PATIENT-REPORTED OUTCOME MEASUREMENT GROUP, OXFORD

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR ADULTS WITH CHRONIC KIDNEY DISEASE

Report to the Department of Health and NHS Kidney Care, 2010

patient-reported

O U T C O M E

measures

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR PEOPLE WITH CHRONIC KIDNEY DISEASE, 2010

Elizabeth Gibbons
Ray Fitzpatrick

Patient-reported Outcome Measurement Group
Department of Public Health
University of Oxford

Copies of this report can be obtained from:

Elizabeth Gibbons
Senior Research Officer
Patient-reported Outcome Measurement Group

tel: +44 (0)1865 289402
e-mail: elizabeth.gibbons@dphpc.ox.ac.uk

Alternatively, it can be downloaded free of charge from the PROM Group website:

http://phi.uhce.ox.ac.uk/

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR PEOPLE WITH CHRONIC KIDNEY DISEASE: 2010

CONTENTS

EXECUTIVE SUMMARY

Aims of the report

The aims of this report are to review the evidence of Patient-Reported Outcome Measures (PROMs) for adults with Chronic Kidney Disease (CKD) to provide a basis for consideration of the systematic use of PROMs in routine kidney care in adults in England.

The methods of the review are described and the results of the search including sources and search terms used to identify specific published research. Details of this evidence are presented for generic, preference based and then condition-specific Patient-Reported Outcome Measures. The report concludes with discussion and recommendations.

Results

GENERIC MEASURES

Five generic measures were included in the review:

a) SF-36
b) SF-20
c) SF-12
d) QWB-SA
e) SIP

HEALTH UTILITY MEASURES

Three preference-based measures were included in the review:

1. EQ-5D
2. SF-6D
3. HUI

RENAL-SPECIFIC MULTIDIMENSIONAL QUESTIONNAIRES

Eight renal-specific measures were included in the review:

a) Quality of Life Index-Dialysis (QLI-D)
b) Kidney Disease Quality of Life-Long Form (KDQOL-LF)
c) Kidney Disease Quality of Life-Short Form (KDQOL-SF)
d) Kidney Disease Questionnaire (KDQ)
e) Kidney Transplant Questionnaire (KTQ)
f) Renal Quality of Life Profile (RQLP)
g) CHOICE Health Experience Questionnaire (CHEQ)
h) Renal Dependant Individualised Quality of Life Questionnaire

RENAL-SPECIFIC SYMPTOM FOCUSED QUESTIONNAIRES.
Patients with ESRD have high symptom burden (Davison et al., 2006a) and the following section reports the evidence of questionnaires to measure kidney disease specific symptoms.

Thirteen symptom focused questionnaires are included in the review:

1. Modified Edmonton Symptom Assessment System (ESAS)
2. Memorial Symptom Assessment Scale-short form (MSAS-SF
3. Dialysis Symptom Index
4. National Kidney Dialysis and Kidney Transplant Study (NKDKTS) symptom checklist
5. Transplant Symptom Occurrence and Distress Scale
6. Patient Outcome Scale--symptom module (POSs)
7. The Rome II
8. Gastrointestinal Rating Scale (GSRS)
9. Gastrointestinal Quality of Life Index (GIQLI)
10. Thirst Distress Scale
11. Multi-dimensional Fatigue Inventory (MFI-20)
12. Fatigue Severity Scale
13. Epworth Sleepiness Scale

Recommendations

There is supporting evidence for use of the SF-36 in the patient with Chronic Kidney Disease. Modest evidence exists for its responsiveness so that further evidence is needed before widespread use to evaluate outcomes and quality of services. The EQ-5D is favoured among the preference-based measures with more supporting evidence, but similar caveats must be expressed about lack of evidence of responsiveness and use in the context of quality and outcomes.

Of the renal-specific multidimensional measures the KDQOL has more supporting evidence. It is not apparent that there is significant value of the longer over the shorter version and the likelihood of lower respondent burden and higher responses rates make the shorter version potentially more appropriate. However although shorter, an 80-item instrument may still be considered long for routine and large scale administration. The amount of testing of the use of different versions of KDQOL in the NHS is still modest with relatively little UK based evidence, so further assessment before it is possible to be confident in recommending the short version of KDQOL.

Given the overlap between instruments such as SF-36 and KDQOL the issue of how instruments are used in combination needs to be flagged up. There is no merit, for example, in using both SF-36 and KDQOL in the same survey whereas the combination of EQ-5D and KDQOL would provide complementary and non-overlapping evidence of patients' perceptions in relation to kidney disease.

Whilst the concept of utilising a short symptom based questionnaire is attractive, the main benefit would only be for screening or identification of the prevalence of symptoms. No symptom focused instrument can be recommended at this stage.

Much of the evidence for this review has been found in evidence of cross-sectional surveys and lack of longitudinal evidence is limitation when considering PROMs for use to assess quality and outcomes. Similarly PROMs are like any outcome measure in that appropriate adjustment of potential confounders is essential to facilitate interpretation of scores from PROMs.

INTRODUCTION

Patient-reported outcome measures (PROMs) offer enormous potential to improve the quality and results of health services. They provide validated evidence of health from the point of view of the user or patient. They may be used to assess levels of health and need in populations, and in users of services they can provide evidence of the outcomes of services for the purposes of audit, quality assurance and comparative performance evaluation. They may also improve the quality of interactions between health professionals and individual service users.

Lord Darzi's Interim Report on the future of the NHS recommends that patient-reported outcome measures (PROMs) should have a greater role in the NHS (Darzi 2007). The new Standard NHS Contract for Acute Services, introduced in April 2008, includes a requirement to report from April 2009 on patient-reported outcome measures (PROMs) for patients undergoing Primary Unilateral Hip or Knee replacements, Groin Hernia surgery or Varicose Vein procedures. Furthermore, Lord Darzi's report 'High Quality Care for All' (2008) outlines policy regarding payments to hospitals based on quality measures as well as volume. These measures include PROMs as a reflection of patients' experiences and views. Guidance has now been issued regarding the routine collection of PROMs for the selected elective procedures (DH, 2008).

There are three broad categories of PROMs: generic health status-, preference-based-, and condition or population-specific-measures. Generic instruments comprise items intended to be relevant to the widest range of patients' conditions and the general population. Preference-based measures are also broad in content but additionally provide utilities or values regarding health (for use in, for example, cost-utility analyses of interventions). Condition-specific instruments are often more focused on a particular disease or health condition (for example, diabetes), a patient population (for example, older people), a specific problem or symptom (for example, pain), or a described function (for example, activities of daily living). For any given area of health, condition-specific instruments may have greater clinical appeal due to the inclusion of content specific to particular conditions, and the likelihood of increased responsiveness to interventions.

It has been recommended that a combination of a generic or utility measure with a specific measure be used in the assessment of patient health outcomes on the grounds that the complementary content of the two types of measure when combined should assess a full range of aspects of health relevant to the particular population concerned. However, consensus is often lacking as to which instrument to use for specific purposes and contexts (Garratt et al., 2002). Structured reviews of PROMs for specific health conditions or populations can provide guidance for selection. An evidenced-based approach strengthens recommendations from these reviews.

Selection criteria have been defined for assessing the quality of existing PROMs (Streiner and Norman, 1995; Fitzpatrick et al., 1998). These include measurement issues, such as reliability, validity, responsiveness, and precision, as well as practical issues, such as acceptability and feasibility. These criteria are briefly summarised in Appendix A.

Chronic Kidney Disease (CKD)

The Department of Health's Vascular Programme addressed shared issues of disease prevention, risk factors and management of specific related conditions: renal disease, heart disease, diabetes and stroke. These priorities are echoed in the Renal NSF (DoH 2005). In addition, the 18 week pathway specified in the NHS Improvement Plan (2004) aims to improve the patient experience. The Commissioning pathway for patients with chronic kidney disease (2008) specifies processes of care for patients from primary care to rehabilitation. Additionally there is a requirement to measure the quality of life of patients.

The definition and classification of Acute Kidney Injury is principally guided by the RIFLE criteria which specifically refer to serum creatinine concentration and volume of urine output. The RIFLE acronym refers to the increasing severity: Risk, Injury and Failure and the outcomes: Loss and End Stage Kidney Disease (Kellum et al., 2008). Persistent AKI is classified as complete loss of renal function greater than 4 weeks in duration. The term AKI includes patients across the full spectrum of acutely unwell patients and not just those receiving RRT.

Chronic Kidney Disease (CKD) has been defined as the presence of kidney damage indicated by a Glomerular Filtration Rate (GFR) less than 60ml/min/1.73m2 for three months or longer (Levey et al., 2005). Five different stages of classification of CKD are proposed based on different levels of GFR by the Kidney Disease Outcomes Quality Initiative (KDOQI, 2002). Winearls and Glassock (2009) offer a modified definition of the KDOQI stages based on GFR as it is considered that the rigidity of the cut-off for diagnosis based on GFR does not account for differences by, for example, age and gender. Advanced kidney disease are those patients with late stage 4 and stage 5 CKD (DH 2009).

Table 1 outlines the definition and staging of kidney disease outlined in the Quality Outcome Framework.

Table 1

Stage	GFR*	Description
1	90+	Normal kidney function but urine findings or structural abnormalities or genetic trait point to kidney disease
2	60-89	Mildly reduced kidney function, and other findings (as for stage 1) point to kidney disease
3	30-59	Moderately reduced kidney function
4	15-29	Severely reduced kidney function
5	<15	Very severe, or established kidney failure

*Glomerular filtration rate. (Source-Quality and Outcomes Framework guidance for GMS contract 2008/9). *Shaded boxes: by definition-Chronic Kidney Disease*

Levey et al. (2009) outline a conceptual model of CKD reflecting the continuum of the disease progression including primary, secondary and tertiary prevention. Primary prevention focuses of identification of people at risk of developing CKD based on the

presences of modifiable risk factors. Secondary prevention is aimed at improving outcomes, management of risk factors and prevention of End Stage Renal Disease (ESRD). Tertiary prevention focuses on improving outcomes and management of complications. End Stage Renal Disease (ESRD) is irreversible and fatal without Renal Replacement Therapy (RRT) or transplantation.

Renal Replacement Therapy (RRT) has a fundamental role in the treatment of CKD and advanced kidney disease. RRT in the form of dialysis can be administered either peritoneal or haemodialysis. This can be carried out in various settings: home, in-hospital units or satellite units.

RRT is often part of the treatment for people with advanced kidney disease. However, transplantation may be the focus for increasing survival and increasing HRQL (Muehrer and Becker. 2005). There are though specific factors which may impact on HRQL post-transplant such as side effects from medications particularly immunosuppressive therapy, presence of co morbid conditions and potential rejection. Rejection can of course result in failure of the transplanted organ and returns the patient to being dialysis-dependant.

Mortality is high in patients with advanced kidney disease often due to the presence of other morbidities. End of life care has been a focus of the NSF for Renal Services (DH 2005) with a quality requirement for End of Life Care. This is echoed in the End of Life Care Strategy (DH 2008) which aims to ensure all people receive high quality care during this phase of their life. Subsequent development of the End of Life Care in Advanced Kidney Disease: A Framework for Implementation (NHS Kidney Care 2009) sets out how to achieve high quality end of life care for patients with kidney disease. In addition, about 25% of people who reach ESRF now elect for conservative kidney care and these people are likely to have other chronic conditions such as heart failure. They may often only live for 2 to 3 years.

