and SF-12 scores will probably be the same. Finally, not all patients had complete SF-36 and SF-12 scores, and as expected baseline characteristics differed between patients with and without complete questionnaires. It is not to be expected that this changes the ability of SF-12 to reproduce SF-36 scores.

In conclusion, SF-12 scores can be used to replicate SF-36 scores in cross-sectional studies and can also be used to detect HRQOL changes over time in cohort studies on dialysis patients. SF-12 and SF-36 were similarly associated with short-term and long-term mortality. However, for individual patients considerable differences were observed between SF-12 and SF-36. We therefore recommend to use the SF-36 for individual purposes.

ACKNOWLEDGEMENTS

We thank the trial nurses, participating dialysis centres and data managers of the NECOSAD study for collection and management of the data. We gratefully thank all patients who participated in the NECOSAD study. This study was supported by grants from the Dutch Kidney Foundation (E.018) and the Dutch National Health Insurance Board (OG97/005). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

REFERENCES

(1) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-483.

(2) Merkus MP, Jager KJ, Dekker FW, Boeschoten EW, Stevens P, Krediet RT. Quality of life in patients on chronic dialysis: self-assessment 3 months after the start of treatment. The Necosad Study Group. Am J Kidney Dis 1997;29:584-592.

(3) Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220-233.

(4) Gandek B, Ware JE, Aaronson NK et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998;51:1171-1178.

(5) Muller-Nordhorn J, Roll S, Willich SN. Comparison of the short form (SF)-12 health status instrument with the SF-36 in patients with coronary heart disease. Heart 2004;90:523-527.

(6) Pickard AS, Johnson JA, Penn A, Lau F, Noseworthy T. Replicability of SF-36 summary scores by the SF-12 in stroke patients. Stroke 1999;30:1213-1217.

(7) Pakpour AH, Nourozi S, Molsted S, Harrison AP, Nourozi K, Fridlund B. Validity and reliability of short form-12 questionnaire in Iranian haemodialysis patients. Iran J Kidney Dis 2011;5:175-181.

(8) Lacson E Jr, Xu J, Lin SF, Dean SG, Lazarus JM, Hakim RM. A comparison of SF-36 and SF-12 composite scores and subsequent hospitalization and mortality risks in long-term dialysis patients. Clin J Am Soc Nephrol 2010;5:252-260.

(9) Osthus TB, Preljevic VT, Sandvik L et al. Mortality and health-related quality of life in prevalent dialysis patients: Comparison between 12-items and 36-items short-form health survey. Health Qual Life Outcomes 2012;10:46.

(10) Mapes DL, Bragg-Gresham JL, Bommer J et al. Health-related quality of life in the Dialysis Outcomes and Practice Patterns Study (DOPPS). Am J Kidney Dis 2004;44:54-60.

(11) Wu AW, Fink NE, Marsh-Manzi JV et al. Changes in quality of life during haemodialysis and peritoneal dialysis treatment: generic and disease specific measures. J Am Soc Nephrol 2004;15:743-753.

(12) Kalantar-Zadeh K, Kopple JD, Block G, Humphreys MH. Association among SF36 quality of life measures and nutrition, hospitalization, and mortality in haemodialysis. J Am Soc Nephrol 2001;12:2797-2806.

(13) Lopes AA, Bragg-Gresham JL, Satayathum S et al. Health-related quality of life and associated outcomes among haemodialysis patients of different ethnicities in the United States: the Dialysis Outcomes and Practice Patterns Study (DOPPS). Am J Kidney Dis 2003;41:605-615.

(14) Valdes C, Garcia-Mendoza M, Rebollo P, Ortega T, Ortega F. Mental health at the third month of haemodialysis as a predictor of short-term survival. Nephrol Dial Transplant 2006;21:3223-3230.

(15) Lopez RK, Garcia Lopez FJ, de Alvaro MF, Alonso J. Perceived mental health at the start of dialysis as a predictor of morbidity and mortality in patients with end-stage renal disease (CALVIDIA Study). Nephrol Dial Transplant 2004;19:2347-2353.

(16) Davies SJ, Russell L, Bryan J, Phillips L, Russell GI. Comorbidity, urea kinetics, and appetite in continuous ambulatory peritoneal dialysis patients: their interrelationship and prediction of survival. Am J Kidney Dis 1995;26:353-361.

(17) Aaronson NK, Muller M, Cohen PD et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. J Clin Epidemiol 1998;51:1055-1068.

(18) Farivar SS, Cunningham WE, Hays RD. Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. Health Qual Life Outcomes 2007;5:54.

(19) Liem YS, Bosch JL, Hunink MG. Preference-based quality of life of patients on renal replacement therapy: a systematic review and meta-analysis. Value Health 2008;11:733-741.

(20) Wyld M, Morton RL, Hayen A, Howard K, Webster AC. A systematic review and meta-analysis of utility-based quality of life in chronic kidney disease treatments. PLoS Med 2012;9:e1001307.

(21) Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53-72.

(22) Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095-1108.

(23) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-310.

(24) Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials 1996;17:343-346.

(25) Jenkinson C, Layte R, Jenkinson D et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? J Public Health Med 1997;19:179-186.

(26) Kiely JM, Brasel KJ, Guse CE, Weigelt JA. Correlation of SF-12 and SF-36 in a trauma population. J Surg Res 2006;132:214-218.

(27) Muller-Nordhorn J, Nolte CH, Rossnagel K et al. The use of the 12-item short-form health status instrument in a longitudinal study of patients with stroke and transient ischaemic attack. Neuroepidemiology 2005;24:196-202.

(28) Rubenach S, Shadbolt B, McCallum J, Nakamura T. Assessing health-related quality of life following myocardial infarction: is the SF-12 useful? J Clin Epidemiol 2002;55:306-309.

(29) Riddle DL, Lee KT, Stratford PW. Use of SF-36 and SF-12 health status measures: a quantitative comparison for groups versus individual patients. Med Care 2001;39:867-878.

(30) Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349-362.

(31) Ware JESK. SF-36 health survey : manual and interpretation guide. Boston: Health Institute, New England Medical Center, 1993.