USING PATIENT-REPORTED OUTCOME MEASURES (PROMS) TO PROMOTE QUALITY OF CARE AND SAFETY IN THE MANAGEMENT OF PATIENTS WITH ADVANCED CHRONIC KIDNEY DISEASE

by

OLALEKAN LEE AIYEGBUSI

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Centre for Patient-Reported Outcome Research (CPROR)
Institute of Applied Health Research
College of Medical and Dental Sciences
University of Birmingham

2018

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

The research reported in this thesis explored the use of patient-reported outcome measures (PROMs) in the routine clinical management of patients with advanced chronic kidney disease (CKD). Mixed-methods were employed to (i) evaluate the evidence to support the use of PROMs in patients with advanced CKD, (ii) explore the opinions of patients and clinicians on the use of electronic PROMs (ePROMs) in routine renal practice, and (iii) develop and test the usability of a renal ePROM prototype with patients.

A narrative review, a systematic review and analyses of the Renal Impairment In Secondary Care (RIISC) dataset were conducted to evaluate the evidence to support the use of PROMs in patient with advanced CKD. The narrative review found that the use of ePROMs may provide early warnings of deterioration in patients; reduce the need for patient follow-up visits by up to 50%; reduce hospitalization rates, and improve health-related quality of life (HRQOL) and survival. The systematic review found evidence to support the use of the KDQOL-36 PROM in pre-dialysis patients (although there is a need for further validation within the UK population); evidence was found to support the use of the KDQOL-36 and KDQOL-SF in patients on dialysis; and the ESRD-SCLTM was recommended for kidney transplant patients. Survival analyses using the RIISC dataset found that, along with other clinical and laboratory variables, patient reports of problems on the EQ-5D dimensions of self-care and usual activities significantly increased the risk for all-cause mortality. The likelihood of reporting problems with self-care and usual activities increased significantly with unemployment, retirement, co-morbidities.

A qualitative study was conducted to explore the opinions of patients and clinicians on the use of ePROMs. The qualitative study found that patients may be willing to

complete ePROMs on a regular basis and the selected questionnaires may be perceived as acceptable and relevant. Clinicians expressed a keen interest in PROMs and in PROMs research and acknowledged that PROMs have a potentially important role to play in the future care of patients with CKD. However, clinicians expressed concerns that the use of an ePROM system could raise patient expectations to unrealistic levels, medicalise ‘normal’ psychological/emotional responses to illness, and expose clinicians to the risk of litigation. The usability study demonstrated that patients with advanced CKD may find a Renal ePROM system easy to use and acceptable for reporting their symptoms remotely.

In conclusion, the research reported in the thesis demonstrated the existence of substantial evidence to support the introduction of PROMs in routine renal practice. The thesis also highlighted the need for more real world evidence of the impact of using ePROMs on patient outcomes in order to facilitate adoption of ePROM systems by patients and clinicians.

Dedication

To Mum, Dad, Femo, Fumzi, Bubbles and Roger

How wonderful life is with you in my world

Acknowledgements

Immense thanks to my supervisors, Professor Melanie Calvert, Dr Derek Kyte, Professor Paul Cockwell and Professor Tom Marshall, for the giving me the opportunity to work on this PhD project. I truly appreciate your thoughtful and honest guidance every step of the way.

Thank you to the Health Foundation for funding my PhD (PhD Awards for Improvement Science, reference: 7452) and organising the fantastic summer schools and networking events.

Thanks to Drs Anita Slade, Adrian Gheorghe, Tom Keeley, Anthony Fenton, Christel McMullan and Susan Bayliss for their contributions at various stages of the research.

Special thanks to Mrs Mary Dutton and Mrs Natalie Walmsley-Allen from the Renal Unit, University Hospitals Birmingham NHS Foundation Trust, for going the extra mile to secure R &D approvals and assist with participant recruitment. Thanks to Michelle Grinham for her assistance with room bookings at the Centre for Rare Diseases for the usability test sessions; to Anita Walker, Karen Biddle and Kevin Shepherd for their willingness to answer my numerous queries and sort out all things administrative. Thanks to Samantha Cruz Rivera, Khalid Ahmed, Ameeta Retzer, Janet Jones, Grace Turner, Fatima Isa, Nicola Anderson and Harjeet Bachu for their advice and moral support.

Thanks to all the members of the Patient Advisory Group (PAG), Centre for Patient-Reported Outcomes Research (CPROR), University of Birmingham, for their comments and suggestions throughout the project.

Authorship statement

The research for this thesis was conducted by me (OLA), with guidance from my supervisors: Prof. Melanie Calvert (MC), Dr Derek Kyte (DK), Prof. Paul Cockwell (PC) and Prof. Tom Marshall (TM). Specific contributions to each article are detailed below:

Narrative review paper (Chapter 1) – OLA drafted the manuscript. The manuscript was reviewed, edited and critiqued for intellectual content by MC, DK, PC, OLA and Nicola Anderson (NA).

Systematic review protocol (Chapter 2) – The study was conceived and designed by MC, DK, PC, TM and OLA. The search strategy was developed by OLA, MC, DK, PC and TM. Susan Bayliss assisted with the development of the search strategy for the systematic review. The selection criteria were developed by OLA, MC, DK, PC, TM, Drs Thomas Keeley (TK) and Adrian Gheorghe (AG). OLA piloted the search strategy on the databases. OLA drafted the protocol manuscript. The manuscript was reviewed, edited and critiqued for intellectual content by all authors.

Mixed-methods protocol (Chapter 2) – The study was designed by MC, DK, PC, TM, OLA, Mary Dutton (MD) and Dr Anita Slade (AS). OLA drafted the protocol manuscript. The manuscript was reviewed, edited and critiqued for intellectual content by MC, DK, PC, TM, OLA, MD, AS and members of the patient advisory group (PAG).

Systematic review paper (Chapter 3) - The study was conceived and designed by MC, DK, PC, TM and OLA. AG and TK were involved in the screening and selection of eligible studies. AS contributed to the COSMIN analysis. OLA drafted the initial manuscript. The manuscript was reviewed, edited and critiqued for intellectual content by all authors.

RIISC data analysis paper (Chapter 4) – The study was conceived and designed by MC, DK, PC, TM and OLA. OLA conducted the statistical analysis assisted by Dr Anthony Fenton (AF). OLA drafted the manuscript which was reviewed, edited and critiqued for intellectual content by MC, DK, PC, TM, OLA, AF and Dr Stephanie Stringer (SS).

Qualitative paper (Chapter 5) – The study was conceived and designed by MC, DK, PC, TM and OLA. Dr Christel McMullan (CM) reviewed the methodology. MD and NWA assisted with patient recruitment. OLA conducted all the interviews. The focus group session was moderated by AS. OLA coded and analysed all the transcripts. OLA drafted the initial manuscript which was reviewed, edited and critiqued for intellectual content by all the authors.

Usability testing paper (Chapter 6) – The study was conceived and designed by MC, DK, PC, TM and OLA. Ram Auti (RA) developed the ePROM system to the specifications of DK and OLA. MD and NWA assisted with patient recruitments. OLA conducted all the test sessions. OLA drafted the initial manuscript which was reviewed, edited and critiqued for intellectual content by all the authors.

Table of Contents

Chapter 1: Introduction 1

1.1 Background 2

1.2 Chronic kidney disease 2

1.2.1 Definition and prevalence 2

1.2.2 Staging 3

1.3 Patient-reported outcomes 3

1.3.1 Definition 3

1.3.2 Types of PROMs 4

1.3.3 Electronic PROMs 5

1.4 Applications of PROMs 5

1.4.1 Established applications of PROMs 5

1.4.2 The use of PROMs in clinical practice 6

1.5 Justification for research 7

1.6 Thesis aims and structure 8

1.6.1 Research Questions 8

1.6.2 Aims 9

1.6.3 Structure 9

Patient-centred approach to measuring quality in kidney care (Publication 1) 17

Chapter 2: Methods 26

2.1 Background 27

2.2 Definitions of research paradigm 27

2.3 Paradigms and research approaches 28

2.4 Research approach for this research 29

2.5 Research methods 30

2.6 Research designs 36

Measurement properties of patient-reported outcome measures (PROMs) in CKD: a systematic review protocol (Publication 2) 41

Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety: a mixed-methods project protocol (Publication 3) 47

Chapter 3: Psychometric properties of PROMs used in adult patients with chronic kidney disease - systematic review results (Publication 4) 60

Chapter 4: Relationship between health-related quality of life, clinical outcomes of CKD and clinical variables – analysis of the RIISC dataset 90

Chapter 5: Patient and clinician perspectives on the use of electronic patient-reported outcome measures (ePROMs): Qualitative study (Publication 5) 126

Reflexivity 162

Interaction with patients 162

Interaction with clinicians 163

Chapter 6: Development and usability testing of a renal ePROM prototype (Publication 6) 167

Chapter 7: Discussion 178

7.1 Summary of principal findings 179

7.1.1. Evidence to support the use of PROMs in patients with CKD 179

7.1.2. Opinions of patients and clinicians on the use of ePROMs in routine renal practice 181

7.1.3. Usability testing of a renal ePROM prototype 182

7.2 Interpretation and implications of findings 182

7.2.1. Assessment of evidence supporting the use of PROMs 182

7.2.2 Implementation of an ePROM intervention 184

7.2.3. Development of an ePROM intervention - Input from stakeholders 187

7.3 Strengths and limitations 191

7.4 Future research 195

7.4.1. Development of ePROM systems 195

7.4.2. Implementation of ePROMs intervention 197

7.4.3. Real life implementation of ePROMs systems 198

7.5 Conclusions 200

Appendices 208

Appendix 1: Systematic review search strategy 209

Appendix 2: Summary of study results (systematic review) 217

Appendix 3: Characteristics of included studies (systematic review) 232

Appendix 4: Characteristics of included PROMs (systematic review) 246

Appendix 5: Patient saturation table 254

Appendix 6: Clinician saturation table 263

Appendix 7: Supplementary table of quotations 281

Appendix 8: Comments on selected questionnaires 292

Appendix 9: Illustrative quotes for the selected questionnaires 294

Appendix 10: HRA approval letter 299

Appendix 11: Patient invitation letter 307

Appendix 12: Patient information sheet (Qualitative study) 308

Appendix 13: Clinician information sheet (Qualitative study) 315

Appendix 14: Patient consent form (Qualitative study) 322

Appendix 15: Topic guide for patient interviews 324

Appendix 16: Topic guide for clinician focus group/interviews 326

Appendix 17: Review of frameworks for research impact (Co-authored publication) 327

List of Tables	Page
Chapter 1 - Narrative review	
Description of patient-reported outcome measures	21
Description of patient-reported experience measures	22
Chapter 2 - Mixed-methods protocol	
Description of questionnaires	49
Definitions of domains, measurement properties and aspects of measurement properties	54
Chapter 3 - Systematic review	
Quality criteria for measurement properties	67
Levels of evidence for the quality of the measurement property	68
Description of PROMs evaluated	70
Methodological quality of included studies	74
Evidence synthesis of PROMs used in CKD	76
Chapter 4 - RIISC analysis	
Baseline demographic, clinical and laboratory data	102
Univariable survival analyses for composite outcome	105
Multivariable survival analyses for composite outcome	106
Univariable survival analyses for all-cause mortality	108
Multivariable survival analyses for all-cause mortality	109
Univariable survival analyses for progression to ESRD	111
Multivariable survival analyses (Cox and Competing risk) for progression to ESRD	112
Logistic regression analyses for Self-care	115
Logistic regression analyses for Usual Activities	117
Chapter 5 - Qualitative study	
Patient interview guide	152
Clinician interview guide	152
Patient characteristics	153
Clinician characteristics	154
Description of interviews and saturation tables	155
Illustrative Quotations by Theme and sub-themes	156
Chapter 6 - Usability testing	
Patient demographics	173
Time requirements and error information	173
Usability and satisfaction with Renal ePROM	174