The Liverpool Care Pathway for the Dying Patient (LCP) has been adapted for patients with CKD. This pathway specifically provides drug guidance for the control and management of symptoms for patients with advanced kidney disease in the last days of life. It is known that opioid based analgesics accumulate in renal failure which can cause toxicity. The drug guidance in the LCP considers this in its protocol.

Patients with advanced kidney disease regardless of treatment modality have a high symptom burden (Davison et al., 2006a; Murphy et al., 2009) with the most common symptoms being tiredness, sleep disturbance, poor appetite, pruritis, and pain and decreased well-being. Other symptoms referred to are, cramps, abdominal pain, palpitations, oedema (Murtagh et al., 2007). Symptoms are defined as the intensity (severity); timing (duration and frequency of occurrence); distress (degree of discomfort or bother) (Jablonski 2007). There is also considerable impact on the psychosocial aspects of people's lives with increased demand on carers and family members of people with kidney disease.

It has been reported that there is a lack of significant association between clinical measures and HRQL and symptoms burden. Some evidence is reported of an association between anaemia and patient reported health and that treatment with the

administration of erythropoietin associated with improvements in physical and psychological well-being.

Symptom management is often the focus of palliative care for patients with advanced kidney disease including patients receiving RRT and those patients for which dialysis is not the optimal choice of treatment- for example who have multiple comorbid conditions and poor functional status (Murtagh et al., 2007).

It is clear that there is an illness trajectory for patients with kidney disease which is classified both physiologically (GFR) and by subsequent management strategies; Conservative Management, RRT, or Transplantation. All of these can lead to an End of Life pathway where the focus is on more supportive and palliative care and ensuing symptom control. Health-related Quality of Life (HRQL) can be the major focus of the outcomes of care for these patients as mortality is so high. PROMs therefore become a valuable method of eliciting patient experience during the illness trajectory.

The identification of appropriate PROMs is the remit of this present review commissioned by both by NHS Kidney Care and the Department of Health.

Structure of the report

The methods of the review are described, including search strategies and search terms used to identify specific published research regarding PROMs for kidney disease. The report focuses on evidence and recommendations for PROMs. Evidence is presented for generic, preference based and renal-specific PROMs (including Symptom based instruments) for each illness phase where available: Conservative management; patients receiving RRT and Post transplant patients.

Methods

Methods adopted were as described in previous reviews performed by the PROM group, Oxford. Comprehensive searches were conducted; articles retrieved were assessed for relevance and evidence of measurement performance and operational characteristics abstracted for each PROM identified.

a) Search sources and terms
Several sources were searched to identify relevant articles:

The primary source of evidence was the bibliographic database compiled by the PROM group in 2002 with funding from the Department of Health and hosted by the University of Oxford [1]. In 2005, it became the property of the NHS Information Centre for Health & Social Care. The PROM database comprises 30, 350 records up to the end of 2006 (16,054 online up to December 2005) downloaded from several electronic databases using a comprehensive search strategy (available on request). These records had been assessed as eligible for inclusion in the bibliography and assigned keywords. These records were searched using the keyword ‘renal’ and by instrument identified during the review.

[1] Available online at http://phi.uhce.ox.ac.uk

Supplementary searches included scanning the reference lists of key articles, checking instrument websites, where found, and drawing on other bibliographic resources. Hand searching of titles of key journal from 2007 to 2009 was conducted. The following journals were selected:

- Journal American Society of Nephrology
- Kidney International
- American Journal of Kidney Disease
- Nephrology Dialysis Transplantation
- Health and Quality of Life Outcomes
- Medical Care
- Quality of Life Research

The National Institute for Health Research: Health Technology Assessment Programme, published research was also searched.

In addition, PubMed records for the past two years (2007-2009) were searched using a comprehensive search filter devised by Terwee et al., (2009) (*in press, but permission obtained from Author)* and further developed by the PROM group and Outreach Librarian at the University of Oxford.

b) Inclusion criteria

Published articles were included if they provided evidence of measurement and/or practical properties of relevant PROMs (Fitzpatrick et al., 1998).

-Population

- Patients with Acute Kidney Injury.
- Patients with advanced kidney disease or ESRD receiving conservative management, RRT, or Transplantation
- Patients with ESRD not receiving RRT but palliative care
- Patients post renal transplant

- English speaking populations

-Study design selection

- Studies where a principal PROM is being evaluated
- Studies evaluating several PROMs concurrently
- Trials or studies evaluating the effectiveness of interventions; where a PROM is used as an endpoint.
- Prospective studies measuring patient-reported outcomes where data is available for a PROM in terms of measurement performance or operational characteristics

-Specific inclusion criteria for generic, preference- and disease-specific instruments

- The instrument is patient-reported
- There is published evidence of measurement reliability, validity or responsiveness following completion in the specified patient population

- Evidence is available from English language publications, and instrument evaluations conducted in populations within UK, North America, or Australasia.
- The PROM ideally will be multi-dimensional. It is at the reviewer's discretion to include instruments which are specific to a health condition but have a narrow focus, for example a specific dimension of health, such as, symptoms.

-Exclusion criteria

- Clinician-assessed instruments
- Studies evaluating the performance of non-patient reported measures of functioning or health status where a PROM is used as a comparator indicator
- Studies with very small samples, i.e fewer than 50 participants

c) Data extraction

For all PROMs included in the review, evidence is reported for the following measurement criteria:

- reliability
- validity
- responsiveness
- precision

Operational characteristics such as patient acceptability and feasibility of administration for staff are also reported.

d) Assessment of methodological quality of PROMs

Assessment and evaluation of the PROMs was performed by means of the criteria described in Appendix A.

Searches identified nearly 6000 potentially relevant records. When assessed against the review inclusion criteria, 78 articles were included in the review (Table 2).

Table 2. Number of articles identified by the literature review

Source	***Results of search***	***Number of articles included in review***
PROM database: 30,350	460	40
Pubmed	5472	22
Hand searching	-	16
TOTAL	-	**78**

GENERIC MEASURES

Five generic measures are included in the review:

1. SF-36
2. SF-20
3. SF-12
4. QWB-SA
5. Sickness Impact Profile

1. SF-36

The SF-36 is a generic health status instrument with 36 items assessing health across eight domains (Ware, 1997), namely bodily pain (BP: two items), general health perceptions (GH: five items), mental health (MH: five items), physical functioning (PF: ten items), role limitations due to emotional health problems (RE : three items), role limitations due to physical health problems (RP: four items), social functioning (SF: two items), and vitality (VT: four items). An additional health transition item, not included in the final score, assesses change in health. All items use categorical response options (range: 2-6 options). Scoring uses a weighted scoring algorithm and a computer-based programme is recommended. Eight domain scores give a health profile; scores are transformed into a scale from 0 to 100, where 100 denotes the best health. Scores can be calculated when up to half of the items are omitted. Two component summary scores for physical and mental health (MPS and MCS, respectively) can also be calculated.

Twenty-four studies were included providing good evidence of performance. Six of these studies included the SF-36 as part of a kidney-specific instrument, the Kidney Disease Quality of Life questionnaire. One study evaluated the SF-36 in a post renal transplant population and the others were with patients with ESRD receiving RRT. Eight studies were conducted in the UK.

Patients receiving RRT

High internal consistency has been reported for all domains with alpha's greater than 0.80 (Hays et al., 1994; Ozminkowski et al., 1997; Wight et al., 1998; Wu et al., 2004; O'Sullivan and McCarthy 2007).

The internal structure is supported for seven of the SF-36 domains with exception of the SF domain (Ozminkowski et al.,1997). The factor structure is supported in Lowrie et al., (2003).

Discriminative validity is supported. Most domains of the SF-36 discriminated between patients receiving haemodialysis and CAPD. Scores were significantly lower for Physical Functioning and General Health Perceptions. The Role Emotional and Mental Health domains were nearer to population norms. These scores were lower in

the CAPD patient group than those receiving haemodialysis (Khan 1995, UK). Further support is reported with scores discriminating between patients with ESRD receiving haemodialysis (n=1000) which were lower than population norms (mean score 35 vs. 50 US reference) (Cleary and Drennan 2005; DeOreo 1997). Further evidence is given in Johansen et al. (2001) with haemodialysis patients PCS and PF scores statistically significantly lower than population norms but MCS scores near to population norms (Lowrie et al., 2003; Knight et al., 2003, Wu et al., 2004).

Statistically significantly lower scores were reported for patients with ESRD receiving CAPD than population norms. This study also evaluated the QOL of carers of these patients and although their scores were higher than for patients they were still significantly lower than population norms (Lin-sun Fan et al., 2008). Following commencement of a home PD programme, there was a statistically significant improvement in the SF domain but not for others.

Scores have been reported lower for patients with more severe kidney disease as hypothesised (Ozminkowski et al.,1997). Similar discriminative properties have been reported comparing scores from ESRD patients and other chronic conditions (Asthma/COPD, Diabetes) (Hays et al 1994) and between the presence of a disability and co-morbidities (Manns et al., 2003). A significant decline in SF-36 scores is reported in older patients (aged 70 or older) receiving HD (Unruh et al., 2008).

Both Physical and Mental component scores discriminated patients with haemoglobin levels of ≥ or ≤ 11 Hb g/dl with significantly lower scores for patients classified as anaemic (≤ 11 Hb g/dl) as was expected (Plantinga et al.. 2007). Furthermore, there was an incremental benefit in quality of life (PCS and MCS adjusted scores) for patients who had an increase of 1g/dl Hb over a six month period.

Patients receiving HD in renal satellite units and those in main renal units reported similar scores which were comparable to other studies reported here with MCS scores near population norms and PCS significantly lower (Roderick et al., 2005 UK). Lower MCS scores have been reported for patients receiving haemodialysis compared to CAPD patients in this study. Higher scores for Vitality are reported for haemodialysis patients which is consistent with hypotheses (Ozminkowski et al.,1997; Wu et al., 2004). It is recognised that patients with ESRD receiving RRT have reduced energy levels and therefore do not participate in physical activity. Compounding factors include the presence of anaemia and other co morbidities (Brenner and Brohart 2008). Higher BP was reported for patients receiving PD compared to HD both at baseline and one year follow-up (Wu et al., 2004). Lower scores for Physical Functioning and Vitality in patients receiving haemodialysis compared to population norms reported in O'Sullivan and McCarthy (2007). Significant association is reported between general fatigue and physical functioning with decreased levels of fatigue related to increasing physical function as would be expected (Mittal et al., 2001).

Some evidence of predictive validity is reported. In a large study (14,815) age-adjusted multivariate analysis indicated both PCS and MCS scores less than 30 were independently associated with 1-year mortality (hazard ratio MCS: 1.48 [1.32-1.64]; PSC: 1.62 [1.36-1.92]; Knight et al 2003). Additional analysis reported that a 10-point

decline in MCS and in PCS scores was associated with a significant additional increase in mortality (Knight et al., 2003). Survival analysis indicated that an increase of PCS core by 5 points was associated with a 10% increase in the probability of survival and the risk of death twice as probable when the PCS score was lower than the median value (35). Lower scores on the PCS were also predictive of hospitalisation but not for the MCS (DeOreo 1997). When the PCS score was lower than the median value of 36, the odds ratio for hospitalisation was 1.92 (0.91 to 4.01) Mittal et al., 2001). In a larger population study of over 13,000 dialysis patients, The PCS and MCS scores were significantly associated with hospitalisation and mortality rate (Lowrie et al., 2003).

Patients who skipped a haemodialysis session were more likely to have higher PCS scores than those who did not but lower MCS scores (DeOreo 1997).

Convergent validity is reported with moderate correlation of scores between the PCS with physical activity measured by accelerometry and the Human Activity Profile (Johansen et al., 2001).

A cross-sectional study of patients with end stage renal failure at one unit found that patients who had received transplants had significantly more favourable SF-36 scores after controlling for a wide range of potential con-founders (Wight et al., 1998).

Some evidence for the responsiveness of SF-36 is found in a trial of human erythropoietin therapy to address symptoms of anaemia in dialysis patients (Beusterien et al., (1996). Assessed at inception of therapy and on average 99 days later, patients in receipt of therapy experienced significantly improved scores for vitality, physical function, social function and mental health dimensions.

Additional evidence of responsiveness of SF-36 for this patient group is found in a longitudinal study of haemodialysis and peritoneal dialysis patients assessed at baseline and one-year follow-up (Wu et al., 2004).. Adjusting for a range of potential confounders statistically significant improvement on GH and PF domains of SF-36 for haemodialysis patients (Wu et al., 2004).

Ceiling effects have been reported for RP, BP, SF and RE and floor effects for RF and RE. The greatest floor effect was reported for RP (50% of responses at lowest possible score) and ceiling effect for RE (47% of responses at highest possible score) (Wight et al., 1998). Further floor effects are reported for RP and RE (Klemens et al., 1994; Hays et al. 1994) and ceiling effects for SF and RE (Klemens et al., 1994; Hays et al. 1994).

A 45% response rate was reported to a postal administration to people with ESRD and receiving home peritoneal dialysis (Lin & Curhan 2008). However higher response rates have been reported with 80% of patients returning questionnaires completed at home. These patients were given the questionnaire whilst receiving RRT in a hospital setting. In this study, it was considered that patients should not complete the SF-36 whilst receiving RRT as this can be a distressing experience which may bias the responses (Wight et al., 1998 UK). Some patients in the study by Wight et al (1998) stated that the SF-36 took longer than 15 minutes to complete. 82 % response rate was

reported in the study by Roderick et al., (2005 UK) but this was interview administered.

Data quality was pre-defined in a study by Klemens et al., (1994) as 'excellent' for all 36 items computable and 'satisfactory' if at least half of the items have been completed. 62% of data was 'excellent' and 30% 'satisfactory'.

Transplant
Evidence of responsiveness of SF-36 was found in a study of post transplant patients receiving different immunosuppressant regimes (Russ et al., 2007). Significant changes over time and differences between treatment arms were found for vitality, general health and social function dimensions of SF-36, with the authors speculating that improved energy made it possible for patients to maintain more socially active lives.

SF-36 as part of KDQOL

The evidence of the SF-36 administered as part of the KDQOL is reported separately here. Details of the complete instrument are reported in the next section.

Patients receiving RRT
The PF scale in a small study had acceptable reproducibility (Johansen et al., 2001). High internal consistency reliability of the SF-36 domains is reported for the self-administered KDQOL-LF for all domains except GH, VT and SF (Unruh et al., 2003). The interview administered method yielded similar results. Item discriminant validity is supported for both methods of administration. There were significantly higher scores on RP and RE in interview administration in Unruh et al. (2003)

The SF-36 domains included in the KDQOL-SF discriminated between patients receiving RRT and UK population norms (Carmichael et al., 2000). The PSC and MCS discriminated between patients with co morbidities, and expected demographic differences were observed in a large international study including UK (Lopes et al., 2007) and also in a study by Lee et al., (2005 UK). Further discriminative properties are reported between patients with CKD and population norms with lower scores for PCS and near population norms for MCS which are consistent with other studies (Johanson et al., 2007). Lower scores were reported for patients receiving HD or PD than those that had received a renal transplant (Lee et al., 2005 UK).

Floor effects for RP and RE are reported for interview administered and self-administered questionnaires. Ceiling effects are reported for RP, SF and RE for self-administration and interview administered. Generally ceiling effects were greater for the interview administered method (Unruh et al., 2003).

A 66% response rate is reported Johanson et al., (2007). The questionnaires were self-completion in the dialysis unit.

Transplant
Discriminative properties are reported for General Health and Vitality domains with patients post renal transplant having higher scores than those patients on a waiting list (Sureshkumar et al., 2005).

2. SF-20

The Medical Outcomes Study (MOS) 20-item Short Form Health Survey (SF-20) is a 20-item abbreviation of the same Rand instrument from which the SF-36 originates. The abridged instrument was intended to reduce respondent burden, whilst comprehensively addressing important issues in health status measurement.

The SF-20 assesses health across six domains, namely bodily pain (BP: one item), general health perception (GH: five items), physical function (PF: six items), mental health (MH: five items), social function (SF: one item), and role function (RF: two items) Items have categorical response options (range: 3-6 options); several items have reversed scoring. Domain item summation scores are transformed into a scale from 0 to 100, where higher values denote better health.

Patients receiving RRT
Two studies report some evidence of construct validity and patient acceptability of SF-20.

In Meers et al. (1992) the SF-20 discriminated between patients with ESRD receiving dialysis and patients who had received a transplant with significantly higher scores across all domains for the transplant patients. The SF-20 discriminated between patients receiving CAPD and those receiving different modalities; haemoglobin level and co-morbidities (Morton et al., 1996). The SF-20 was administered in a peritoneal dialysis outpatient clinic and staff did not find it intrusive to the running of the clinic. Minimal burden to patients was reported.

3. SF-12

A shorter 12 item version of SF-36 was developed using regression analysis; 12 items were selected that reproduced 90% of the variance in the overall Physical and Mental Health components of the SF-36. A computer-based scoring algorithm is used to calculate scores; Physical Component Summary (PCS) and Mental (MCS) Component Summary scales are generated using norm-based methods. Scores are transformed to have a mean value of 50, standard deviation (SD) 10, where scores above or below 50 are above or below average physical or mental well-being, respectively.

Patients receiving RRT
Limited evidence is reported for the SF-12 in relation to kidney disease. Curtin et al. (2004) obtained a response rate 85%.

Some weak but significant correlation of scores is reported between the PCS and MCS and knowledge indices in Curtin et al. (2004). A high response rate is reported in this study.

Conservative management
Anaemia as a result of CKD is common and associated with iron deficiency. Often patients need oral iron supplementation or intravenous replacement. The SF-12 and KDCS of the KDQOL were used in a trial of oral Vs. intravenous iron supplementation. Small but significant differences in scores in the same direction were observed for the intravenous group of patients. This was associated with significant improvement in haemoglobin levels in these patients (Agarwal et al., 2006).

4. Quality of Well-Being Scale-Self-Administered (QWB-SA)

A self-administered version of the Quality of Well-Being Scale has been developed, the QWB-SA (Andresen et al., 1998). The QWB-SA has five sections: acute and chronic symptoms and problems (58 items), self-care (two items), mobility (i.e. use of transportation) (3 items) and physical functioning (8 items), and performance of usual activity (three items). In addition, there are three questions on overall health, and four demographic items, giving a total of 81.

Patients receiving RRT
Only limited evidence was found in relation to kidney disease. Moderate correlation is reported between the 10 and 7 subscale KDCS and QWB-SA and SF-6D in ESRD patients receiving RRT (Saban et al., 2008).

5. Sickness Impact Profile
This instrument has 12 domains reflecting the disability focus of quality of life with total scores ranging from 0 to 100. Physical, Psychosocial and Total scores can be calculated.

Transplant
Statistically significantly different scores were observed pre and post renal transplant in 293 patients as would be expected (Cetingok et al., 2004).

PREFERENCE-BASED MEASURES

Three preference-based measures evaluated with people with ESRD receiving RRT were included in the review:

1. EQ-5D
2. SF-6D
3. HUI

1. EQ-5D

The European Quality of Life instrument (EuroQol[2]), now generally known as the EQ-5D, was developed by researchers in five European countries to provide an instrument with a core set of generic health status items. There are two sections to the EuroQol: the EQ-5D and the EQ thermometer. The EQ-5D assesses health across five domains: anxiety/depression (AD), mobility (M), pain/discomfort (PD), self-care (SC), and usual activities (UA). Each domain has one item and a three-point categorical response scale; health 'today' is assessed. Weights based upon societal valuations of health states are used to calculate an index score. A score profile can be reported. The EQ thermometer is a single 20-cm vertical visual analogue scale with a range of 0 to 100, where 0 is the worst and 100 the best imaginable health.

Patients receiving RRT

Five studies report some evidence from evaluations with people with ESRD receiving different replacement therapies. Three were conducted in the UK.

Comparative performance is reported in a UK study (Gerard et al., 2004). The SF-36 (from which SF-6D can be derived) and the EQ-5D were evaluated with 626 patients undergoing haemodialysis. Mean health state utilities were comparable for both the EQ-5D and SF-6D but with different distributions. Agreement between the two preference measures was poor. Construct validity for EQ-5D was supported with fair to moderate correlations reported for EQ-5D with analogous domains. The EQ-5D discriminated between the presence of a disability and co-morbidities. Importantly, higher response rates were reported for the EQ-5D than for the SF-6D (93% vs 79%) (Gerard et al., 2004).

Other studies provide evidence of construct validity. Lower utility scores were reported for patients receiving HD or PD than those that had received a renal transplant (Lee et al., 2005 UK). Statistically significant differences in EQ-5D utilities were reported for patients receiving HD compared to UK population norms (Roderick et al., 2005 UK).

[2] http://www.euroqol.org/home.html

Very limited evidence was found for the responsiveness of EQ-5D in longitudinal studies despite the fact that a recent meta-analysis found evidence of extensive use of EQ-5D in relation to renal replacement therapy (Liem et al., 2008). A study carried out to evaluate possible benefits of night-time over conventional haemodialysis found that whilst some significant improvements were found via kidney specific measures (KDQoL), no significant change over time was observed for EQ-5D (Culleton et al., 2007; Manns et al., 2009).

A lower response rate of 33% was reported for EQ-5D in the study by Lee and colleagues (2005). However it is difficult to disentangle effects of EQ-5D from other instruments in this postal survey.

No studies were identified for AKI, and post transplant patients.

2. SF 6D

The SF-6D was constructed from eight items selected from the SF-12 and three from the SF-36 transformed into six items assessing physical functioning, role limitations, social functioning, pain, mental health and vitality (Brazier et al., 2002). Valuations from the general public of combinations of health states, together with regression techniques allow a single index value of health status to be estimated from responses to the 11 items.

Patients receiving RRT

Three studies provided some evidence for the SF-6D in relation to kidney disease, one in the UK. The results of SF-6D in the study by Gerard and colleagues (2004) have already been mentioned, providing evidence of construct validity but lower response rates compared with EQ-5D. Other studies find evidence of construct validity, for example correlations of SF-6D with Beck Depression scores and the HUI (Davison et al., 2008a), with kidney disease symptoms and the QWB-SA (Saban et al., 2008).

Substantial ceiling effects are reported in Davison et al. (2008a; Canada) with 23% to 39% of responses at the higher end of the scales for RL, SF and Pain). Floor effects are reported for RL (35%) (Davison et al., 2008; Canada).

No studies were identified for AKI, and post transplant patients.

3. HUI 2 & 3

The Health Utilities Index (HUI) was designed as a comprehensive framework for describing health status and health-related quality of life for use in clinical studies, population health surveys, and economic evaluations (Feeny et al., 1995).