Participants' comments and interviewer's observations	174
List of tasks	176

List of Figures	Page
Chapter 1 - Narrative review	
Interrelationships between PROMs, PREMs, elements of high-quality care, and PCC	19
Chapter 2 - mixed-methods protocol	
Screenshot of the myhealth@QEHB login page	50
Chapter 3 - Systematic review	
Flow diagram	69
Chapter 5 - Qualitative study	
Thematic schema	160
Chapter 6 - Usability testing	
Screenshot of electronic KDQOL-36 questionnaire	172
Screenshot of the progress buttons	172
Screenshot of the myhealth@QEHB login page	172

List of abbreviations

AE	Adverse Events
ASyMS	Advanced Symptom Management System
BDI	Beck Depression Inventory
BMI	Body Mass Index
CAD	Coronary Artery Disease
CAPD	Continuous Ambulatory Peritoneal dialysis
CAT	Computer Adaptive Testing
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	Chief Investigator
CINAHL+	Cumulative Index to Nursing and Allied Health Literature
CKD	Chronic Kidney Disease
COPD	Chronic Obstructive Pulmonary Disease
COREQ	Consolidated criteria for REporting Qualitative research
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
CRF	Chronic Renal Failure
CVI	Content Validity Index
DM	Diabetes Mellitus
EFA	Exploratory Factor Analysis
eGFR	estimated Glomerular Filtration Rate
EMBASE	Excerpta Medica Database
ePROMs	Electronic Patient-Reported Outcome Measures
eRAPID	Electronic patient self-Reporting of Adverse-events: Patient

	Information and aDvice
ESRD	End Stage Renal Disease
EQ-5D	EuroQol Five Dimension
GCP	Good Clinical Practice
GFI	Goodness-of-Fit Index
HD	Haemodialysis
HRQOL	Health-Related Quality of Life
ICC	Internal Correlation Coefficients
IHD	Ischemic Heart Disease
IMD	Index of Multiple Deprivation
KDCS	Kidney Disease Composite Summary
KDIGO	Kidney Disease Improving Global Outcomes
KMO	Kaiser-Meyer-Olkin
KPS	Karnofsky Performance Scale
MCS	Mental Composite Summary score
MEDLINE	Medical Literature Analysis and Retrieval System Online
MHC	Mental Health Composite
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
ONS	Office of National Statistics
PCA	Principal Component Analysis
PCS	Physical Composite Summary score
PD	Peritoneal Dialysis
PHC	Physical Health Composite

PI	Principal Investigators
PICs	Participant Identification Centres
PREMs	Patient-Reported Experience Measures
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PROs	Patient-Reported Outcomes
PROMs	Patient-Reported Outcome Measures
PROMIS	Patient Reported Outcomes Measurement Information System
PROSPERO	Prospective register of systematic reviews
PVD	Peripheral Vascular Disease
REC	Research Ethics Committee
RCT	Randomized Controlled Trial
RIISC	Renal Impairment In Secondary Care
RMSEA	Root Mean Square Error of Approximation
RRT	Renal Replacement Therapy
SD	Standard Deviation
SE	Standard error
SES	Socio-Economic Status
SF-12	Short-Form Health Survey 12-item questionnaire
SF-36	Short-Form Health Survey 36-item questionnaire
SIP	Sickness Impact Profile
SOPs	Standard Operating Procedures
STROBE	STrengthening the Reporting of OBservational studies in Epidemiology
UHB	University Hospitals Birmingham
UK	United Kingdom
US	United States

QALY Quality-adjusted life year

Formatting

This thesis was formatted according to the University of Birmingham 'Alternative Format' thesis guidelines which allows the inclusion of published articles as well as chapters formatted for submission to peer-reviewed journals.

The work presented here has been published within six peer-reviewed journals. The publications are:

- **Aiyegbusi, O. L.,** Kyte, D., Cockwell, P., Anderson, N. and Calvert, M. (2017) 'A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures', *Curr Opin Nephrol Hypertens,* 26(6), pp. 442-449.

- **Aiyegbusi, O. L.,** Kyte, D., Cockwell, P., Marshall, T., Keeley, T., Gheorghe, A. and Calvert, M. (2016) 'Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol', *BMJ Open,* 6(10), pp. e012014.

- **Aiyegbusi, O. L.,** Kyte, D., Cockwell, P., Marshall, T., Dutton, M., Slade, A., Marklew, N., Price, G., Verdi, R., Waters, J., Sharpe, K. and Calvert, M. (2017) 'Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol', *BMJ Open,* 7(6).

- **Aiyegbusi, O. L.,** Kyte, D., Cockwell, P., Marshall, T., Gheorghe, A., Keeley, T., Slade, A. and Calvert, M. (2017) 'Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review', *PLOS ONE,* 12(6), pp. e0179733.

- **Aiyegbusi, O.L.,** Kyte, D., Cockwell, P., Marshall, T., Dutton, M., Walmsley-Allen, N., Slade, A., McMullan C. and Calvert, M. Patient and clinician perspectives on electronic patient-reported outcome measures (ePROMs) in the management of advanced CKD: a qualitative study. *American Journal of Kidney Diseases.*

- **Aiyegbusi, O. L.,** Kyte, D., Cockwell, P., Marshall, T., Dutton, M., Walmsley-Allen, N., Auti, R. and Calvert, M. (2018) 'Development and usability testing of an electronic patient-reported outcome measure (ePROM) system for patients with advanced chronic kidney disease', *Computers in Biology and Medicine,* 101, pp. 120-127.

This thesis also includes a chapter prepared for submission to peer-reviewed journals (Chapter 4).

Please note:

- The page numbers for the articles are not included in the thesis pagination sequence.
- To aid clarity, reference lists are included at the end of each chapter and not at the end of the thesis.

Excerpts from the University of Birmingham Alternative Format Thesis Guidelines:

"Since the alternative format thesis includes copies or offprints of journal articles, book chapters etc., which already have page numbers, the pages of the publications themselves will not be included in the pagination sequence of the submission."

"The incorporation of publication-style chapters in the thesis will inevitably lead to some duplication since each publication-style chapter will have self-contained components that will overlap with parts of the other sections of the thesis…"

Full regulations at:

http://www.birmingham.ac.uk/Documents/university/legal/regulations-part7.pdf

https://intranet.birmingham.ac.uk/as/studentservices/graduateschool/documents/public/rsa/alternative-format-thesis-guidelines.pdf

Chapter 1: Introduction

1.1 Background

The PhD research reported in this thesis explores the use of PROMs in the management of patients with advanced CKD in routine renal practice. The studies evaluate the evidence to support the use of PROMs; explores patient and clinician opinions of PROMs; and investigates the usability of a renal ePROM prototype. This chapter introduces and provides the justification for the research.

1.2 Chronic kidney disease

1.2.1 Definition and prevalence

Chronic kidney disease (CKD) is a broad term that describes various disorders that affect kidney structure and function.[1] CKD may be formally defined as the existence of kidney damage and/or an estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m^2 for 3 months or longer, irrespective of cause.[2] CKD has become a global health issue[3] affecting up to 16% of the adult population in the developed world and is associated with increased morbidity and mortality as it progresses to end-stage renal disease (ESRD).[4 5] Patients with CKD often experience multiple 'clusters' of five to seven symptoms[6 7] such as: fatigue, pain, itchiness, sleep difficulties and cognitive impairment.[7 8] This symptom burden may be high irrespective of CKD stage and[6 7] may negatively impact patients' lifestyle and perceived health-related quality of life (HRQOL).[7 9]

There are also major financial healthcare costs associated with the management of CKD.[10 11] Kerr et al. reported that between 2009 and 2010, the total estimated expenditure attributable to CKD was over a billion pounds in England with renal replacement therapy (RRT) accounting for nearly half of these costs.[10 12] In the US,

the Medicare fee-for-service (FFS) spends approximately $44.4 billion annually managing stages 3 and 4 CKD.[13]

1.2.2 Staging

The Kidney Disease Outcomes Quality Initiative[14] (KDOQI) proposed five different stages of classification of CKD based on the levels of eGFR. Stage 1, is normal kidney function with ≥90ml/min/1.73m^2 but evidence of kidney disease from urinalysis, structural anomalies or genetic traits. At stage 2 there is mild reduction in renal function (eGFR 60 – 89ml/min/1.73m^2) and evidence of kidney disease as for stage 1.[14] Kidney function is moderately reduced at stage 3a (eGFR 45 - 59 ml/min/1.73m^2) and further reduced at stage 3b (eGFR range of 30 - 44 ml/min/1.73m^2).[14] Kidney function is severely reduced at stage 4 (eGFR 15 - 29 ml/min/1.73m^2). At stage 5, eGFR is less than 15 ml/min/1.73m^2.[14]

The KDOQI staging has been criticised as being too inflexible because it does not consider diversity in renal function (due to age, gender and ethnicity) within the general population.[15-17] It has been revised over the years and currently CKD is classified based on cause, the GFR categories above and albuminuria category.[18]

1.3 Patient-reported outcomes

1.3.1 Definition

As described earlier, patients with CKD often experience impairments of their health-related quality of life (HRQOL) due to their symptom burden.[7 9] It is therefore important that patients' assessment of their HRQOL and/or symptoms is obtained and considered in conjunction with clinician assessments, in order to obtain a more complete picture of the impact of CKD and its treatment on patients.[19 20] This may increase clinician awareness of patients' condition[21] and ensure that aspects of care

valued by patients are prioritised appropriately.[19 20] According to the US Food and Drug Administration (FDA), a patient-reported outcome (PRO) is "…any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else."[22] PROs can be measured using validated questionnaires known as patient-reported outcome measures (PROMs).[12]

1.3.2 Types of PROMs

Broadly speaking, PROMs can be categorized into two groups: generic measures and condition-specific measures. Generic PROMs provide information about overall health status and encompass items relevant to the widest range of patients' conditions and the general population.[23 24] They are used for making comparisons between patients with different medical conditions as well as comparisons between patients and healthy individuals in order to quantify the burden of disease on health-related quality of life (HRQOL) in a population.[25] A commonly used generic PROM is the 36-item Short-Form questionnaire (SF-36) developed by RAND Corporation.[26] While generic PROMs offer the opportunity to compare outcomes across various clinical conditions, the trade-off is a loss of responsiveness and sensitivity in many patient groups.[23] Generic PROMs are therefore less suitable for the measurement of patient outcomes at individual patient level.[23]

Preference-based or utility measures are a subset of generic measures which provide utilities or values pertaining to health and can be used in cost-utility analysis to assist with resource allocation.[27] One of the most commonly used preference-based PROMs is the EuroQol 5-dimension (EQ-5D), which originally had 3 response levels (EQ-5D-3L). Over the years, the measure has been widely translated. Value sets based on preferences directly obtained from representative respondents within

the general population, have been produced for many countries including the UK.[28] The measure has since been revised and now has 5 response levels (EQ-5D-5L). Valuation studies for the EQ-5D-5L are on-going and in the meantime a response mapping 'crosswalk' model has been developed which maps responses to the EQ-5D-5L to the EQ-5D-3L value sets using the crosswalk value sets.[29]

Condition-specific measures focus on a particular health condition (for example, asthma or kidney disease), a patient population (for example, adults), or a specific symptom (for example, fatigue). Due to the inclusion of content specific to particular conditions, they may be more sensitive to specific symptoms and therefore possess greater clinical appeal and increased responsiveness to interventions.[23 30]

1.3.3 Electronic PROMs

Recent technological advances have facilitated the electronic collection of PROMs (ePROMs) using devices such as personal laptops, smartphones and tablets.[31] By removing the need for manual data entry by clinical staff, the use of ePROMs could reduce the level of human error involved in the collection process and lower administrative burden.[32-34] A number of studies have demonstrated that data obtained using paper versions and ePROMs are basically equivalent provided the electronic modifications are minimal.[32 35 36]