The original HUI has been superseded by HUI2 and HUI3. HUI2 consists of seven attributes or dimension, namely, sensation, mobility, emotion, cognition, self-care, pain, and fertility. The Health Utilities Index Mark 3 (HUI3) consists of eight dimensions, rated by members of the general population as the most important dimensions of health status, namely, vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain or discomfort. For each attribute, there are five or six

levels of functioning, ranging from highly impaired to normal, defined in terms of capacity rather than performance, to avoid confounding abilities with preferences. A combination of levels across the attributes constitutes a health state; utility scores, based on community preferences, can be obtained for each health state using an algorithm, with 0 representing death and 1 perfect health.

Patients receiving RRT

In relation to kidney disease, two studies provided some evidence of construct validity. In the study by Davison et al. (2009), the mean preference scores were lower than population norms. Moderate correlation is reported between the SF-6D and HUI scores (Davison et al., 2008b)

The HUI2 and HUI3 discriminated between kidney disease severity, pre-dialysis and patients receiving dialysis and Beck Depression scores (Davison et al., 2008b).

The content of the HUI3 is considered to reflect the cognitive, pain and emotional deficits associated with CKD.

Substantial ceiling effects are reported for six of the 8 attributes (26 to 88%) in Davison et al. (2008b; Canada). No floor effects were reported in this study population.

No studies were identified for AKI, and post transplant patients.

RENAL-SPECIFIC MULTI-DIMENSIONAL QUESTIONNAIRES

Eight renal-specific multidimensional measures are included in the review:

1. Quality of Life Index-Dialysis (QLI-D)
2. Kidney Disease Quality of Life-Long Form (KDQOL-LF)
3. Kidney Disease Quality of Life-Short Form (KDQOL-SF)
4. Kidney Disease Questionnaire (KDQ)
5. Kidney Transplant Questionnaire (KTQ)
6. Renal Quality of Life Profile (RQLP)
7. CHOICE Health Experience Questionnaire (CHEQ)
8. Renal Dependant Individualised Quality of Life Questionnaire

1. Quality of Life Index-D (QLI-D)

The Quality of Life Index (QLI) was developed in the USA during the 1980s as a measure of morbidity for application in both normal and unwell populations (Ferrans and Powers, 1985). The original instrument, with the addition of six dialysis-specific items, was developed and tested in patients receiving haemodialysis (Ferrans and Powers, 1985); factor analysis confirmed instrument construction.

The instrument comprises two sections assessing respondent satisfaction and relative importance of each domain, respectively. Each section has 32 items, with eight items per domain. Six-point ordinal response scales range from ‘very dissatisfied’ or ‘very unimportant’ (1), to ‘very satisfied’ or ‘very important’ (6). Scoring is complicated and the developers recommend a computer programme. In summary, importance scores are used to weight satisfaction scores; index or domain scores range from 0 to 30, where higher scores indicate better quality of life.

Patients receiving RRT

Regarding evidence in relation to kidney disease, reliability was supported for the QLI in a one month time interval for dialysis patients. High internal consistency was also reported in a small study by (Ferrans and Power 1985) and reproduced in a further study with a larger sample of patients (Ferrans and Powers, 1992).

Additional items for haemodialysis patients relating to treatment were added to each section (Satisfaction with various domains and Importance of the domain to the individual) in Ferrans and Powers 1985). Items were endorsed by patients receiving haemodialysis. A transplant version is also available which included two items relating to the potential for a successful transplant. This is for patients receiving haemodialysis and on the transplant list.

A four factor structure was supported in Ferrans and Powers (1992) of Health and functioning, socioeconomic, psychological/spiritual and family. A high order factor was revealed representing Quality of Life.

The QLI-D discriminated between socioeconomic variables (income) pre-specified by the developers (1992). People in a lower income bracket reported statistically significantly lower scores than the higher income group. Further support is provided in the study by Ferrans and Powers (1992) in terms of education and age. Further discriminative properties are provided with decreasing scores reported with concomitant increase in number of symptoms reported (Jablonski, 2007).

Moderate correlation of QLI-D scores with a life satisfaction questionnaire has been reported (Ferrans and Powers 1985). Further convergent validity is supported for each domain and life satisfaction, with higher correlations for the Psychological/spiritual domain. Moderate correlation was reported between scores from the QLI-D and other patient-reported measures of symptoms and psychological adjustment to disease (Killingworth and Van Der Akker, 1996, UK). Moderate correlation of scores has been reported between QLI-D and symptoms (Jablonski, 2007).

A larger population was recruited in another study by the developers (Ferrans and Powers, 1992). This included 349 patients from a haemodialysis unit and questionnaires mailed to patients. 20% of patients had missing values greater than 15% and overall computable responses were available from 46% of participants invited. A 46% response rate was obtained to postal administration of the questionnaire (Ferrans and Powers, 1992). A total of 86% (CAPD patients) and 48% (HD patients) returned questionnaires in a UK study (Killingworth and Van Der Akker, 1996).

Transplant
Statistically significant scores were observed pre and post renal transplant in 293 patients as expected (Cetingok et al., 2004).

2. Kidney Disease Quality of Life-Long Form (KDQOL-LF) instrument (dialysis version).

The long form of the KDQOL™ (134 items) was developed in 1994 by the Kidney Disease Quality of Life Working Group. The KDQOL-LF includes the original SF-36 items (PCS and MCS) together with additional items assessing kidney specific issues: Symptoms and problems (34 items), Effects of kidney disease (20), Sleep quality (9), Burden of kidney disease (4), Cognitive function (6), Social support (4), Dialysis staff encouragement (6) and Patient satisfaction (2). These are referred to as the Kidney Disease Component Summary (KDCS).

Patients receiving RRT
Three studies are included in the current review, one from the UK.

Exploratory factor analysis of the SF-36 items and the kidney-specific scales yielded four distinct dimensions: Physical Functioning (PCS), Mental Health (MCP), Kidney-specific and Patient satisfaction (KDCS) (Hays et al., 1994).

The evidence for the KDCS is reported here.

High internal consistency (alphas greater than 0.80) was reported during development of all domains with the exception of Quality of social interaction and Sleep (Hays et al., 1994).

High internal consistency reliability is reported for the self-administered mode of administration of the KDQOL-LF kidney domains except Staff encouragement (Unruh et al., 2003). Similar results were reported for the interview-administered mode. Rao et al. (2000) report acceptable internal consistency of items and domains from the Symptoms/Problems and Effects of Kidney disease scales. Item discriminative validity is supported for both methods of administration.

Validity is reported for the Symptoms/problems scale which was sensitive to differences in the number of good days reported for a typical week in the last month (Hays et al., 1994).

No change in scores were reported in the study by Roderick et al. (2005, UK) with the exception of the patient satisfaction domain when comparing outcomes from people receiving HD in renal satellite units and those receiving HD in main renal units. This was consistent with other measures in the study.

Floor effects (60%) were reported for the Work status domain and Ceiling effects for Sexual functioning (39%) (Hays et al.,1994). No floor or ceiling effects were found in the study by Unruh et al., (2003).

43% of patients chose to be interviewed as mode of administration in a large study (978 HD patients). These patients tended to be older than the participants in the study who chose self-report (Unruh et al., 2003). Item completion rate was 93 to 99%. No differences were reported for mode of administration. This study examined bias in mode of administration and was part of a trial of RRT regimes. There were significantly higher scores on the Effects of kidney disease scale for the interview administered group of patients. A response rate of 82 % was reported in the study by Roderick et al. (2005, UK) but this was interview administered.

Conservative management
SF-36 and KDQOL scores discriminated between patients with corrected anaemia with significantly higher scores for patients with Hb in the target range of 12.0-13.0g/dL. FACT-Fatigue and FACT-Anaemia scales were also used in this study (Alexander et al., 2007).

Anaemia as a result of CKD is common and associated with iron deficiency. Often patients need oral iron supplementation or intravenous replacement. The SF-12 and KDCS of the KDQOL were used in a trial of oral Vs. intravenous iron supplementation. Small but significant differences in scores in the same direction were observed for the intravenous group of patients. This was associated with significant improvement in haemoglobin levels in these patients (Agarwal et al., 2006).

-KDQOL-Cognitive Functioning domain

Patients receiving RRT

The CF sub-scale was evaluated specifically as a patient-reported method of screening for cognitive impairment in a small study (160 participants) of older people with CKD. Some ceiling effects were reported with 26% of patients scores at the upper scale (Kurella et al., 2004). Modest correlation of scores were reported between the CF sub-scale and 3MS (r=0.31) (clinician assessed cognitive functioning test) but stronger correlation between CF scores and patient-reported depression (r=0.56) (Geriatric Depression Scale). Screening properties were explored and varying cut-off points of scores were examined for dementia diagnostic properties. Using a reference score of 3MS <80 as indicative of cognitive impairment, a CF cut-point score at 33 correctly classified the largest number of patients (83%). The sensitivity of this cut-point was only 15% but specificity 98%. Raising the cut-point to 60 enhanced the positive predictive value (PPV) and slightly decreased the negative predictive value (NPV). This study also considered the comparative costs of clinician assessed cognitive function tests and patient-reported methods (KDQOL-CF) and concluded that cost savings could be made if the CF-scale is used as a screening tool. Those patients with scores less than 60 only could then be offered more comprehensive cognitive assessment. (Kurella et al., 2004).

A larger study population of over 2000 patients who had recently started RRT reported significantly lower CF sub-scores in a group of patients who reported sleep difficulties (Kutner et al., 2007)

Independent predictors of the CF score have been reported as ESRD status, other co-morbidities, sleep medications, and biochemical concentrations (serum phosphorus; serum albumin) (Kurella et al 2004; Kutner et al., 2007).

3. Kidney Disease Quality of Life Short Form (KDQOL-SF Version. 1.3)

Item selection for a short version (KDQOL-SF) was conducted empirically by regression analyses on data derived from the longer version with results endorsed by patients and experts (Hays, 1995). The 80 item instrument comprise the 36 items from the SF-36, 43 items addressing all of the scales of the original KDQOL and a single item addressing overall health. It has high internal consistency for most domains with the exception of Quality of social interaction. Cognitive function domain did not reach alpha greater than 0.80 (Hays, 1995).

Patients receiving RRT

The KDCS (excluding sexual activity) was evaluated in Saban et al. (2008). Confirmatory factor analysis resulted in strong variances for the ten domains with the exception of patient satisfaction, work status and dialysis staff encouragement. Strong correlation is reported between domain scores for the SF and LF.

All domains of KDQOL-SF discriminated between different RRT modalities with expected differences were observed for scores for specific domains. Regression analysis indicated that treatment modality was the most important variable determining kidney disease burden (Carmichael et al., 2000;UK). The KDQOL-SF discriminated between the presence of a disability and co-morbidities (Manns et al., 2003) and between patients with expected demographic differences in a large international study including UK (Lopes et al., 2007). Lower scores were highly predictive of mortality and hospitalisation in a large international study (Mapes et al., 2003).

Lower scores were reported for patients receiving HD or PD than those that had received a renal transplant as expected (Lee et al., 2005 UK). No difference in scores on KDCS between patients receiving home haemodialysis and home peritoneal dialysis which was consistent with other health status measures in the study (Fong et al., 2007; Manns et al., 2009)

A low response rate of 33% was reported in Lee et al. (2005, UK). This was a large study (target population 1251) and other questionnaires included in the postal survey were the EQ-5D. It is feasible that the burden of completion of all these questionnaires without supervision was considerable.

A range of higher response rates have been reported in other studies: 70% to 80% (Carmichael et al., 2000, UK); 67% (Fong et al., 2007) and 66% (Johanson et al., 2007).

KDQOL-36- This is referred to in some text but no published evaluations have been identified. It includes the following items: Items 1-12: SF-12, Burden of kidney disease (4), Symptoms/problems (12), Effects of kidney disease (8). Scoring is suggested the same as in the manual by the developers. An online fee based scoring system is available on the Life Options Program (Medical Education Institute) (MEI) http://www.kdqol-complete.org/

4. Kidney Disease Questionnaire (KDQ)

This 26 itemed questionnaire was developed in Canada (Laupacis et al., 1991) with the involvement of patients receiving haemodialysis and empirically by factor analysis. Five domains included are: Physical symptoms (6 individualised symptoms identified by the patient); Fatigue: 6); Depression (5); Relationships (6); Frustration (3). Responses are scored in a 7 point Likert scale during the last 2 weeks. It is reported to take 10 to 15 minutes to complete.

Patients receiving RRT
Reproducibility is supported with ICCs above 0.80 for all domains. Construct validity is reported with moderate correlations with analogous domains using the SIP. Trial data (reported in Laupacis et al., 1992), provide support for responsiveness with significant improvement in scores for patients receiving treatment for anaemia which was consistent with score changes on the SIP.

Conservative management
Lefebvre et al., 2006 used the KDQ together with the LASA (not reported here as developed for people with cancer related anaemia) in a large study of patients with CKD not receiving RRT and receiving epoetin alfa for anaemia. There was a significant improvement over time of scores on the KDQ which was associated with an increase in haemoglobin levels to the normal range. Correlation of scores and levels of haemoglobin were modest (0.21 for Frustration to 0.46 for Physical symptoms).

5. Kidney Transplant Questionnaire

The KTQ was developed by Laupacis et al. (1993) with the involvement of renal transplant patients and clinical experts. 25 items are classified in 5 domains: Physical symptoms, Fatigue, Uncertainty/fear, Appearance and Emotions. Responses are obtained on a 7 point Likert scale. Construct validity is reported with moderate correlation of scores between specific domains and similar domains from the SIP. High internal consistency is reported for Fatigue and Emotional but lower alphas reported for other domains. Reproducibility is reported in patients between 6 and 12 months post-operatively with ICCs above 0.70. There were significantly different scores in patients pre and post transplant supporting responsiveness.

The KTQ scores were sensitive to change in a group of post-transplant patients receiving different immunosuppressant regimes (including UK population) (Russ et al., 2007). Scores for patients receiving a new regime were significantly different on KTQ Fatigue, Emotions and Appearance domains. Similar changes were detected in SF-36 scores for similar domains.

6. Renal Quality of Life Profile (RQLP)

The Renal Quality of Life profile (RQLP) is a 43 itemed questionnaire with a 5 point Likert scale for responses. Five dimensions include: Eating and drinking, Physical activities, Leisure time, Psychosocial activities and Impact of treatment. It was developed adopting a comprehensive methodology involving patients and clinicians in the UK (Salek. 1999).

Patients receiving RRT
Principal component factor analysis supported the five dimensions.

A high response rate is reported in Barton et al., (2009).The RQLP scores were responsive to change in a trial of pharmacy care compared to standard care for patients receiving HD. Effect sizes were moderate (Barton et al., 2009).

Moderate correlation is reported between the RQLP and SF-36 dimensions which were similar in construct (Boyd et al., 2009).

7. CHOICE Health Experience Questionnaire (CHEQ)

The Choices for Healthy Outcomes in Caring for ESRD (CHOICE) study was designed to evaluate the effectiveness of alternative dialysis prescription. As part of the CHOICE study, the CHEQ as patient-reported HRQOL instrument was developed to specifically complement the SF-36; be sensitive to dialysis treatment modalities and regimes; and be useful for longitudinal evaluation. A comprehensive, patient-centred approach was used during development. Items were derived from interviews with patients; literature; and clinicians' expertise (Wu et al., 2001). The questionnaire has 83 items addressing 21 domains: the 8 domains of the SF-36, 8 additional generic domains (cognitive functioning, sexual functioning, sleep, work, recreation, travel, finances, and general quality of life); and 5 ESRD-specific domains (diet, freedom, body image, dialysis access. The original study by Wu and colleagues (2001) provided some evidence for the reliability and validity of the scales.

Patients receiving RRT

Adequate internal consistency is reported for most domains in Wu et al. (2004). Test re-test reliability correlations range from 0.55 for body image to 0.79 for finance. Differences were observed for dimensions of the instrument between treatment modalities such as haemodialysis and peritoneal in Bass et al. (2004).

8. Renal Dependant Individualised Quality of Life Questionnaire

This instrument was developed out of an instrument used in relation to diabetes, the Audit of Diabetes Dependent Quality of Life (ADDQoL) diabetes-specific individualized quality of life questionnaire. From a small study with patients in eight U.K. renal clinics each of the 13 ADDQoL items were found relevant and important for renal patients. Additional items were also identified by patients including physical appearance, dependency, freedom, restrictions of fluid intake, and societal prejudice (Bradley, 1997, UK). No psychometric data for the new instrument were reported. No further studies using the instrument were identified.

RENAL-SPECIFIC SYMPTOM FOCUSED QUESTIONNAIRES.

Patients with ESRD have high symptom burden (Davison et al., 2006a) and the following section reports the evidence of questionnaires to measure kidney disease specific symptoms.

Thirteen symptom focused questionnaires are included in the review:

1. Modified Edmonton Symptom Assessment System (ESAS)
2. Memorial Symptom Assessment Scale-short form (MSAS-SF
3. Dialysis Symptom Index
4. National Kidney Dialysis and Kidney Transplant Study (NKDKTS) symptom checklist
5. Transplant Symptom Occurrence and Distress Scale
6. Patient Outcome Scale--symptom module (POSs)
7. The Rome II
8. Gastrointestinal Rating Scale (GSRS)
9. Gastrointestinal Quality of Life Index (GIQLI)
10. Thirst Distress Scale
11. Multi-dimensional Fatigue Inventory (MFI-20)
12. Fatigue Severity Scale
13. Epworth Sleepiness Scale

1. Modified Edmonton Symptom Assessment System (ESAS)

The Modified Edmonton Symptom Assessment System **(**ESAS) was originally developed for use with cancer patients and measures physical and psychological symptom distress. It consists of nine Visual Analogue Scales (VAS) with 0-10 scale (o= none, 10 =severe) for each symptom- pain, activity, nausea, depression, anxiety, drowsiness, appetite, well-being, and shortness of breath. A 10^{th} item 'pruritis' was added to the original measure. A total symptom score is calculated by summing scores for all ten symptoms (range from 0 to 100). An unlabelled VAS is included to allow patients to identify a non listed symptom which is important to them.

Patients receiving RRT

Patients receiving peritoneal and haemodialysis were surveyed to identify common symptoms (Davison et al., 2006b). Cross-sectional validity is supported with significant correlation with specific symptom items and related domains from the SF-12 and KDQOL-SF. The overall symptom scale correlated highly with the KDQOL-SF symptom/problems items, effects of kidney disease and burden of kidney disease as would be expected in both cross-sectional analysis and longitudinal (Davison et al., 2006a; Davison et al., 2006b). Reproducibility is reported with ICC 0.70 for patients with a one week re-test period.

It is reported to be useful in clinical practice with limited staff and patient burden. Ease of interpretation and analysis is considered an attractive feature. This has been evaluated extensively with cancer patient including those receiving palliative care. It may have several uses for kidney disease patients as CKD has high symptom burden.

In cross-sectional and longitudinal evaluation, most patients had 7 of the ten symptoms listed in the measure. Of these, the most common were tiredness, decreased well-being, pruritis and pain (Davison et al., 2006b).

2. Memorial Symptom Assessment Scale-short form (MSAS-SF)

The MSAS-SF was developed to assess the presence, severity and frequency of 32 symptoms in patients with cancer.

Patients receiving RRT
Several of these symptoms were found to be prevalent and severe in chronic haemodialysis patients (Weisbord et al., 2003; Murtagh et al., 2007, UK). A 62% response rate was reported in Murtagh et al., 2007).

3. Dialysis Symptom Index

Patients receiving RRT
This 30 itemed index was developed from the MSAS-SF. Responses are obtained on a 5 point Likert scale. It measures the level of distress associated with the symptom. This was evaluated in chronic haemodialysis patients and several symptoms were found to be prevalent and severe (Weisbord et al., 2004). A comprehensive methodology was used to develop the index including item and content evaluation of four existing measures, focus groups with patients and renal services providers; expert endorsement of items and formal evaluation with patients.

Poor correlation between renal provider's assessment of symptoms and patients with providers underestimating the severity and number of symptoms (Weisbord et al., 2007)

Carreon et al. (2008) reported high prevalence of symptoms in chronic haemodialysis patients.

Similar scores were obtained in a recent study with similar number of symptoms for both ESRD and CKD patients (Abel-Kader et al., 2009). Results were comparable for other measures in this study (SF-36; PHQ).

4. National Kidney Dialysis and Kidney Transplant Study (NKDKTS) symptom checklist.

The NKDKTS symptom checklist was developed in 1987 (Evans et al.) with involvement of patients with ESRD receiving dialysis and patients who had received a kidney transplant. The questionnaire has 13 items measuring anaemia symptom frequency. Good evidence of discriminative validity between these two groups of patients was reported during the development phase. Initially it was developed for use in clinical trials in ESRD-related anaemia. No further evaluations are reported until a recent study (Spiegal et al., 2009). Factor analysis supports a single domain structure and internal consistency good. Reproducibility is acceptable with ICCs greater than 0.60. The questionnaire scores discriminated between patients with different

haemoglobin levels with those with a haemoglobin level less than 10g/dl reporting greater symptoms as hypothesised. No floor or ceiling effects were reported.

5. Transplant Symptom Occurrence and Distress Scale

This questionnaire was developed for people post heart transplant. It contains 27 items relating to the use of immunosuppressive drugs and assesses the symptoms and distress with a 5 point Likert scale. Limited evidence is reported for use with patients' post-renal transplant. Reliability is reported in Moons et al. (2001) cited in Zarifan (2006). Content validity is supported in Zarifan (2006) with patients (n=100) reporting occurrence of all symptoms and construct validity with decreased quality of life reported as the presence of symptoms increased.

6. Patient Outcome Scale--symptom module (POSs)

Conservative management
This index has been applied in patients with stage 4-5 CKD who are being managed conservatively without dialysis (Murphy et al., 2009 UK). This 15 item scale identifies the presence and severity of symptoms during the last three days. Two further symptoms specific to renal disease have been added (itch and restless legs). The short time frame for responses in this instrument is valid specifically as these patients have limited survival time. High symptom burden was significantly associated with comorbidity in patients with advanced kidney disease and receiving conservative management (Murphy et al., 2009).

7. The Rome II

Patients receiving RRT
The Rome II questionnaire has been applied in patients with ESRD receiving RRT to establish the prevalence of gastro-intestinal symptoms. Patients with ESRD were more likely to report IBS symptoms than community controls using the Rome II questionnaire which is specific to gastro-intestinal symptoms (Cano et al., 2007 UK).