1.4 Applications of PROMs

1.4.1 Established applications of PROMs

PROMs were originally designed for use in clinical trials[37] to measure treatment benefit or harm.[38] They were predominantly used as secondary outcome measures to support primary outcome(s) which were often related to clinical or laboratory parameters.[39] For example, in a lung cancer trial, the results of the European

Organization for Research and Treatment Core quality of life questionnaire (EORTC QLQ-C30) were used to demonstrate that the combination of gemcitabine and carboplatin was not only more effective but led to fewer side effects than conventional treatment.[40]

In settings such as palliative care or trials where there are no objective or measurable clinical outcomes to back efficacy claims, such as those for pain relieving drugs, PROs may be used as primary outcomes.[39]

PROMs have also been used for other purposes. Pharmaceutical companies employ PROM data to demonstrate that the value of their products go beyond clinical efficacy and safety thus gaining a competitive edge over competitors. Such data is also used to facilitate product approvals and support product labelling claims for new medicines.[22] PROMs data can also inform clinical guidelines (such as NICE guidelines) and influence changes in healthcare policy.[41]

1.4.2 The use of PROMs in clinical practice

The use of PROMs in routine clinical practice has been relatively scarce. PROM data has been used for hospital performance assessment and for benchmarking purposes.[42 43] However, in recent years, there has been an upsurge in interest in the utilisation of PROMs as part of the clinical management of individual patients.[30 44 45]

A few countries notably the US, Denmark and the UK have developed ePROM systems for use in routine clinical practice.[46-50]

The development and use of an ePROM system is purported to facilitate the remote monitoring of patients' health status providing early warnings of deterioration and the opportunity to intervene earlier and improve clinical outcomes.[51-53] It could foster better patient-clinician communication;[50] provide information on aspects of health

valued by patients and not captured in clinical measures;[54] and inform and promote shared decision making.[30 55 56]

1.5 Justification for research

Over the past two decades, numerous studies have suggested various potential benefits that could be derived from the use of PROMs in routine patient care. However, a number of reviews found modest evidence of PROMs impact in routine clinical practice.[55 57-59] The studies reported evidence of improvements in patient-clinician communication and the detection of psychological problems with the use of PROMs.[55 57-59] Methodological as well as implementation issues were cited as possible reasons for the very little evidence of impact on patient outcomes.[55 57 59] It should be noted that the studies included in these reviews were mostly oncology-related and conducted over a decade ago. A number of implementation guidelines and studies have been published since.[44 60-62] At present, the majority of the evidence to support the use of PROMs in routine practice is derived from oncology[19 46 52 63-65] and there are very few published reviews of PROMs impact on patient outcomes in routine renal practice.

When selecting PROMs for routine practice or research purposes, it is essential that their psychometric properties are considered carefully to ensure that useful data is derived from their use.[66 67] Evidence of psychometric properties of PROMs used in CKD were reported in a systematic review by Gibbons and Fitzpatrick.[24] However, this review was conducted nearly a decade ago.[24] As important methodological advances and development/validation studies have been published since the review's publication, it was considered due for an update.[68-70]

Findings from oncology studies have shown that PROMs results may signal serious adverse events in patients before clinical and/or laboratory data.[19 65] However,

research into the predictive abilities of HRQOL data for CKD outcomes and its relationship with clinical and laboratory parameters are very limited. Evidence of the existence of these relationships may mean that patients at risk of rapid deterioration may be identified before their clinical data is available.

Finally, there is also limited CKD-specific literature on patient and clinician opinions of PROMs/ePROMs and patient experience of using a renal ePROM system. It is important that the specific opinions of patients with CKD and renal clinicians are known as recent research shows that the opinions and use of PROMs depend to a large extent on context of use (such as severity of medical condition, setting and existing patient-clinician relationship) and varies widely between medical specialities.[21]

1.6 Thesis aims and structure

This thesis describes a programme of research which addresses the research gaps highlighted in the previous section through a series of studies described in more detail below.

1.6.1 Research Questions

The following research questions were generated for this PhD project:

(1) how are PROMs currently being used in nephrology and what evidence of utility or impact is available?; (2) what is the quality of the psychometric evidence to support PROMs used in adult patients with CKD?; (3) what is the relationship between patients' HRQOL, clinical variables and the risk of progression to ESRD?; (4) what are the opinions of patients and clinicians on the use of ePROMs in routine renal practice?; and (5) what are the views of patients on a renal ePROM system developed in-house?

1.6.2 Aims

The first aim of the thesis, therefore, was to evaluate the evidence to support the use of PROMs in routine renal practice (research questions 1 – 3). The second aim was to explore the opinions of patients and clinicians from the University Hospitals Birmingham NHS Foundation Trust (UHB) and generate feedback which could inform the development of a renal ePROM system (research question 4). A third aim was to develop and test a renal ePROM prototype with patients and obtain their feedback on its usability (question 5).

1.6.3 Structure

Chapters 1, 3, 4 addresses the first aim of the PhD project, Chapter 5 deals with the second and Chapter 6 addresses the third.

Chapter 1 introduces the thesis, defining and describing CKD and PROMs. It includes a narrative review (published in Current Opinions in Nephrology and Hypertension in 2017) which focuses on the current applications of PROMs in the field of renal medicine and summarises recently published evidence of its impact on processes of care and patient outcomes. Chapter 2 presents the methodology for this PhD research in an initial section on the selection of methods and two protocol papers. One protocol paper is for the systematic review of PROMs used in CKD (published in BMJ Open in 2016) and another relates to qualitative research and usability testing (published in BMJ Open in 2017).

Chapter 3 presents a comprehensive systematic review of PROMs used in CKD. It outlines the psychometric evidence for PROMs used in patients with CKD pre-dialysis, dialysed patients and transplant recipients. This review was published in PLOS One in 2017.

Chapter 4 presents a statistical analysis of the Renal Impairment In Secondary Care (RIISC) database, which focuses on quantifying the relationships between HRQOL (through the collection of EQ-5D-3L), clinical variables and progression to ESRD. This study has been written up for submission to a peer reviewed journal for publication.

Chapter 5 reports the qualitative study of the opinions of patients and clinicians on the use of PROMs in routine renal practice. A paper based on this study is currently under review with a journal.

Chapter 6 presents the results of a usability study conducted with patients to assess the usability and acceptability of a renal ePROM system developed by CPROR and UHB.

Finally, Chapter 7 discusses the principal findings of this PhD research, considers their interpretation and implications, highlights the strengths and limitations of the work and outlines suggestions for future research and final conclusions.

References

1. Levey AS, Coresh J. Chronic kidney disease. *Lancet (London, England).* 2012;379(9811):165-180.
2. Levey AS, de Jong PE, Coresh J, et al. The definition, classification, and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. *Kidney international.* 2011;80(1):17-28.
3. Eckardt K-U, Coresh J, Devuyst O, et al. Evolving importance of kidney disease: from subspecialty to global health burden. *The Lancet*.382(9887):158-169.
4. Matsushita K, van der Velde M, Astor BC, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. *Lancet (London, England).* 2010;375(9731):2073-2081.
5. Nitsch D, Grams M, Sang Y, et al. Associations of estimated glomerular filtration rate and albuminuria with mortality and renal failure by sex: a meta-analysis. *BMJ (Clinical research ed.).* 2013;346:f324.
6. Brown SA, Tyrer FC, Clarke AL, et al. Symptom burden in patients with chronic kidney disease not requiring renal replacement therapy. *Clin Kidney J.* 2017;10(6):788-796.
7. Jablonski A. The multidimensional characteristics of symptoms reported by patients on hemodialysis. *Nephrol Nurs J.* 2007;34(1):29-37; quiz 38.
8. Meyer T, Hostetter, T. Medical Progress: Uremia. *The New England journal of medicine.* 2007;13(357):1316-1325.
9. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *Journal of Renal Care.* 2013;39(3):140-150.
10. Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association.* 2012;27 Suppl 3:iii73-80.
11. *US Renal Data System 2015 Annual Data Report: Epidemiology of Kidney Disease in the United States. Chapter 6: Medicare Expenditures for Persons With CKD.* 2016.
12. Aiyegbusi OL, Kyte D, Cockwell P, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. *PloS one.* 2017;12(6):e0179733.
13. Honeycutt AA, Segel JE, Zhuo X, Hoerger TJ, Imai K, Williams D. Medical Costs of CKD in the Medicare Population. *Journal of the American Society of Nephrology.* 2013;24(9):1478-1483.
14. KDOQI. National Kidney Foundation. K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease: Evaluation, Classification and Stratification. *American journal of kidney diseases : the official journal of the National Kidney Foundation.* 2002;39(S1-S266).
15. Glassock RJ, Winearls C. Screening for CKD with eGFR: doubts and dangers. *Clinical journal of the American Society of Nephrology : CJASN.* 2008;3(5):1563-1568.
16. Lamb EJ, Brettell EA, Cockwell P, et al. The eGFR-C study: accuracy of glomerular filtration rate (GFR) estimation using creatinine and cystatin C and albuminuria for monitoring disease progression in patients with stage 3 chronic kidney disease--prospective longitudinal study in a multiethnic population. *BMC nephrology.* 2014;15:13.
17. Winearls CG, Glassock RJ. Dissecting and refining the staging of chronic kidney disease. *Kidney international.* 2009;75(10):1009-1014.
18. KDIGO. *KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease.* 2013.

19. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. *Journal of the National Cancer Institute.* 2011;103(24):1808-1810.
20. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. *The American journal of managed care.* 2008;14(8):530-539.
21. Greenhalgh J, Gooding K, Gibbons E, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. *Journal of Patient-Reported Outcomes.* 2018;2(1):42.
22. FDA. Patient-reported outcome measures: use in medicinal product development to support labeling claims. *Guidance for industry.* 2009.
23. Bryan S, Davis J, Broesch J, et al. Choosing your partner for the PROM: a review of evidence on patient-reported outcome measures for use in primary and community care. *Healthcare policy = Politiques de sante.* 2014;10(2):38-51.
24. Gibbons E, Fitzpatrick , R. A structured review of patient-reported outcome measures for people with chronic kidney disease. *Department of Public Health University of Oxford.* 2010.
25. Calvert MJ, Freemantle N. Use of health-related quality of life in prescribing research. Part 1: why evaluate health-related quality of life? *Journal of clinical pharmacy and therapeutics.* 2003;28(6):513-521.
26. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical care.* 1992;30(6):473-483.
27. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. *Annual review of public health.* 2000;21:587-611.
28. Dolan P. Modeling valuations for EuroQol health states. *Medical care.* 1997;35(11):1095-1108.
29. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2012;15(5):708-715.
30. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2012;21(8):1305-1314.
31. Jensen RE, Rothrock NE, DeWitt EM, et al. The role of technical advances in the adoption and integration of patient-reported outcomes in clinical care. *Medical care.* 2015;53(2):153-159.
32. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2009;12(4):419-429.
33. Velikova G, Wright EP, Smith AB, et al. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 1999;17(3):998-1007.
34. Lee SJ, Kavanaugh A, Lenert L. Electronic and computer-generated patient questionnaires in standard care. *Best Practice & Research Clinical Rheumatology.* 2007;21(4):637-647.
35. Campbell N, Ali F, Finlay AY, Salek SS. Equivalence of electronic and paper-based patient-reported outcome measures. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2015;24(8):1949-1961.
36. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review.

Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2008;11(2):322-333.
37. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ (Clinical research ed.).* 2002;324(7351):1417.
38. Kyte D, Ives J, Draper H, Calvert M. Current practices in patient-reported outcome (PRO) data collection in clinical trials: a cross-sectional survey of UK trial staff and management. *BMJ Open.* 2016;6(10).
39. Gnanasakthy A, Lewis S, Clark M, Mordin M, DeMuro C. Potential of patient-reported outcomes as nonprimary endpoints in clinical trials. *Health and quality of life outcomes.* 2013;11:83.
40. Lee SM, James LE, Qian W, et al. Comparison of gemcitabine and carboplatin versus cisplatin and etoposide for patients with poor-prognosis small cell lung cancer. *Thorax.* 2009;64(1):75-80.
41. Cleland JGF, Daubert J-C, Erdmann E, et al. The Effect of Cardiac Resynchronization on Morbidity and Mortality in Heart Failure. *New England Journal of Medicine.* 2005;352(15):1539-1549.
42. Black N. Patient reported outcome measures could help transform healthcare. *BMJ (Clinical research ed.).* 2013;346:f167.
43. Devlin N, Appleby J. Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making. *London: Kings Fund and Office of Health Economics.* 2010.
44. Aaronson N, Elliott T, Greenhalgh J, et al. User's Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice. *International Society for Quality of Life Research.* 2015.
45. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2009;18(1):115-123.
46. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. *CA: a cancer journal for clinicians.* 2012;62(5):337-347.
47. Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. *Interactive journal of medical research.* 2014;3(1):e5.
48. Schwartzberg L. Electronic Patient-Reported Outcomes: The Time Is Ripe for Integration Into Patient Care and Clinical Research. *American Society of Clinical Oncology educational book. American Society of Clinical Oncology. Meeting.* 2016;35:e89-96.
49. Krawczyk M, Sawatzky R. Relational use of an electronic quality of life and practice support system in hospital palliative consult care: A pilot study. *Palliat Support Care.* 2018:1-6.
50. Velikova G, Booth L, Smith AB, et al. Measuring Quality of Life in Routine Oncology Practice Improves Communication and Patient Well-Being: A Randomized Controlled Trial. *Journal of Clinical Oncology.* 2004;22(4):714-724.
51. Pittman ZCL, John SG, McIntyre CW. Collection of daily patient reported outcomes is feasible and demonstrates differential patient experience in chronic kidney disease. *Hemodialysis International.* 2017;21(2):265-273.
52. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *Journal of Clinical Oncology.* 2016;34(6):557-565.
53. Aiyegbusi OL, Kyte D, Cockwell P, Anderson N, Calvert M. A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures. *Curr Opin Nephrol Hypertens.* 2017;26(6):442-449.

54. Bren L. 'The importance of patient-reported outcomes…It's all about the patients'. *FDA Consumer.* 2006.
55. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of evaluation in clinical practice.* 1999;5(4):401-416.
56. Higginson IJ, Carr AJ. Measuring quality of life: Using quality of life measures in the clinical setting. *BMJ (Clinical research ed.).* 2001;322(7297):1297-1300.
57. Valderas JM, Kotzeva A, Espallargues M, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2008;17(2):179-193.
58. Espallargues M, Valderas JM, Alonso J. Provision of feedback on perceived health status to health care professionals: a systematic review of its impact. *Medical care.* 2000;38(2):175-186.
59. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *Journal of evaluation in clinical practice.* 2006;12(5):559-568.
60. Antunes B, Harding R, Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: A systematic review of facilitators and barriers. *Palliative Medicine.* 2014;28(2):158-175.
61. Osoba D. Translating the science of patient-reported outcomes assessment into clinical practice. *Journal of the National Cancer Institute. Monographs.* 2007(37):5-11.
62. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *Journal of Comparative Effectiveness Research.* 2016;5(5):507-519.
63. Jensen RE, Snyder CF, Abernethy AP, et al. Review of electronic patient-reported outcomes systems used in cancer clinical care. *Journal of oncology practice.* 2014;10(4):e215-222.
64. Andikyan V, Rezk Y, Einstein MH, et al. A prospective study of the feasibility and acceptability of a Web-based, electronic patient-reported outcome system in assessing patient recovery after major gynecologic cancer surgery. *Gynecologic oncology.* 2012;127(2):273-277.
65. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. *Journal of the National Cancer Institute.* 2009;101(23):1624-1632.
66. Snyder CF, Watson ME, Jackson JD, Cella D, Halyard MY. Patient-reported outcome instrument selection: designing a measurement strategy. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2007;10 Suppl 2:S76-85.
67. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. *Lancet (London, England).* 2009;374(9687):369-370.
68. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2012;21(4):651-657.
69. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2010;19(4):539-549.

70. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology.* 2007;60(1):34-42.

The narrative review that follows summarizes the use of PROMs/ePROMs in patients with CKD. It discusses the use of PROMs in routine renal practice as well as in clinical trials. The review also highlights recent studies with significant findings in terms of impact of ePROMs on patient outcomes and or clinical processes.

This review has been published in Current Opinions in Nephrology and Hypertension (Impact Factor (IF) 3.85) as:

Aiyegbusi OL, Kyte D, Cockwell P, Anderson N, Calvert M. A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures. Curr Opin Nephrol Hypertens. 2017. Epub 2017/08/15. doi: 10.1097/mnh.0000000000000357. PubMed PMID: 28806191.

A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures

Chapter 2: Methods

2.1 Background

The desire to acquire knowledge, to constantly question beliefs, are vital philosophical elements that drive and advance research.[1 2] Exploring the philosophical underpinnings or paradigms of research can inform the selection of appropriate research approaches and assist researchers with the clarification of their research questions.[1 2]

2.2 Definitions of research paradigm

Paradigms can be described as “worldviews or all-encompassing ways of experiencing and thinking about the world, including beliefs about morals, values, and aesthetics”.[3-5] A narrower interpretation provided by Kuhn 1970 focuses on worldviews pertaining to issues within the philosophy of knowledge, i.e. the nature of knowledge and the suitable ways of generating it.[3] This version, which is related to epistemological stance, is the most frequently encountered within social science research.[3] Kuhn also referred to paradigms “as shared beliefs within a community of researchers who share a consensus about which questions are most meaningful and which procedures are most appropriate for answering those questions”.[3] The most specific interpretation of paradigms is that they are model examples or “exemplars for how research is done in a given field.”[3 6]

Morgan viewed these varied interpretations of research paradigms as inter-related and nested within each other.[3] He suggested that researchers use model examples to demonstrate the crucial elements of their field and reflect their shared beliefs about research questions and methods. These shared beliefs are in turn based on epistemological stances which are components of each researcher’s broad worldview.[3]

2.3 Paradigms and research approaches

Guba & Lincoln[7] described five paradigms for social and behavioural research namely; positivism, post-positivism, critical theory, participatory research and constructivism (also referred to as interpretivism).[7] At the extremes of the spectrum are positivism and constructivism with different assumptions about the nature of reality or being (ontology) and its study (epistemology).[8] The positivist paradigm asserts that a single "objective reality exists which is independent of human behaviour"[9] and can be observed empirically and explained with logical reasoning.[10] Interpretivism on the other hand, maintain that multiple realities exist and are socially constructed and constantly changing.[10]

A relatively newer paradigm championed by a number of researchers is pragmatism.[3 11-13] Unlike the older paradigms, pragmatism is less concerned with issues of ontology and epistemology focusing instead on the appropriateness of research question(s). Deweyan pragmatism emphasises human experience and the constant interaction between belief and action which lead to the creation of meaning.[11 14 15] Dewey proposed that beliefs arise from interpretation of prior actions while actions are based on interpretations of existing beliefs.[11 14 15]

Paradigmatic positions are often considered to be strongly associated, although not exclusively,[16] with specific design approaches.[17] Positivism is linked to a quantitative approach while constructivism is associated with a qualitative approach. The pragmatic view is that the choice of approach should be decided by the nature and purpose of research questions, acknowledging that a combination of approaches might be required to adequately address research questions.[17] This pluralist position makes pragmatism the philosophical basis for the mixed-methods approach.[11 14 15]

2.4 Research approach for this research

Research approaches and associated philosophical positions were considered early to determine the most appropriate for this body of research. The varied nature of the research questions outlined in Chapter 1, suggested that a combination of methods would be required to effectively address them. Therefore, a mixed-methods approach, underpinned by the pragmatic paradigm, was chosen for this research work. Mixed-methods research is formally defined as "*the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language in a single study or set of related studies*."[18]

Although researchers have been combining methods for decades, mixed-methods approach only became recognised as a distinct approach within the last two decades.[19] In the field of health outcomes research, mixed-methods approach has been recognised as crucial for instrument development and/or evaluation.[19 20] Quantitative and qualitative methods such as factor analysis and cognitive interviews need to be combined during development or validation work to ensure the reliability and validity of instruments.[19] In addition, the use of mixed-methods research with an emphasis on patient and public involvement (PPI) has been recommended to ensure that HRQOL interventions are scientifically rigorous and relevant and acceptable to patients.[20 21]

By utilising a mixed-methods approach, which incorporated PPI, this PhD research drew on the strengths and minimised the weaknesses of single approach research; obtained a more complete picture of the topic as findings were triangulated and synthesised across methods; and subsequently arrived at stronger and more confident conclusions.[13 22 23] The patient advisory group, consisting of patients with

CKD, helped with the selection of measures for this project and also reviewed study materials.

2.5 Research Methods

As mentioned in the previous section, the choice of methods to employ for the project was governed by the nature of each question. Therefore, a narrative review was conducted to obtain a broad perspective on the use of PROMs in renal medicine and summarise recently published evidence of its impact on outcomes (Chapter 1).

A systematic review, which facilitate the assessment and synthesis of findings across numerous studies whilst minimizing bias,[24] was conducted to evaluate PROMs used in CKD. The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist was used to appraise the methodological quality of the selected studies and facilitate the synthesis of the review findings (Chapters 2 & 3). This systematic review helped identify suitable measures to take forward in the research work and the gaps in evidence that may be addressed in future work.

Quantitative statistical methods including regression models made it possible to explore and quantify the relationships between numeric clinical data such as eGFR and HRQOL (EQ-5D index scores) from the RIISC dataset. These analyses helped identify which factors are predictive of an increased risk of progression to ESRD in patients with advanced CKD (Chapter 4).

Qualitative methods were utilized to explore the opinions of patients and clinicians in order to elicit a deeper understanding of the factors that may influence the adoption of a renal ePROM system (Chapters 2 & 5).[25]

Sampling for the qualitative research was influenced by theoretical and practical considerations. The sampling process was guided by a 4-point approach as

described by Robinson[2] and involves the determination of: (i) sample universe (ii) sample recruitment (iii) sample size (iv) sampling technique.

Sample universe - The sampling universe was defined by the eligibility criteria, which provided a cohesive sample required to address the study aims.[26 27] The mixed-methods protocol included as part of this chapter provides detailed information about the eligibility criteria.[28]

Sample recruitment - Participants were recruited from the Renal Unit at the Queen Elizabeth Hospital Birmingham (QEHB) as the ePROM system is being designed for the management of patients managed by this organization.

Sample size - For the practical purpose of planning the study, sample size requirements were estimated at the design stage based on experience from previous similar qualitative studies conducted by the research team.[28 29] However, these were flexible, as the attainment of data saturation was considered more important.[26]

Sampling technique - Luborsky and Rubinstein[30] describe five sampling techniques namely[30]:

- Convenience sampling – Participants are selected based on the fact that they are easily accessible to the researcher.[30]
- Purposive sampling – Participants are intentionally selected to represent explicitly predefined traits or criteria.[30]
- Snowballing – This technique utilises participants as referral sources for other potentially eligible participants that may be recruited for a study.[30]
- Quota sampling – A predetermined number or percentage of study participants are recruited from a specific sub-group of the sample population.[30]
- Case study – A single individual or institution is selected as the total universe. A variant of case study technique is the key-informant approach where an expert provides knowledgeable information on a specialized topic.[30 31]

Sampling techniques were chosen for the qualitative study based on the theoretical understanding of the research topic.[27] Purposive sampling was employed as it ensured that study participants, whose suitability was determined by the eligibility criteria, were representative of the target populations for the ePROM intervention. This meant that the study explored perspectives that were relevant to the development of the ePROM system.[27] The sampling universe is more defined with purposive sampling than convenience sampling, meaning that recruitment might be more challenging. As the eligibility criteria did not purposively sample for hard-to reach subgroups, this may lead to their underrepresentation within the study sample. Therefore, quota sampling was combined with purposive sampling to ensure that eligible patients from minority ethnic groups participated in the study. The decision to recruit up to 30% of participants from these groups was to ensure that the study reflected the ethnic diversity of the patient catchment area. One notable issue with quota sampling is that if the views of the target sub-group differ substantially from that of the rest of the sample, more participants may be required in order to achieve data saturation.[32] Therefore, the codes generated from the interview transcripts of participants from minority ethnic groups were continually checked to detect any divergence from those obtained for the rest of the sample.