8. Gastrointestinal Rating Scale (GSRS)

Transplant
The Gastrointestinal Rating Scale (GSRS) was specifically developed as a symptom checklist for people receiving immunosuppressive regimes. Fifteen items assess the impact of upper and lower gastrointestinal symptoms. This has been evaluated in post renal transplant patients. Good internal consistency is reported (Kleinman et al., 2006). Moderate correlation of scores is reported (≥0.40) between GIQLI domains and Psychological Well-being scales with the exception of PGWB-positive well-being and self control. Weak correlation of scores was reported for GSRS scores and EQ-5D (Kleinman et al., 2006). Responsiveness is reported with statistically significant improvement in scores for patients receiving enteric-coated immunosuppressive medications (Bolin et al., 2007).

9. Gastrointestinal Quality of Life Index (GIQLI)

Transplant

Gastrointestinal Quality of Life Index (GIQLI) is a 36 itemed instrument designed to assess the impact of GI symptoms on daily life in five domains (Symptoms, Emotional status, Physical function and Social function. This has been evaluated in post renal transplant patients. Good internal consistency is reported (Kleinman et al., 2006). Moderate correlation of scores is reported (≥0.40) between GIQLI domains and Psychological Well-being scales and EQ-6D (Kleinman et al., 2006).

10. Thirst Distress Scale

Patients receiving RRT

Patients with ESRD receiving RRT have additional burden of daily dietary and fluid restrictions and the latter in particular reported as a stressful aspect of self-management and often compliance is poor; the consequences of this can be serious. Non-compliance can result in interdialytic weight gain (IWG) which impacts significantly on respiratory and cardiovascular functioning. The Thirst Distress Scale was developed with an underpinning model of symptom evaluation and symptom response. Symptom evaluation refers to the frequency, duration, intensity etc of symptoms and symptom response may include emotional and physiological effects. Item derivation included interviews with patients, expert review and literature reviews. Thirty-one items addressing distress (12), duration (3) and frequency (16) are included and responses obtained on a 5 point Likert scale. Internal consistency is satisfactory (0.78). A single factor is supported in factor analysis (Welch. 2002). No further studies have been identified.

11. Multi-dimensional fatigue Inventory (MFI-20)

Multi-dimensional fatigue Inventory (MFI-20) is a 20 itemed instrument which measures 5 dimensions of fatigue; general, mental, physical, reduced motivation and reduced activity. Four items in each sub-scale are scored on a 6 point Likert scale.

Patients receiving RRT

Previous studies have reported low internal consistency suggesting the need for other items measuring fatigue (McCann and Boore 2000). High internal consistency is though reported in one study (O'Sullivan and McCarthy, 2006, Ireland).

12. Fatigue Severity Scale

All patients

Bonner et al., (2008) evaluated levels of fatigue using the Fatigue Severity Scale with patients with different classifications of kidney disease. The instrument was discriminative and scores were significantly higher indicating more fatigue in pre-dialysis patients, those with diabetic nephropathy and those using PD more fatigued that HD patients. Scores for all patients in this study were significantly higher than population norms. There was no difference in scores though between patients with different levels of haemoglobin and urea levels.

13. Epworth Sleepiness Scale

Patients receiving RRT

Epworth Sleepiness Scale is an 8 item measure of sleepiness with scores ranging from 0 to 24 with values of 10 and greater indicating significant sleepiness. It is well documents that patients with CKD receiving RRT have sleep problems with shorter duration, and less efficient sleep. The ESS discriminates between patients on HD and community controls with HD patients having higher scores (Unruh et al., 2008).

Acute Kidney Injury

Of the instruments identified in this review, no evaluations reporting psychometric criteria were identified with an acute kidney injury population. A recent review focusing on long-term outcomes for patients following acute renal failure cites six studies which report on HRQL and functional status of survivors of critical illness and organ failure (specifically renal failure) (Bagshaw 2006). The Nottingham Health profile featured in three studies and the EQ-5D in one other although this was with a Finnish population. No psychometric evidence was reported.

CONCLUSIONS AND RECOMMENDATIONS

Table 3 shows the appraisals of the evidence for each of the PROMs identified in this review.

Most of the instruments included in this review have been evaluated with people receiving RRT. Some evidence is found for evaluations with patients post renal transplant and less for those being managed conservatively.

With regard to the generic measures identified in the review, it is clear from the evidence reported and summarised in table 3 that the SF-36 is the only generic measure with extensive, good measurement properties and operational characteristics. Furthermore, the psychometric criteria and operational performance is replicable when administered as a stand alone measure and when used in combination with the Kidney Disease Domains in the KDQOL questionnaires. Convincing evidence is reported of its discriminative properties in patients with ESRD. Of particular interest is that several studies report lower than population mean scores for the PCS, but near normal scores for the MCS. Other valuable properties include predictive validity. Several studies report lower scores to be significantly associated with subsequent mortality and utilisation of healthcare resources. Similar response rates are reported whether administered as a stand alone measure or as part of the KDQOL. A very important limitation of the available evidence is that only modest testing of responsiveness could be found. If used to assess quality and the impact of services, this measurement property would be of particular importance.

Of the three preference-based measures identified in this review, evidence for the EQ-5D is more favourable. Three of the 4 studies found for EQ-5D in relation to renal disease were conducted in the UK. The evidence presented suggests good discriminative properties and high response rates for completion. Some ceiling effects have been observed. It has been recommended for use in a number of relevant contexts in the NHS. It is mandated for use as a PROM in relation to four elective surgical procedures on the basis of evidence of its performance. It has been recommended to be piloted in the NHS in relation to a range of six long term conditions. It is referred to as a quality of life measure for use as outcome measure in relation to commissioning pathways for kidney disease. As with SF-36 the absence of evidence of responsiveness is a concern if the EQ-5D is to be considered a potential measure in longitudinal studies examining quality and the impact of services.

The renal-specific measure with the most substantial supporting evidence is the KDQOL (whether in longer or shorter format). As reported, this measure includes the SF-36 but the Kidney-disease domains can be scored as a Component score (KDCS). The Long Form (total of 134 items) may be burdensome for some patients. The Short-Form (total of 80 items) is possibly a more attractive version than the longer form, although it may still be considered demanding in terms of length. Some positive evidence is found for the QLI-D.

Given the volume and weight of evidence, the following questionnaires are suggested as the most favourable options depending on context and purpose of measurement:

The SF-36 should be considered for further evaluation where a more comprehensive assessment is needed of a patient's general health status than is provided by, for example, the EQ-5D.

The EQ-5D should be considered for further evaluation as a preference measure. A particular attraction is its brevity, lower response burden and the ability to derive utility values which have been based on UK samples.

The KDQOL instrument includes the SF-36 but additional items specific to renal disease. It is not apparent that there is significant value of the longer over the shorter version and the likelihood of lower respondent burden and higher responses rates make the shorter version potentially more appropriate. However although shorter, an 80-item instrument may still be considered long for routine and large scale administration. The amount of testing of the use of different versions of KDQOL in the NHS is still modest so further assessment before it is possible to be confident in recommending the short version of KDQOL.

Given the overlap between instruments such as SF-36 and KDQOL the issue of how instruments are used in combination needs to be flagged up. There is no merit, for example, in using both SF-36 and KDQOL in the same survey whereas the combination of EQ-5D and KDQOL would provide complementary and non-overlapping evidence of patients' perceptions in relation to kidney disease.

Symptoms

Several narrow focused and single dimension questionnaires are included in the review. These focus on symptoms patients with CKD experience. Some instruments have been developed specifically with and for CKD patients; for example the Dialysis Symptom Index which was developed to identify symptoms specific to those experienced by people receiving RRT. The National Kidney Dialysis and Kidney Transplant Study (NKDKTS) symptom checklist and Transplant Symptom Occurrence and Distress Scale have been developed and evaluated for patients post transplant.

General symptom based instruments have been applied and evaluated with CKD patients (ESAS and MSAS-SF). These were primarily developed with people with cancer as they have high symptom burden. Patients with CKD are considered to have comparable presence of symptoms and related impact.

Patients post transplant will have life long immunosuppressive therapy which is not without unpleasant side effects. Three instruments were included to measure the impact and severity of gastrointestinal symptoms patients experience whilst receiving this therapy (Rome II, GSRS and GIQLI). Intensive thirst associated with fluid restriction during non-dialysis period can be intense and the Thirst Distress Scale attempts to identify the severity and impact on the patient.

Fatigue is a common debilitating symptom for all CKD patients and can be associated with anaemia- The MFI-20, FSS and ESS have been evaluated although these can be applied in most patients who experience fatigue and sleep disorders.

The Patient Outcome Scale--symptom module (POSs) is a short questionnaire with a 3 day recall period designed to identify symptom impact during the end of life and includes two renal specific symptoms. Limited evidence is reported.

Although these short questionnaires which have very specific symptoms may appear attractive in clinical practice as screening tools or to identify the presence and severity of symptoms, they are limited in their scope to measure broader aspects of the health and the impact of the disease.

All of the included questionnaires measuring symptoms have limited psychometric evidence and therefore we cannot make sharp recommendations. What should be emphasised though is the importance of measuring the impact of symptoms for these patients. It may be that the KDCS symptom domain included in the KDQOL captures what is important to patients.

General comment

So much of the evidence assessed in this review examines PROMs in the context of cross-sectional studies. In the context of assessing quality and outcomes of services, it is more likely that PROMs will be used longitudinally to provide some estimate of positive or negative effects of services. There is very little evidence of such uses of PROMs to assess quality and outcomes of services. Furthermore PROMs can only be meaningfully interpreted in study designs that attempt to take account of co-morbidities, socio-demographic factors and other potential confounders of relationships. Appropriately designed pilots are needed to inform decisions about the use of PROMs in relation to quality and outcomes of services.

Table 3: Appraisal of PROMs included in the review

PROM	Reproducibility	Internal consistency	Validity Content	Construct	Responsiveness	Interpretability	Precision	Acceptability	Feasibility
Generic measures									
SF-36		X	X	XXX	XXXX	XX		X	
SF-20			X	X				X	X
SF-12				X				X	
QWB-SA				X					
SIP					X				
Preference-based measures									
EQ-5D		n/a	X	XX				X	
SF-6D				X				X	
HUI				X					
Renal-specific measures									
QLI-D	X	X	X	X X	X			X	
KDQOL-LF		X	X	XX	X			X	
KDQOL-SF			X	XX				X	
KDQ	X		X	X	X				
KTQ			X		X				
RQLP			X	X				X	
CHEQ	X	X	X	X					

Appendix A: Appraisal of the methodological quality of PROMs

A simple rating scale (Table 1) was used to rate the sum total of evidence available for each dimension or criterion against which PROMs were assessed. The dimensions or criteria are summarised in Table 2.