Snowballing was combined with purposive sampling for clinician recruitment in order to allow the quick identification of eligible, information-rich clinician participants for the study. However, selection bias might have occurred as individuals may be more likely to suggest others within their social network who share their personal views.[33] Patient recruitment was conducted by renal research nurses as stipulated by the Research Ethics Committee. Further information about the recruitment process may be found in the published protocol submitted with this chapter.[28]

Three analytical methods were considered for the qualitative study namely: grounded theory[34], qualitative content analysis[35] and thematic analysis.[36] Grounded theory is an inductive method of developing new theories to explain a phenomenon.[34] In grounded theory, the researcher starts with the systematic collection and concurrent analysis of data and then the proposal of a hypothesis or theory rather than starting with a hypothesis or theory to prove or disprove.[37] The emerging theory or hypotheses is therefore 'grounded' in the data and modified by the researcher according to the findings from the data.[38]

Qualitative content analysis focuses on the characteristics of language as a mode of communication and involves the systematic analysis of the content and context of text data (interview transcripts, field notes etc.) for subsequent subjective interpretation.[35] Content analysis is therefore useful for exploring communicative messages.[39] The researcher systematically codes and identifies themes focusing on thematic presentation and frequency of occurrence.[35]

Thematic analysis is a method for identifying, analysing, and reporting themes generated from a data set.[36] Whilst thematic analysis is not considered as a separate analytic method by some authors, who consider it as a process utilised by many methods,[40 41] other researchers have advocated for it to be recognised as a distinct analytic method.[36 42 43] The PhD researcher shares the stance of the latter school of thought.

As the aim of the study was not to generate hypothesis or theories, it was decided that grounded theory was not required. Content analysis was also considered inappropriate as the use of language or the frequency of occurrence of themes was not the focus of the study. Thematic analysis was deemed suitable for this study as it has the potential to facilitate the identification of patterns of meaning or themes

across the data set[36] and generate practical insights which may inform the development of the ePROM system. Furthermore, thematic analysis is not bound to any pre-existing theoretical or epistemological positions[36] and can therefore be used across a range of epistemologies and research questions.[42] This versatility meant that thematic analysis was compatible with the pragmatic approach of this study.[36]

One-to-one interviews are suitable for situations where independent responses to questions are preferred or where participants might be less likely to provide candid views in a focus group setting.[44-49] However, one-to-one interviews may be time-consuming and resource-intensive.[44-49] Semi-structured one-to-one interviews rather than focus groups were conducted with patients, as the research team felt some might be reluctant to share personal views and information or find it distressing discussing their condition in a group setting.[45-49] In addition, the majority of the patients preferred to participate in the study on their clinic days. As patients have different clinic appointments, finding a mutually convenient date and time for a group meeting would have been difficult. A semi-structured format was chosen as it ensured that the key objectives of the study were addressed while providing the opportunity to explore ideas and views that might arise during discussions with participants.[50]

In contrast to one-to-one interviews, there is some evidence that focus groups may provide a more stimulating experience for participants as it provides participants the opportunity to interact, challenge and or concur with the ideas and opinions of others.[51 52] Focus groups may be useful for studies focused on socially marginalized groups[53] and act as a forum for change.[54 55] However, due to conformance or conflict avoidance, participants may modify their opinions as discussions progress making it difficult for researchers to arrive at definite conclusions on the subject matter.[56] For

this reason it has been suggested that focus groups should be supported by other methods.[51]

The general advice in literature is that focus groups should be sufficiently homogeneous to encourage free flowing discussions.[55 57] There was a strong possibility that a joint focus group with patients and clinicians would make it difficult for patients in particular to share their opinions freely. Clinicians (nurses and doctors) were invited for focus groups as it was felt that their interactions might generate informative discussion and rich data.[51 52 58] Follow up semi-structured interviews were planned to further explore the opinions expressed during the focus group discussions and triangulate findings.[51] Conducting individual interviews meant that each clinician had the opportunity to express their own views and not just concur with the general opinions which might occur in the group setting of a focus group.[56] This also allowed clinicians who were unable to join the focus group, due to work commitments, to participate in the study.

Saturation has been defined as the point where 'data adequacy' is achieved or "the point when no new information is obtained from additional qualitative data".[26] Rothman et al., defined saturation in qualitative PRO research as "the point in the data collection process when no new concept-relevant information is being elicited from individual interviews or focus groups, or no new information is deemed missing during cognitive interviewing".[59 60]

The adequacy of sample size for qualitative studies is linked to data saturation. In other words, when saturation is achieved, the sample size is adequate. For practical reasons, such as the application for ethical approval, an estimate of sample size is required at the beginning of a study. However, the point at which saturation is achieved during a study determines its final sample size.

A number of factors such as the nature of the topic, study scope and design, the richness of the data provided by each participant,[26 32 61] and the experience and skill of the researcher[32] may influence the number of interviews/focus groups required before saturation is achieved.

For the qualitative study (Chapter 5), data collection and analysis continued until data saturation was achieved for both participant groups.[32 62] Saturation tables, which document the initial and subsequent occurrences of codes, were used to assess data saturation.[59] These tables were examined at intervals during the data collection and analysis in order to detect the point at which data saturation was achieved. Determining the point of data saturation was important as it meant that no new information would be obtained from further interviews and so the interview process could be stopped.

Usability testing methods were employed for the testing of an initial version of the renal ePROM system with patients (Chapters 2 & 6). These usability techniques made it possible to assess first-hand the acceptability of the system and identify technical and non-technical issues which needed rectification.[63-66]

2.6 Research designs

The research designs for the systematic review (Chapter 3), the qualitative study (Chapter 5) and the usability testing (Chapter 6) were published in two protocols:

(a) The protocol outlining the research design for the systematic review has been published in BMJ Open (Impact Factor (IF) 2.37) as:

Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Keeley T, Gheorghe A, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol. BMJ Open.

2016;6(10):e012014. Epub 2016/10/14. doi: 10.1136/bmjopen-2016-012014. PubMed PMID: 27733411.

(b) The protocol outlining the research designs for both the qualitative study and usability testing has been published in BMJ Open (Impact Factor (IF) 2.37) as:

Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Dutton M, Slade A, et al. Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol. BMJ Open. 2017;7(6).

Please note that the research design for the RIISC analysis is reported in the paper included as Chapter 4.

References

1. Easterby-Smith M, Thorpe R, Lowe A. *Management Research: an Introduction*. London: Sage, 1991.
2. Proctor S. Linking philosophy and method in the research process: the case for realism. Nurse researcher 1998;**5**(4):73-90.
3. Morgan DL. Paradigms Lost and Pragmatism Regained. Journal of Mixed Methods Research 2007;**1**(1):48-76.
4. Rossman G, Rallis S. *Learning in the field: An introduction to qualitative research.* 2nd ed. Thousand Oaks, CA: Sage, 2003.
5. Masterman M. The nature of paradigms. In: Lakatos I, Musgrave A, eds. Criticism and the growth of knowledge. New York: Cambridge University Press, 1970.
6. Hanson WE, Creswell JW, Clark VLP, et al. Mixed methods research designs in counseling psychology. Journal of Counseling Psychology 2005;**52**(2):224-35.
7. Guba E, Lincoln Y. Paradigmatic controversies, contradictions, and emerging confluences. In: Denzin N, Lincoln Y, eds. Handbook of qualitative research. 3rd ed. Thousand Oaks, CA: Sage, 2005:191-215.
8. Dewey J. Human nature and conduct. In: Boydston J, Murphy G, eds. The middle works of John Dewey, 1899-1924. Carbondale: Southern Illinois University Press, 2008:1-227.
9. Crossan F. Research philosophy: towards an understanding. Nurse researcher 2003;**11**(1):46-55.
10. Hudson LA, Ozanne JL. Alternative Ways of Seeking Knowledge in Consumer Research. Journal of Consumer Research 1988;**14**(4):508-21.
11. Morgan DL. Pragmatism as a Paradigm for Social Research. Qualitative Inquiry 2014;**20**(8):1045-53.
12. Patton M. Paradigms and pragmatism. In: Fetterman D, ed. Qualitative approaches to evaluation in educational research. Thousand Oaks, CA: SAGE, 1988:116-37.
13. Onwuegbuzie AJ, Leech NL. On Becoming a Pragmatic Researcher: The Importance of Combining Quantitative and Qualitative Research Methodologies. International Journal of Social Research Methodology 2005;**8**(5):375-87.
14. Dewey J. Experience and nature. In: Boydston J, Hook S, eds. The later works of John Dewey, 1925-1953. Carbondale Southern Illinois University Press, 2008:1-437.
15. Dewey J. Reconstruction in philosophy. In: Boydston J, Ross R, eds. The middle works of John Dewey, 1899-1924. Carbondale: Southern Illinois University Press, 2008:77-202.
16. Bryman A. *Quantity and Quality in Social Research*. 2nd ed. London: Routledge, 2004.
17. Creswell JW. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 4th ed: SAGE Publications, Inc, 2013.
18. Collins KMT, Onwuegbuzie AJ, Sutton IL. A model incorporating the rationale and purpose for conducting mixed-methods research in special education and beyond. Learning Disabilities: a Contemporary Journal 2006;**4**(1):67-100.
19. Onwuegbuzie AJ, Bustamante RM, Nelson JA. Mixed Research as a Tool for Developing Quantitative Instruments. Journal of Mixed Methods Research 2009;**4**(1):56-78.
20. Regnault A, Willgoss T, Barbic S, et al. Towards the use of mixed methods inquiry as best practice in health outcomes research. Journal of Patient-Reported Outcomes 2018;**2**(1):19.
21. Wiering B, Boer D, Delnoij D. Patient involvement in the development of patient-reported outcome measures: a scoping review. Health Expect 2016;**20**(1):11-23.
22. Johnson BR, Onwuegbuzie AJ. Mixed Methods Research: A Research Paradigm Whose Time Has Come. Educational Researcher 2004;**33**(7):14-26.
23. Polit DF, Beck C, Hungler BP. *Essentials of Nursing Research: Methods, Appraisal, and Utilization.* 5th ed. Philadelphia: Lippincott Williams & Wilkins, 2002.
24. Mulrow CD. Systematic Reviews: Rationale for systematic reviews. BMJ (Clinical research ed) 1994;**309**(6954):597.