Table 1: Psychometric and operational criteria

0	*not reported (no evaluation completed)*
—	*Evaluation evidence available indicating poor performance of instrument*
+	*Some limited evidence in favour*
++	*Good evidence in favour*
++ +	*Excellent evidence in favour.*

Table 2 Appraisal criteria for assessing PROMs

Appraisal component	**Definition/test**	**Criteria for acceptability**
Reliability		
Test-retest reliability	The stability of a measuring instrument over time; assessed by administering the instrument to respondents on two different occasions and examining the correlation between test and re-test scores	Test re-test reliability correlations for summary scores 0.70 for group comparisons
Internal consistency	The extent to which items comprising a scale measure the same construct (e.g. homogeneity of items in a scale); assessed by Cronbach's alpha's and item-total correlations	Cronbach's alphas for summary scores ≥0.70 for group comparisons Item-total correlations ≥ 0.20
Validity		
Content validity	The extent to which the content of a scale is representative of the conceptual domain it is intended to cover; assessed qualitatively during the questionnaire development phase through pre-testing with patients. Expert opinion and literature review	Qualitative evidence from pre-testing with patients, expert opinion and literature review that items in the scale represent the construct being measured Patients involved in the development stage and item generation
Construct validity	Evidence that the scale is correlated with other measures of the same or similar constructs in the hypothesised direction; assessed on the basis of correlations between the measure and other similar measures	High correlations between the scale and relevant constructs preferably based on a priori hypothesis with predicted strength of correlation
	The ability of the scale to differentiate known-groups; assessed by comparing scores for sub-groups who are expected to differ on the construct being measured (e.g a clinical group and control group)	Statistically significant differences between known groups and/or a difference of expected magnitude
Responsiveness	The ability of a scale to detect significant change over time; assessed by comparing scores before and after an intervention of known efficacy (on the basis of various methods including t-tests, effect sizes (ES), standardised response means (SRM) or responsiveness statistics	Statistically significant changes on scores from pre to post-treatment and/or difference of expected magnitude
Precision	The ability of an instrument to measure accurately across full spectrum of a construct	Floor/ceiling effects for summary scores <15%
Practical properties		
Acceptability	Acceptability of an instrument reflects' respondents' willingness to complete it and impacts on quality of data	Low levels of incomplete data or non-response
Feasibility/burden	The time, energy, financial resources, personnel or other resources required of respondents or those administering the instrument	Reasonable time and resources to collect, process and analyse the data.

REFERENCES

Abdel-Kader K, Unruh ML, Weisbord SD. Symptom burden, depression, and quality of life in chronic and end-stage kidney disease Clin J Am Soc Nephrol. 2009 Jun;4(6):1057-64. Epub 2009 May 7

Agarwal R, Rizkala AR, Bastani B, Kaskas MO, Leehey DJ, Besarab A. A randomized controlled trial of oral versus intravenous iron in chronic kidney disease. *Am J Nephrol* 2006; 26(5):445-454.

Alexander M, Kewalramani R, Agodoa I, Globe D. Association of anemia correction with health related quality of life in patients not on dialysis. *Curr Med Res Opin* 2007; 23(12):2997-3008.

Andresen EM, Rothenberg BM, Kaplan RM. Performance of a self-administered mailed version of the Quality of Well-Being questionnaire/QWB-SA among older adults. *Medical Care* 1998; 36(9):1349-1360.

Bagshaw SM. The long-term outcome after acute renal failure. *Current Opinion in Critical Care* 2006; 12(6):561-566.

Barton A, Boyd A. Chavez A, Manley H. Health-related quality of life is maintained in hemodialysis patients receiving pharmaceutical care: A 2-year randomized, controlled study *Hemodialysis International,* Volume 13, Number 1, 2009 , pp. 72-79(8)

Bass EB, Wills S, Fink NE, Jenckes MW, Sadler JH, Levey AS, Meyer K, Powe NR. How strong are patient preferences in choices between dialysis modalities and doses? *Am J Kidney Dis*. 2004 Oct;44(4):695-705

Beusterien KM, Nissenson AR, Port FK, Kelly M, Steinwald MB, Ware Jr JE. The effects of recombinant human erythropoietin on functional health and well-being in chronic dialysis patients. *Journal of the American Society of Nephrology*, Vol 7, 763-773

Bolin P, Tanriover B, Zibari GB, Lynn ML, Pirsch JD, Chan L et al. Improvement in 3-month patient-reported gastrointestinal symptoms after conversion from mycophenolate mofetil to enteric-coated mycophenolate sodium in renal transplant patients. *Transplantation* 2007; 84(11):1443-1451.

Bonner A, Wellard S, Caltabiano M. Levels of fatigue in people with ESRD living in far North Queensland. *Journal of Clinical Nursing* 17, 90-98

Boyd A, McQuade C, Conner T, Manley H, Barton A. Correlation between the renal quality of life profile and SF-36 in a United States haemodialysis population. *Haemodialysis International*. 2009; 13: 89-90

Bradley C. Design of a renal-dependent individualized quality of life questionnaire. *Advances in Peritoneal Dialysis* 1997; 13:116-120.

Brazier JE, Roberts j, Deverill M: The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 2002; 21 :271-292.

Brenner I, Brohart K. Weekly energy expenditure and quality of life in hemodialysis patients. *CANNT J* 2008; 18(4):36-40.

Browne J, Jamieson L, Lewsey J van der Meulan J, Black N, Cairns J, Lamping D, Smith S, Copley L, Horrocks J. Patient Reported Outcome Measures (PROMs) in Elective Surgery (2007). Health Services research Unit, London School of Hygiene and tropical Medicine.

Cano AE, Neil AK, Kang JY, Barnabas A, Eastwood JB, Nelson SR et al. Gastrointestinal symptoms in patients with end-stage renal disease undergoing treatment by hemodialysis or peritoneal dialysis. *Am J Gastroenterol* 2007; 102(9):1990-1997.

Carmichael P, Popoola J, John I, Stevens PE, Carmichael AR. Assessment of quality of life in a single centre dialysis population using the KDQOL-SF questionnaire. *Quality of Life Research* 2000; 9(2):195-205.

Carreon M, Fried LF, Palevsky PM, Kimmel PL, Arnold RM, Weisbord SD. Clinical correlates and treatment of bone/joint pain and difficulty with sexual arousal in patients on maintenance hemodialysis. *Hemodial Int.* 2008 Apr;12(2):268-74

Cetingok M, Winsett RP, Hathaway DK. A comparative study of quality of life among the age groups of kidney transplant recipients. *Progress in Transplantation* 2004; 14(1):33-38.

Cleary J, Drennan J. Quality of life of patients on haemodialysis for end-stage renal disease. *Journal of Advanced Nursing* 2005 Sep;51(6):577-86

Culleton BF, Walsh W, Klarenbach SW, Mortis G, Scott-Douglas N, Quinn RR, Tonelli M, Donnelly S, Friedrich M, Kumar A, Mahallati H, Hemmelgarn BR/. Manns BJ. Effect of Frequent Nocturnal Hemodialysis vs Conventional Hemodialysis on Left Ventricular Mass and Quality of Life: A Randomized Controlled Trial *JAMA.* 2007; 298(11):1291-1299.

Curtin,R.B.; Sitter,D.C.B.; Schatell,D.; Chewning,B.A. Self-management, knowledge, and functioning and well-being of patients on haemodialysis. *Nephrology-Nursing-Journal.* 2004. 31 (4): 378-396

Darzi A. Our NHS Our Future: NHS Next Stage Review Interim Report. Department of Health, London, October 2007.

Darzi A. Our NHS Our Future: High Quality Care for All. NHS Next Stage Review Final Report. Department of Health, London, June 2008.

Davison SN, Jhangri GS, Feeny DH. Comparing the Health Utilities Index Mark 3 (HUI3) with the Short Form-36 Preference-Based SF-6D in Chronic Kidney Disease. *Value Health* 2008a.

Davison SN, Jhangri GS, Feeny DH. Evidence on the construct validity of the Health Utilities Index Mark 2 and Mark 3 in patients with chronic kidney disease. *Qual Life Res* 2008b; 17(6):933-942.

Davison SN, Jhangri GS, Johnson JA. Cross-sectional validity of a modified Edmonton symptom assessment system in dialysis patients: A simple assessment of symptom burden. *Kidney International* 2006a; 69(9):1621-1625.

Davison SN, Jhangri GS, Johnson JA. Longitudinal validation of a modified Edmonton symptom assessment system (ESAS) in haemodialysis patients. *Nephrol Dial Transplant* 2006b; 21(11):3189-3195

DeOreo PB. Hemodialysis patient-assessed functional health status predicts continued survival, hospitalization, and dialysis-attendance compliance. *American Journal of Kidney Diseases* 1997; 30(2):204-212.

Department of Health. Our Health, Our Care, Our Say: A New Direction for Community Services. White Paper, London 2006.

Department of Health. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs) for the NHS in England 2009/10. Department of Health, London, December 2008

Department of Health. National Service Framework for Renal Services (2005) Department of Health, London, December 2008

Department of Health. The NHS Improvement Plan: Putting people at the heart of public services 2004

Department of Health. 18 week Commissioning Pathway for Chronic Kidney Disease (2008). Department of Health

Department of Health Quality and Outcomes Framework guidance for GMS contract 2008/9

Department of Health. End of Life Care in Advanced Kidney Disease: A Framework for Implementation 2009

Department of Health Renal NSF Team and Marie Curie Palliative Care Institute. 2008. Guidelines for LCP Prescribing in Advanced Chronic Kidney Disease

Feeny D, Furlong W, Boyle M, Torrance GW. Multi-Attribute Health Status Classification Systems: Health Utilities Index. *PharmacoEconomics* 1995; 7(6):490-502.

Ferrans CE, Powers MJ. Quality of Life Index: development and psychometric properties. *Advances in Nursing Science* 1985; 8(1):15-24.

Ferrans CE, Powers MJ. Psychometric assessment of the Quality of Life Index. *Research in Nursing and Health* 1992; 15(1):29-38.

Fong E, Bargman JM, Chan CT. Cross-sectional comparison of quality of life and illness intrusiveness in patients who are treated with nocturnal home hemodialysis versus peritoneal dialysis. *Clin J Am Soc Nephrol* 2007; 2(6):1195-1200.

Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 2(14).

Fitzpatrick R, Bowling A, Gibbons E, Haywood K, Jenkinson C, Mackintosh A, Peters M. A Structured Review of Patient-Reported Measures in Relation to Selected Chronic Conditions, Perceptions of Quality of Care, and Carer Impact. Oxford: National Centre for Health Outcomes Development, November 2006.

Garratt AM, Schmidt L, Mackintosh A, Fitzpatrick R. (2002) Quality of life measurement: bibliographic study of patient assessed health outcome measures. *British Medical Journal;* 324 (7351): 1417-1421.

Gerard K, Nicholson T, Mullee M, Mehta R, Roderick P. EQ-5D versus SF-6D in an Older, Chronically Ill Patient Group. *Applied Health Economics and Health Policy 2004* 2004; 3(2).

Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Development of the Kidney Disease Quality of Life/KDQOL instrument. *Quality of Life Research* 1994; 3(5):329-338.

Hays R, Kallich J, Mapes D, Coons S, Amin N, Carter W, Kamberg C. Kidney Disease Quality of Life Short Form (KDQOL-SF ™), Version 1.3 A Manual for Use and Scoring. RAND 1995.

Jablonski A. Level of symptom relief and the need for palliative care in the hemodialysis population. *Journal of Hospice and Palliative Nursing* 2007; 9(1):50-60.

Johansen KL, Painter P, Kent-Braun JA, Ng AV, Carey S, Da Silva M et al. Validation of questionnaires to estimate physical activity and functioning in end-stage renal disease. *Kidney International* 2001; 59(3):1121-1127.

Johansen KL, Chertow GM. Chronic kidney disease mineral bone disorder and health-related quality of life among incident end-stage renal-disease patients. *J Ren Nutr* 2007; 17(5):305-313.

Kellum JA, Bellomo R, Ronco C. The concept of acute kidney injury and the RIFLE criteria. *Contrib Nephrol*. 2007;156:10-6.

Khan IH, Garratt AM, Kumar A, Cody DJ, Catto GRD, Edward N et al. Patients' perception of health on renal replacement therapy: evaluation using a new instrument. *Nephrology, Dialysis and Transplantation* 1995; 10(5):684-689.

Killingworth A, Van den Akker O. The quality of life of renal dialysis patients: trying to find the missing measurement. *International Journal of Nursing Studies* 1996; 33(1):107-120.

Kleinman L, Kilburg A, Machnicki G, Faull R, Walker R, Prasad R et al. Using GI-specific patient outcome measures in renal transplant patients: validation of the GSRS and GIQLI. *Quality of Life Research* 2006; 15(7):1223-1232.

Knight EL, Ofsthun N, Teng M, Lazarus JM, Curhan GC. The association between mental health, physical function, and hemodialysis mortality. *Kidney International* 2003; 63(5):1843-1851.

Kurella M, Luan J, Yaffe K, Chertow GM. Validation of the kidney disease quality of life (KDQOL) cognitive function subscale. *Kidney International* 2004; 66(6)(6):2361-2367.

Kutner NG, Zhang R, Huang Y, Bliwise DL. Association of sleep difficulty with Kidney Disease Quality of Life cognitive function score reported by patients who recently started dialysis. *Clin J Am Soc Nephrol* 2007; 2(2):284-289.

Laupacis A, Wong C, Churchill DN, Cohen A, Barre P, Lazaravits A et al. The use of generic and specific quality of life measures in hemodialysis patients treated with erythropoietin. *Controlled Clinical Trials* 1991; 12(4 Suppl.):168S-179S.

Laupacis A, Muirhead N, Keown PA, Wong C. A disease-specific questionnaire for assessing quality of life in patients on hemodialysis. *Nephron* 1992; 60(3):302-306.

Lee J, Hodgson D, Chow E, Bezjak A, Catton P, Tsuji D et al. A phase II trial of palliative radiotherapy for metastatic renal cell carcinoma. *Cancer*- 2005; 104(9):1894-1900.

Lefebvre P, Vekeman F, Sarokhan B, Enny C, Provenzano R, Cremieux PY. Relationship between hemoglobin level and quality of life in anemic patients with chronic kidney disease receiving epoetin alfa. *Current Medical Research & Opinion* 2006; 22(10):1929-1937.

Levey AS, Eckardt KU, Tsukamoto Y. Definition and classification of chronic kidney disease: A position statement from Kidney Disease: Improving global outcomes (KDIGO) *Kidney International* 67: 2089-2100, 2005

Liem Y S, Bosch JL, Myriam Hunink. Preference-Based Quality of Life of Patients on Renal Replacement Therapy: A Systematic Review and Meta-Analysis *Value in Health.* Volume 11 Issue 4, Pages 733 -741

Lin,J.; Curhan,G.C. Kidney function decline and physical function in women. *Nephrol Dial Transplant*. 2008 Sep;23(9):2827-33. Epub 2008 Apr 8.

Lopes AA, Bragg-Gresham JL, Goodkin DA, Fukuhara S, Mapes DL, Young EW et al. Factors associated with health-related quality of life among hemodialysis patients in the DOPPS. *Qual Life Res* 2007; 16(4):545-557.

Lowrie EG, Curtin RB, LePain N, Schatell D. Medical Outcomes Study Short Form-36: a consistent and powerful predictor of morbidity and mortality in dialysis patients. *American Journal of Kidney Diseases* 2003; 41(6):1286-1292.

Manns B, Johnson JA, Taub K, Mortis G, Ghali WA, Donaldson C. Quality of life in patients treated with hemodialysis or peritoneal dialysis: what are the important determinants? *Clin Nephrol*. 2003 Nov;60(5):341-51.

Manns BJ, Walsh MW, Culleton BF, Hemmelgarn B, Tonelli M, Schorr M et al. Nocturnal hemodialysis does not improve overall measures of quality of life compared to conventional hemodialysis. *Kidney Int* 2009; 75(5):542-549.

Mapes DL, Lopes AA, Satayathum S, McCullough KP, Goodkin DA, Locatelli F, Fukuhara S, Young EW, Kurokawa K, Saito A, Bommer J, Wolfe RA, Held PJ, Port FK. Health-related quality of life as a predictor of mortality and hospitalisation: the Dialysis Outcomes and Practice Patterns Study (DOPPS) *Kidney Int.* 2003 Jul;64(1):339-49.

McCann K, Boore JR. Fatigue in persons with renal failure who require maintenance haemodialysis. *Journal of Advanced Nursing* 2000; 32(5):1132-1142.

Meers C, Lang J, McMurray M, Morton AR, Singer MA, Hopman WM et al. Measuring and predicting outcomes in ESRD patients. *J CANNT* 1993; 3(2):15-16.

Mittal SK, Ahern L, Flaster E, Maesaka JK, Fishbane S. Self-assessed physical and mental function of haemodialysis patients. *Nephrol Dial Transplant.* 2001 Jul;16(7):1387-94.

Morton AR, Singer MA, Meers C, Lang J, McMurray M, Hopman WM et al. Assessment of health status in peritoneal dialysis patients: a potential outcome measure. *Clinical Nephrology* 1996; 45(3):199-204.

Muehrer RJ, Becker BN. Life after transplantation: new transitions in quality of life and psychological distress. *Semin Dial.* 2005 Mar-Apr;18(2):124-31

Murphy EL, Murtagh FE, Carey I, Sheerin NS. Understanding symptoms in patients with advanced chronic kidney disease managed without dialysis: use of a short patient-completed assessment tool. *Nephron Clin Pract* 2009; 111(1):c74-c80.

Murtagh FE, Addington-Hall JM, Edmonds PM, Donohoe P, Carey I, Jenkins K et al. Symptoms in advanced renal disease: a cross-sectional survey of symptom prevalence in stage 5 chronic kidney disease managed without dialysis. *J Palliat Med* 2007; 10(6):1266-1276.

National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification and stratification. Am J Kidney Dis 2002; 39 (Suppl 2) S1- S246.

O'Sullivan D and McCarthy G. An exploration of the relationship between fatigue and physical functioning in patients with end stage renal disease receiving haemodialysis. *Journal of Clinical Nursing* Volume 16, Issue 11c, Pages 276-284.

Ozminkowski RJ, White AJ, Hassol A, Murphy M. General health of End Stage Renal Disease Program beneficiaries. *Health-Care Financing Review* 1997; 19(1):121-144.

Rao S, Carter WB, Mapes DL, Kallich JD, Kamberg CJ, Spritzer KL et al. Development of subscales from the symptoms/problems and effects of kidney disease scales of the Kidney Disease Quality of Life instrument. *Clinical Therapeutics* 2000; 22(9):1099-1111.

Roderick P, Nicholson T, Armitage A, Mehta R, Mullee M, Gerard K, et al. An evaluation of the costs, effectiveness and quality of renal replacement therapy provision in renal satellite units in England and Wales. *Health Technol Assess* 2005; ((24)

Russ G, Jamieson N, Oberbauer R, Arias M, Murgia MG, Blancho G et al. Three-year health-related quality-of-life outcomes for sirolimus-treated kidney transplant patients after elimination of cyclosporine. *Transpl Int* 2007; 20(10):875-883.

Saban KL, Stroupe KT, Bryant FB, Reda DJ, Browning MM, Hynes DM. Comparison of health-related quality of life measures for chronic renal failure: quality of well-being scale, short-form-6D, and the kidney disease quality of life instrument. *Qual Life Res* 2008; 17(8):1103-1115.

Salek MS. Quality of life in patients with end-stage renal disease. *Journal of Applied Therapeutic Research* 1999; 2(3):163-170.

Secretary of State for Health The NHS Improvement Plan 2004. London, HMSO, 2004.

Spiegel DM, Evans RW, Gitlin M, Mayne TJ. Psychometric evaluation of the National KidneyDialysis and Kidney Transplantation Study symptom checklist: reliability and validity. *Nephrol Dial Transplant* 2009; 24(2):619-625.

Streiner DL, Norman GR. (1995) Health Measurement Scales. A practical guide to their development and use. Oxford Medical Publications, Inc. Second Edition.

Sureshkumar KK, Patel BM, Markatos A, Nghiem DD, Marcus RJ. Quality of life after organ transplantation in type 1 diabetics with end-stage renal disease. *Clinical Transplantation* 2006; 20(1):19-25.

Unruh ML, Yan G, Radeva M, Hays RD, Benz R, Athienites N, V et al. Bias in assessment of health-related quality of life in a hemodialysis population: a comparison of self-administered and interviewer-administered surveys in the HEMO Study. *Journal of the American Society of Nephrology* 2003; 14(8):2132-2141.

Unruh ML, Newman AB, Larive B, Amanda DM, Miskulin DC, Greene T et al. The influence of age on changes in health-related quality of life over three years in a cohort undergoing hemodialysis. *J Am Geriatr Soc* 2008; 56(9):1608-1617.

Ware JE. (1997) SF-36 Health Survey. Manual and Interpretation Guide. The Health Institute, New England Medical Centre. Boston, MA. Nimrod Press. Second Edition.

Ware JE. SF-36 Health Survey. Manual and Interpretation Guide. Boston, MA: The Health Institute, New England Medical Centre, 1997. 2nd edition.

Weisbord SD, Fried LF, Arnold RM, Rotondi AJ, Fine MJ, Levenson DJ et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *Journal of Pain and Symptom Management* 2004; 27(3):226-240.

Welch JL. Development of the Thirst Distress Scale. *Nephrology Nursing Journal* 2002; 29(4):337-343

Weisbord S, Carmody S, Bruns F, Rotondi A, CohenL, Zeidel M and Arnold M. Symptom burden, quality of life, advance care planning and the potential value of palliative care in severely ill haemodialysis patients *Nephrol Dial Transplant* (2003) 18: 1345–1352

Weisbord S, Fried L, Mor K, Resnick A, Unruh M, Palevsky P, Levenson D, Cooksey S, Fine M, Kimmel P, and Arnold R. Renal Provider Recognition of Symptoms in Patients on Maintenance. Hemodialysis *Clin J Am Soc Nephrol* 2: 960-967, 2007

Winearls CG, Glassock RJ. Dissecting and refining the staging of chronic kidney disease. Kidney International (2009) 75, 1009-1014

Wight JP, Edwards L, Brazier JE, Walters SJ, Payne JN, Brown CB. The SF-36 as an outcome measure of services for end-stage renal failure. *Quality in Health-Care* 1998; 7(4):209-221.

Wu AW, Fink NE, Cagney KA, Bass EB, Rubin HR, Meyer KB et al. Developing a health-related quality of life measure for end-stage renal disease: the CHOICE Health Experience Questionnaire. *American Journal of Kidney Diseases* 2001; 37(1):11-21.

Wu AW, Fink NE, Marsh-Manzi JVR, Meyer KB, Finkelstein FO, Chapman MM et al. Changes in quality of life during hemodialysis and peritoneal dialysis treatment: generic and disease specific measures. *Journal of the American Society of Nephrology* 2004; 15(3):743-753.

Zarifian A. Symptom occurrence, symptom distress, and quality of life in renal transplant recipients. *Nephrology Nursing Journal* 2006; 33(6):609-619.