25. Braun V, Clarke V. *Successful qualitative research: A practical guide for beginners*: Sage, 2013.
26. Morse JM. The Significance of Saturation. Qual Health Res 1995;**5**(2):147-49.
27. Robinson OC. Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. Qualitative Research in Psychology 2014;**11**(1):25-41.
28. Aiyegbusi OL, Kyte D, Cockwell P, et al. Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol. BMJ Open 2017;**7**(6):e016687.
29. Morse JM. Determining Sample Size. Qual Health Res 2000;**10**(1):3-5.
30. Luborsky MR, Rubinstein RL. Sampling in Qualitative Research: Rationale, Issues, and Methods. Research on aging 1995;**17**(1):89-113.
31. Patton MQ. *Qualitative Research and Evaluation Methods*. 3rd ed. Thousand Oaks: Sage, 2002.
32. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. Field Methods 2006;**18**(1):59-82.
33. Atkinson R, Flint J. Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies. . Social Research Update Department of Sociology, University of Surrey 2001.
34. Glaser BG, Strauss AL. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine, 1967.
35. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;**15**(9):1277-88.
36. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology 2006;**3**(2):77-101.
37. Corbin J, Strauss A. Grounded theory research: Procedures, canons and evaluative criteria. . Qualitative Sociology 1990;**13**(1):7 - 8.
38. Koerber A, McMichael L. Qualitative Sampling Methods: A Primer for Technical Communicators. Journal of Business and Technical Communication 2008;**22**(4):454-73.
39. Maier MA. Content Analysis: Advantages and Disadvantages. In: Allen M, ed. The SAGE Encyclopedia of Communication Research Methods. Thousand Oaks, California: SAGE Publications, Inc, 2017.
40. Boyatzis R. *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage, 1998.
41. Holloway I, Todres L. The status of method: Flexibility, consistency and coherence. Qualitative Research 2003;**3**(3):345-57.
42. Nowell LS, Norris JM, White DE, et al. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. International Journal of Qualitative Methods 2017;**16**(1):1609406917733847.
43. Thorne S. Data analysis in qualitative research. Evidence Based Nursing 2000;**3**(3):68.
44. Adams W. Conducting Semi-Structured Interviews, 2015.
45. Tonkiss F. Using focus groups. In: Seale C, ed. Researching Society and Culture. 2 ed. London: Sage Publications, 2004.
46. Fielding N, editor. *Interviewing - SAGE Benchmarks in Social Research Methods. No I*: SAGE Publications, 2003.
47. Fielding N, editor. *Interviewing - SAGE Benchmarks in Social Research Methods. No II*: SAGE Publications, 2003.
48. Fielding N, editor. *Interviewing - SAGE Benchmarks in Social Research Methods. No III*: SAGE Publications, 2003.
49. Fielding N, editor. *Interviewing - SAGE Benchmarks in Social Research Methods. No IV*: SAGE Publications, 2003.
50. Madill A. Interviews and interviewing techniques. APA handbook of research methods in psychology, Vol 1: Foundations, planning, measures, and psychometrics. Washington, DC, US: American Psychological Association, 2012:249-75.

51. Kidd PS, Parshall MB. Getting the focus and the group: enhancing analytical rigor in focus group research. Qual Health Res 2000;**10**(3):293-308.
52. Bristol T, Fern EF. Exploring the atmosphere created by focus group interviews: Comparing consumers' feelings across qualitative techniques. Journal of the Market Research Society 1996;**38**(2):185-95.
53. Madriz EI. Using Focus Groups With Lower Socioeconomic Status Latina Women. Qualitative Inquiry 1998;**4**(1):114-28.
54. Race KEH, Hotch DF, Packer T. Rehabilitation Program Evaluation: Use of Focus Groups to Empower Clients. Evaluation Review 1994;**18**(6):730-40.
55. Gibbs A. Focus Groups University of Surrey, Social Research Update 1997(19).
56. Carey MA, Smith MW. Capturing the Group Effect in Focus Groups: A Special Concern in Analysis. Qual Health Res 1994;**4**(1):123-27.
57. Folch-Lyon E, Trost JF. Conducting Focus Group Sessions. Studies in Family Planning 1981;**12**(12):443-49.
58. Rabiee F. Focus-group interview and data analysis. The Proceedings of the Nutrition Society 2004;**63**(4):655-60.
59. Kerr C, Nixon A, Wild D. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. Expert rev 2010;**10**(3):269-81.
60. Rothman M, Burke L, Erickson P, et al. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(8):1075-83.
61. Bowen GA. Naturalistic inquiry and the saturation concept: a research note. Qualitative Research 2008;**8**(1):137-52.
62. Fusch P, Ness L. Are We There Yet? Data Saturation in Qualitative Research. The Qualitative Report 2015;**20**(9):1408-16.
63. Nielsen J. Usability Engineering. Academic Press Inc 1994:165.
64. Industry Usability Reporting. Secondary Industry Usability Reporting. https://www.nist.gov/itl/iad/industry-usability-reporting.
65. Zbrozek A, Hebert J, Gogates G, et al. Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2013;**16**(4):480-9.
66. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(4):419-29.

Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol

BMJ Open

Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol

Olalekan Lee Aiyegbusi,[1,2] Derek Kyte,[1,2] Paul Cockwell,[1,3] Tom Marshall,[1,2] Thomas Keeley,[1,2] Adrian Gheorghe,[4] Melanie Calvert[1,2]

To cite: Aiyegbusi OL, Kyte D, Cockwell P, *et al*. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol. *BMJ Open* 2016;**6**:e012014. doi:10.1136/bmjopen-2016-012014

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2016-012014).

Received 23 March 2016
Revised 25 July 2016
Accepted 21 September 2016



For numbered affiliations see end of article.

Correspondence to
Dr Derek Kyte;
d.g.kyte@bham.ac.uk

ABSTRACT

Introduction: Chronic kidney disease (CKD) is associated with symptoms that can significantly reduce the quality of life (QoL) of patients. Patient-reported outcome measures (PROMs) may facilitate the assessment of the impact of disease and treatment on the QoL, from a patient perspective. PROMs can be used in research and routine clinical practice.

Methods and analysis: A systematic review of studies evaluating the measurement properties of PROMs in adults with CKD will be conducted. MEDLINE, EMBASE, PsycINFO and CINAHL Plus will be systematically searched from inception. Hand searching of reference lists and citations of included studies will be carried out. 2 reviewers will independently screen the titles and abstracts of all the studies retrieved during the systematic search to determine their eligibility. The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist will be used to appraise the methodological quality of the selected studies following the full-text review. Data on the study population, questionnaire characteristics and measurement properties will be extracted from the selected papers. Finally, a narrative synthesis of extracted data will be undertaken.

Ethics and dissemination: Ethical permissions are not required for this study as data from published research articles will be used. Findings will be disseminated through publication in a peer-reviewed journal and presented at conferences. This systematic review will provide a comprehensive assessment of the measurement properties of PROMs currently available for use in adult patients with CKD and present evidence which may inform the selection of measures for use in research and clinical practice.

Trial registration number: CRD42016035554.

Strengths and limitations of this study

- A key strength of this systematic review is that multiple databases will be searched by multiple independent reviewers, with no language or publication date restrictions, thus minimising the risk of selection bias.
- A further strength is that the methodological quality of selected studies will be evaluated using the validated COSMIN checklist.
- This review will include studies relating to chronic kidney disease in adults and exclude studies of all other kidney conditions and children, this decision has been taken to ensure the results are focused on the study research question.

INTRODUCTION

Chronic kidney disease (CKD) is defined as the existence of kidney damage (ie, pathological abnormalities or markers of damage) for 3 months or more; and/or an estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m^2 for 3 months or longer, with or without kidney damage.[1]

CKD affects up to 16% of adults in the UK[2] and is associated with poor outcomes, with a high proportion of patients dying before reaching end-stage renal disease.[3 4] According to the UK Department of Health, the total cost of providing renal care by the National Health Service (NHS) was £1.64 billion in 2009–2010.[5] Aside from these overt costs, the NHS bears other costs accrued from treating associated conditions while patients and caregivers have to contend with possible loss of income.[5]

Patients with CKD often suffer simultaneously from multiple symptoms related to the condition or side effects of their medical treatment.[6] These clusters of symptoms may exert an adverse effect not only on their physical health but also on their psychological and emotional well-being; giving rise

to what is described as a 'symptom burden'.[7] A review by Almutary *et al*[8] identified 30 symptoms associated with CKD, with 5 symptoms (fatigue, drowsiness, pain, pruritus and dry skin) particularly common in all stages.[8] They concluded that the overall symptom burden was high regardless of disease stage.[8] This symptom burden is now acknowledged as the most important predictor of diminished health-related quality of life (HRQoL) in patients with CKD.[9]

In a clinical context, HRQoL refers to the manner in which the physical, emotional and social well-being of an individual is affected by a disease and/or its treatment.[10] HRQoL can be measured using self-administered, validated questionnaires also known as patient-reported outcome measures (PROMs).[11 12]

PROMs have numerous applications in clinical trials and routine clinical settings. They are employed in clinical trials as measures of effectiveness and pharmaceutical companies may use PROM data to support product approvals and labelling claims.[13] PROMs can aid the reporting of serious adverse events due to drug toxicities[14] and trial data can also influence clinical care and health policy.[15 16]

In routine clinical practice, PROMs are mainly used as tools for benchmarking and hospital performance assessment.[17] However, PROMs also have the potential to assist in the delivery of personalised care.[11 17–21]

Currently, primary and secondary care records in the UK are being linked together such that patient information is more readily available to clinicians when required.[22] The integration of patient-reported outcome (PRO) data with other routinely collected clinical data could provide an opportunity for revolutionising the UK healthcare system by facilitating the delivery of stratified medicine, enhancing clinical audits and assisting with the designing of pragmatic trials.[12 21]

Given the vast array of PROMs in existence,[23] the selection of an appropriate measure for any purpose requires the careful consideration of their measurement properties in order to derive any meaningful benefit from their application. It is crucial that such decisions are backed by the best evidence available, preferably from a recent systematic review.[12]

A scoping search identified two relevant systematic reviews namely: a 2010 review by Gibbons and Fitzpatrick and a 2012 review by Wyld *et al*.[24 25] Although both reviews presented evidence of the measurement properties of PROMs in CKD, there is a realistic possibility that more relevant research work has been performed since their publication.[26] In addition, both studies had eligibility criteria which might have excluded potentially relevant studies.[24 25] Gibbons and Fitzpatrick restricted their review to studies published in English language and conducted in English-speaking populations within the UK, North America or Australasia. They also excluded studies with sample sizes <50.[24] Wyld *et al*[25] focused on studies that assessed preference-based (utility) measures excluding studies that evaluated HRQoL measures that did not have a utility component and only included studies conducted in other languages if they provided an English abstract.

Therefore, after considering the issues above, we concluded that a full systematic review is required. The aim of this systematic review will be to evaluate the measurement properties of PROMs currently available for use in adults with CKD.

This review will include studies that assess the measurement properties of generic, utility as well as disease-specific PROM instruments in patients at any stage of CKD. Recent methodological advances and guidelines[27–29] will be employed to ensure that the most up-to-date and robust evidence is obtained.

METHODS AND ANALYSIS

Design

The protocol has been registered with PROSPERO (registration number CRD42016035554). It was developed using the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) checklist.[29]

The review will be conducted and reported in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines.[30]

Search strategy

The following electronic databases will be systematically searched from inception: MEDLINE (Ovid), EMBASE (Ovid), PsycINFO (Ovid) and CINAHL Plus (EBSCO). Literature search results will be uploaded to Endnote X7 (Thomson Reuters). There will be no publication period or language restrictions.[27] The UK Renal Registry will be searched and expert recommendations from members of the review team will be followed up to help identify any additional measures currently under development.

The search strategy for MEDLINE was developed in consultation with an information specialist at the Institute of Applied Health Research, University of Birmingham. Two existing search filters, the sensitivity search filter developed by Terwee *et al*[31] and the Oxford PROM filter,[32] were combined with key terms for renal disease generated by the review team. The MEDLINE search strategy was adapted for use and tested on all the databases (see online supplementary appendix for the full search strategy).

Selection of studies

To be considered for selection, an article must focus on PROMs used specifically for measuring QoL and/or CKD symptoms (the constructs of interest). In addition, the following inclusion and exclusion criteria will be applied.

Inclusion criteria

1. Articles reporting PROM development in all CKD populations.

2. Articles explicitly reporting the assessment of one or more psychometric properties for PROM(s) in all CKD populations.
3. Articles reporting cross-cultural validation of PROMs in all CKD populations.

Exclusion criteria

1. Clinician-assessed instruments.
2. Instrument development studies solely in patients with acute kidney injury.
3. Instruments developed solely for use in patients below 18 years of age.
4. Trials or studies evaluating the effectiveness of interventions where a PROM questionnaire is used as an end point.
5. Editorials, reviews and conference abstracts.

All titles and abstracts will be screened by two independent reviewers (OLA and TK/AG).

Full-text articles will be obtained for studies that potentially meet the eligibility criteria and will again be independently reviewed by the investigators (OLA and TK/AG).

Abstracts that do not provide the reviewers with enough information to make a decision will be taken forward for full-text screening, thus minimising the risk of missing potentially eligible articles.

Reasons for exclusion at the full-text stage of screening will be recorded. At any stage, if the reviewers are unable to reach a consensus, a third reviewer will be consulted (MC/DK). Hand searching of reference lists and citations of included papers will be conducted.

The process of selection will be summarised using a PRISMA flow diagram.

Appraisal of the methodological quality of selected studies

All eligible papers will be independently appraised by two reviewers using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist.[28 33] The COSMIN checklist is a validated critical appraisal tool designed for the systematic evaluation of the methodological quality of studies of the measurement properties of health measurement instruments.[28]

The COSMIN checklist employs a four-point rating scale which allows the rating of items relating to each measurement property as 'excellent', 'good', 'fair' or 'poor' depending on the methodological quality of each study. If a study meets all the requirements for an item, it is rated 'excellent' for that item. Conversely, if a study fails to meet the requirements for an item, it is given a lower rating commensurate to its quality.[28]

The overall quality rating for each measurement property is determined using the 'worst score counts' method.[28] This means that the methodological quality for each measurement property will be determined by taking the lowest rating of its items.

The following measurement properties as defined by Mokkink *et al*[34] will be evaluated:

1. Reliability—'The proportion of the total variance in the measurements which is because of "true" differences among patients'.[34]
2. Internal consistency—'The degree of the interrelatedness among the items'.[34]
3. Measurement error—'The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured'.[34]
4. Content validity—'The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured'.[34]
5. Construct validity—'The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the HR-PRO instrument validly measures the construct to be measured'.[34]
6. Cross-cultural validity—'The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument'.[34]
7. Criterion validity—'The degree to which the scores of an HR-PRO instrument are an adequate reflection of a "gold standard"'.[34] The consensus by the COSMIN panel was that no gold standard exists for PROMs even though some authors consider widely used instruments as 'gold standards'. An exception made by the panel is the comparison of a shortened measure to the original longer version, in which case, the original version can be regarded as the gold standard.[28 33]
8. Responsiveness—'The ability of an HR-PRO instrument to detect change over time in the construct to be measured'.[34]

Data extraction

Data from selected studies will be extracted independently by two reviewers, using a data collection form, with disagreements resolved through discussion and, if necessary, consultation with a third reviewer. Extracted data will be presented in tables. Where appropriate, the results will be presented in separate sections/tables for instruments assessed in predialysis, dialysis and patients who had kidney transplant.

Data on the following will be extracted:

1. characteristics of the study population (including age, gender, ethnicity and stage of CKD);
2. questionnaire characteristics (including name/version, language, scoring method, domains, number of items);
3. evidence regarding the measurement properties of the questionnaire;
4. setting and purpose for which questionnaire is administered, interpretability and operational characteristics such as patient acceptability, and feasibility of administration for staff will also be reported.

Data synthesis

Two sets of criteria will be used to assess the *quality of the measurement properties*:[27 35]

- The quality criteria proposed by Terwee *et al*[35] will be used to rate the *results of studies of measurement properties as* 'positive', 'indeterminate' or 'negative'.[35]
- The modified criteria reported in Terwee 2011[27] (originally proposed by the Cochrane Back Review Group)[36] will be used to synthesis findings on the measurement property of each PROM across studies in order to ascertain the level of evidence for each instrument while taking into consideration the methodological quality of the selected studies. This overall level of evidence will be rated as 'strong', 'moderate', 'limited', 'conflicting' or 'unknown'.[27]

DISSEMINATION

Findings will be disseminated through publication in a peer-reviewed journal and presented at conferences.

DISCUSSION

We acknowledge that studies will evaluate differing measurement properties; therefore, evidence might be limited or unavailable for some measurement properties.

This systematic review will present a comprehensive assessment of the measurement properties of PROMS currently available for use in adult patients with CKD and provide vital evidence which may inform the selection of measures for use in research and clinical practice.

Author affiliations
[1]Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, UK
[2]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[3]University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Birmingham, UK
[4]Oxford Policy Management Ltd, Level 3 Clarendon Centre, Oxford, UK

Acknowledgements The authors thank Susan Bayliss who assisted with the development of the search strategy as detailed in the methods section of the protocol.

Contributors MC is the guarantor. The study was conceived and designed by MC, DK, PC and TM. OLA drafted the protocol manuscript. OLA, MC, DK, PC and TM contributed to the development of the search strategy. The selection criteria were developed by OLA, MC, DK, PC, TM, TK and AG. OLA piloted the search strategy on the databases. The manuscript was reviewed and the final draft approved by all authors.

Funding This project is funded as part of the Health Foundation's Improvement Science Programme.

Disclaimer The Health Foundation is an independent charity working to improve the quality of healthcare in the UK. The Health Foundation is not involved in any other aspect of the project, such as the design of the project's protocol, data collection, analysis and interpretation of results.

Competing interests None declared.

Ethics approval Ethical approval is not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The full search strategy has been submitted as an online supplementary file.

Open Access This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

REFERENCES

1. Levey AS, De Jong PE, Coresh J, *et al*. The definition, classification, and prognosis of chronic kidney disease: a KDIGO Controversies Conference report. *Kidney Int* 2011;80:17–28.
2. Fraser SD, Roderick PJ, Aitken G, *et al*. Chronic kidney disease, albuminuria and socioeconomic status in the Health Surveys for England 2009 and 2010. *J Public Health (Oxf)* 2014;36:577–86.
3. Gansevoort RT, Matsushita K, van der Velde M, *et al*., Chronic Kidney Disease Prognosis Consortium. Lower estimated GFR and higher albuminuria are associated with adverse kidney outcomes. A collaborative meta-analysis of general and high-risk population cohorts. *Kidney Int* 2011;80:93–104.
4. Stringer S, Sharma P, Dutton M, *et al*. The natural history of, and risk factors for, progressive chronic kidney disease (CKD): the Renal Impairment in Secondary care (RIISC) study; rationale and protocol. *BMC Nephrol* 2013;14:95.
5. Kerr M, Bray B, Medcalf J, *et al*. Estimating the financial cost of chronic kidney disease to the NHS in England. *Nephrol Dial Transplant* 2012;27(Suppl 3):iii73–80.
6. Jablonski A. The multidimensional characteristics of symptoms reported by patients on hemodialysis. *Nephrol Nurs J* 2007;34:29–37; quiz 38.
7. Gapstur RL. Symptom burden: a concept analysis and implications for oncology nurses. *Oncol Nurs Forum* 2007;34:673–80.
8. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39:140–50.
9. 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 Int* 2006;69:1621–5.
10. Fairclough D. Design and analysis of quality of life studies in clinical trials. Chapman & Hall/CRC Press, 2002.
11. Calvert MJ, Freemantle N. Use of health-related quality of life in prescribing research. Part 1: why evaluate health-related quality of life? *J Clin Pharm Ther* 2003;28:513–21.
12. Calvert M, Thwaites R, Kyte D, *et al*. Putting patient-reported outcomes on the 'Big Data Road Map'. *J R Soc Med* 2015;108:299–303.
13. Gnanasakthy A, Lewis S, Clark M, *et al*. Potential of patient-reported outcomes as nonprimary endpoints in clinical trials. *Health Qual Life Outcomes* 2013;11:83.
14. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. *J Natl Cancer Inst* 2011;103:1808–10.
15. Calvert M, Kyte D, Duffy H, *et al*. Patient-reported outcome (PRO) assessment in clinical trials: a systematic review of guidance for trial protocol writers. *PLoS ONE* 2014;9:e110216.
16. Kyte DG, Draper H, Ives J, *et al*. Patient reported outcomes (PROs) in clinical trials: is 'in-trial' guidance lacking? a systematic review. *PLoS One* 2013;8:e60684.
17. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167.
18. Moore A. A yes vote for more patient choice. *Health Serv J* 2014;(Suppl):2–5.
19. Devlin NJ, Appleby J. *Getting the most out of PROMs: putting health outcomes at the heart of NHS decision-making*. London: Kings Fund and Office of Health Economics, 2010.
20. FDA. . Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-drugs-gen/documents/document/ucm193282.pdf (accessed Oct 2016).
21. ABPI, Big data road map. Association of the British Pharmaceutical Industry. 2013. http://www.abpi.org.uk/our-work/library/industry/Documents/ABPI%20big%20data%20road%20map.pdf (accessed Oct 2016).
22. Chris Ham CI, Nick G, Anna D, *et al*. Where next for the NHS reforms? The case for integrated care. The King's Fund, 2011.

23. Bryan S, Davis J, Broesch J, *et al*. Choosing your partner for the PROM: a review of evidence on patient-reported outcome measures for use in primary and community care. *Healthc Policy* 2014;10:38–51.
24. Gibbons E, Fitzpatrick R. A structured review of patient-reported outcome measures for people with chronic kidney disease. Department of Public Health University of Oxford, 2010.
25. Wyld M, Morton RL, Hayen A, *et al*. A systematic review and meta-analysis of utility-based quality of life in chronic kidney disease treatments. *PLoS Med* 2012;9:e1001307.
26. Moher D, Tsertsvadze A, Tricco AC, *et al*. When and how to update systematic reviews. *Cochrane Database Syst Rev* 2008(1):Mr000023.
27. Terwee CB. Protocol for systematic reviews of measurement properties. COSMIN, 2011. http://www.cosmin.nl/images/upload/files/Protocol%20klinimetrische%20review%20version%20nov%202011.pdf (accessed Oct 2016).
28. Terwee CB, Mokkink LB, Knol DL, *et al*. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–7.
29. Shamseer L, Moher D, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015;349:g7647.
30. Liberati A, Altman D, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
31. Terwee CB, Jansma EP, Riphagen II, *et al*. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115–23.
32. PROM Group, A.M., Carolina Casañas i Comabella, Monica Hadi, Elizabeth Gibbons, Ray Fitzpatrick, Nia Roberts, PROM Group Construct & Instrument Type Filters. 2010. http://www.cosmin.nl/images/upload/files/PROM%20Gp%20filtersOCTOBER%202010FINAL.pdf (accessed Oct 2016).
33. Mokkink LB, Terwee CB, Patrick DL, *et al*. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49.
34. Mokkink LB, Terwee CB, Patrick DL, *et al*. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45.
35. Terwee CB, Bot SDM, De Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
36. Van Tulder M, Furlan A, Bombardier C, *et al*. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine* 2003;28:1290–9.

Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol

BMJ Open

Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol

Olalekan Lee Aiyegbusi,[1,2] Derek Kyte,[1,2] Paul Cockwell,[1,3] Tom Marshall,[1,2] Mary Dutton,[3] Anita Slade,[1,2] Neil Marklew,[4] Gary Price,[4] Rav Verdi,[4] Judi Waters,[4] Keeley Sharpe,[4] Melanie Calvert[1,2]

To cite: Aiyegbusi OL, Kyte D, Cockwell P, *et al*. Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol. *BMJ Open* 2017;**7**:e016687. doi:10.1136/bmjopen-2017-016687

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-016687).

Received 2 March 2017
Revised 16 May 2017
Accepted 19 May 2017



For numbered affiliations see end of article.

Correspondence to
Derek Kyte; d.g.kyte@bham.ac.uk

ABSTRACT

Introduction Advanced chronic kidney disease (CKD) has a major effect on the quality of life and health status of patients and requires accurate and responsive management. The use of electronic patient-reported outcome measures (ePROMs) could assist patients with advanced pre-dialysis CKD, and the clinicians responsible for their care, by identifying important changes in symptom burden in real time. We report the protocol for 'Using Patient-Reported Outcome measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney Disease' (PRO-trACK) project, which will explore the feasibility and validity of an ePROM system for use in patients with advanced CKD.

Methods and analysis The project will use a mixed-methods approach in three studies: (1) usability testing of the ePROM system involving up to 30 patients and focusing on acceptability and technical performance/stability; (2) ascertaining the views of patient and clinician stakeholders on the optimal use and administration of the CKD ePROM system—this will involve qualitative face-to-face/telephone interviewing with up to 30 patients or until saturation is achieved, focus groups with up to 15 clinical staff, management and IT team members; (3) psychometric assessment of the system, within a cohort of at least 180 patients with advanced CKD, to establish the measurement properties of the ePROM.

Ethics and dissemination This project was approved by the West Midlands Edgbaston Research Ethics Committee (Reference 17/WM/0010) and received Health Research Authority (HRA) approval on 24 February 2017. The findings from this project will be provided to clinicians at the Department of Renal Medicine, Queen Elizabeth Hospitals, Birmingham (QEHB), NHS England, presented at conferences and to the Kidney Patients' Association, British Kidney Patient Association and the British Renal Society. Articles based on the findings will be written and submitted for publication in peer-reviewed journals.

Strengths and limitations of this study

- ► While there is evidence to support the use of electronic patient-reported outcome measures (ePROMs) in the management of other conditions, notably cancer, the evidence for the use of ePROMs in the management of patients with chronic kidney disease (CKD) is currently limited. The PRO-trACK project will help fill this evidence gap.
- ► By using a mixed-methods approach, the project will provide a rigorous exploration of the acceptability, validity and feasibility of the ePROM system for the management of patients with CKD.
- ► This project will only involve patients with CKD stages 4 and 5 and patients on dialysis for ≤6 months. This is because the ePROM system is presently intended for patients with advanced CKD stages 4 and 5 who we hypothesise are likely to derive the most benefit.

INTRODUCTION

Chronic kidney disease (CKD) refers to a number of disorders affecting the structure and function of kidneys.[1] The definition of CKD is based on sustained reduction in renal function (ie, 'estimated glomerular filtration rate (eGFR) <60 mL/min per 1.73 m^2 for 3 months or more') and/or evidence of structural or functional abnormalities of the kidneys regardless of clinical diagnosis.[1 2] CKD is associated with other long-term conditions such as hypertension, cardiovascular diseases and diabetes that will increase the risk of ill health both when present singly and when associated with other long-term conditions.

CKD causes clinical signs and symptoms, particularly when the disease is advanced.[3] The most commonly experienced are fatigue,

drowsiness, pain, pruritus and dry skin.[4] These symptoms often occur concurrently and may negatively affect patients' daily activities and their physical, emotional and psychological well-being,[5] therefore impacting on the quality of life (QOL) of those affected, particularly as the disease progresses towards end-stage renal disease.[6 7]

The symptoms of CKD progression can be monitored using self-completed questionnaires known as patient-reported outcome measures (PROMs), which capture information about health status from patients' own point of view.[8] Although commonly administered in paper format, PROMs can be completed as electronic patient-reported outcome measures (ePROMs) using multiple digital platforms. This makes it possible to remotely monitor patients and generate 'real time' data about patient symptoms and QOL. As patients with advanced CKD are at risk of deteriorating rapidly and developing cardiovascular complications,[9] the use of an ePROM system may help clinicians detect deterioration of symptoms and assist with the tailoring of treatment to the needs of each patient.[10–12] Health-related issues that matter to patients may also be identified using ePROM data, and this could potentially facilitate communication and shared decision-making between patients and their clinicians.[13–15] In stable patients, the use of ePROMs may reduce the occurrence of unnecessary clinical appointments.[12]

In Denmark, the WestChronic ePROM System has been successfully implemented for tailoring the care of various patient groups,[12] while in the UK, patients with cancer have been successfully monitored for the side effects of chemotherapy using the ePROM Advanced Symptom Management System.[16] However, there is limited information on the use of ePROMs in the management of adult patients with CKD in a routine clinical setting. Therefore, the aim of the project is to explore the feasibility and validity of an ePROM system for monitoring and assisting with the individual management of patients with advanced CKD.

Questionnaire selection and the ePROM system

Selection of measures was informed by (1) a systematic review of measurement properties of PROMs used in patients with CKD[17] and (2) feedback from the patient advisory group (PAG). The systematic review found evidence to support the use of the Kidney Disease Quality of Life-Short Form (KDQOL-SF) and Kidney Disease Quality of Life-36 (KDQOL-36). However, these two measures were validated by very few studies in our population of interest (stages 4 and 5 CKD). The review also identified the Integrated Patient Outcome Scale-Renal (IPOS-Renal), which is currently undergoing validation through use in a number of renal units in the UK.

A PAG met prior to commencing the project and considered the acceptability, burdensomeness and relevance of the KDQOL-SF, KDQOL-36 and the IPOS-Renal for the target CKD group. The PAG members chose the KDQOL-36 and IPOS-Renal as they were brief and easy to understand. Therefore, the decision was made to validate the electronic versions of these in the pre-dialysis population (stages 4 and 5). The EuroQol Five-Dimension Five-Level (EQ-5D-5L) questionnaire will be used as a comparison measure for this validation study. See table 1 for a brief description of these three measures.

Table 1 Description of questionnaires

Measure	Description
Kidney Disease Quality of Life-36 (KDQOL-36)	A 36-item health-related quality of life (HRQOL) measure designed for use in patients with kidney disease undergoing dialysis. Derived from the Kidney Disease Quality of Life-Short Form.[88] There are three disease-specific dimensions, namely (1) symptoms and problems (six items), (2) burden of kidney disease (four items) and (3) effects of kidney disease (eight items). It also includes two summary scales derived from the generic Short Form (SF)-12, namely (1) the physical component summary (six items) and (2) the mental component summary (six items).[89] Response options vary for the items and range from 2 to 6. Questions 1, 8 and 12–36 have five response options; questions 2 and 3 have three response options; questions 4–7 have two response options; and questions 9–11 have six response options each. Total and subscale scores may be calculated using the KDQOL-36 scoring program. The raw scores for each item are converted linearly to a 0 to 100 range with higher scores indicating better HRQOL.[88]
Integrated Patient Outcome Scale-Renal	A symptom-specific measure with 11 questions for use with patients with advanced kidney disease to assess their care needs. The questions relate to common symptoms patients with renal disease experience plus additional items such as information needs, practical issues and family anxiety.[90] The first question has a free text response format. Questions 2 to 9 have five response options while questions 10 and 11 have three response options each. The measure is currently being validated by researchers at the Department of Palliative Care, Policy and Rehabilitation at King's College London. Dimensions yet to be ascertained.
EuroQol Five-Dimension Five-Level questionnaire	A generic utility measure with a self-classifier and a visual analogue scale, which can be used to measure health status.[91 92] The self-classifier includes five dimensions: (1) mobility, (2) self-care, (3) usual activities, (4) pain/discomfort and (5) anxiety/depression. This version of the measure has five levels of severity (response options) for each dimension. It is possible to describe 3125 different health states between 0 (dead) and 1 (perfect health).[91 92]

The KDQOL-36 and the IPOS-Renal are free to use without charge as long as the developers are appropriately acknowledged and cited. The EQ-5D-5L requires prior written consent and payment of licensing fees (if applicable). A licence will be obtained for this project.

The PRO-trACK project will consist of three studies, namely (1) usability testing with patients, (2) qualitative study with patients and clinicians, and (3) validation study with patients.

While the usability testing and qualitative interviews are related, they are distinct studies. The usability testing will focus on the actual experience of patients as they test the ePROM system while the qualitative interviews will explore the broader opinions of patients on the use of ePROMs in the NHS.

The content validation of the KDQOL-36 and the IPOS-Renal by patients and clinicians during the qualitative study as well as the results of the validation study will assist the research team with the final decision on which measure to take forward for the final version of the ePROM system.

The ePROM system will be designed as an electronic method of allowing patients with CKD to remotely self-report their symptoms and QOL using a digital platform that is convenient to them (PC, tablet, smartphone, telephone voice recognition or scanned paper copy), providing important patient-centred data to the patients' clinical team.

The ePROM system will be accessed via the secure electronic patient portal developed by the University Hospitals Birmingham NHS Foundation Trust, known as 'myhealth@QEHB' (see figure 1).[18] myhealth@QEHB currently has 14000 patient users and was awarded the prestigious E- Health Insider award in 2014.[19] Around 1200 patients with renal disease are currently signed up for myhealth@QEHB. At the moment, this is a voluntary system.

METHODS AND ANALYSIS

Project design

In this section, the project setting and eligibility criteria for participants will be described first as this will be the same for the three studies. Aspects of research methods specific to each study will be subsequently discussed separately.

Project setting

Queen Elizabeth Hospitals, Birmingham (QEHB) will be the host site for this project. Clinical staff at the Renal Unit, QEHB and academic researchers at the Centre for Patient-Reported Outcomes Research, University of Birmingham will be responsible for the conduct and management of the project. The nephrology service comprises 21 consultants, 30 junior doctors, and 20 nurses and allied health professionals in the CKD team.

Project participants

Patient participants

For the project, we will recruit adult patients with advanced CKD stages 4 and 5 under the care of the renal services at QEHB, with eGFR $\leq$20 mL/min/1.73 m^2 who have been counselled about treatment modalities for end-stage renal disease (ESRD). In addition, using the renal risk calculator, they must have a >20% projected risk of progressing to ESRD and requiring renal replacement therapy (RRT) or an eGFR <10 mL/min/1.73 m^2 within 2 years.[20] The renal risk calculator is a model designed to use routinely collected laboratory tests to predict the progression of patients with CKD stages 3 to 5 to kidney failure.[20]

We have selected this group of patients as our main target for this project because, even though they have not yet reached ESRD, they are likely to have high symptom burden and a high risk of rapid clinical deterioration to renal failure. We hypothesise this group of patients are likely to derive the most benefit from an ePROM system.



Figure 1 Screenshot of the myhealth@QEHB login page.

The research team also is working on a related project focused on dialysis patients.

Patients who have commenced dialysis within 6 months will also be eligible to participate as we hypothesise they will be able to recall their symptoms and medical needs pre-dialysis. Participants will be required to converse in everyday English and provide informed consent.

Patients who have a recent history of acute kidney injury within the last 3 months, a comorbidity with a high level of symptoms or terminal illness likely to lead to the death within 6 months of participation will be excluded from the study.

We will aim to recruit different sets of eligible patients for each study in order to minimise participant burden. However, if patients voluntarily express an interest, they will be allowed to participate in more than one study as long as the renal team is satisfied with their health status.

Efforts will be made to recruit up to 30% of the study participants from minority ethnic groups to reflect the ethnic diversity of the patient catchment area.

Although our previous work does not show an influence of socioeconomic status on outcomes for CKD,[21] we will be mindful of sample diversity in relation to sociodemographic variables such age, gender, ethnicity and other relevant sociodemographic factors. We will collect data on participant characteristics to monitor this as recruitment and qualitative data collection progress. Although we appreciate that patient populations and research samples do not always represent such diversity, we will try to employ recruitment strategies that optimise our ability to recruit a diverse patient sample.

Clinicians and other professional staff

Clinicians who manage patients with CKD at the Renal Unit, QEHB and members of the myHealth team and hospital management staff who provide consent will be recruited for this project.

Recruitment methods

A member of the renal research team at QEHB will screen patients for eligibility using the electronic screening tools that are used for clinical purposes. This will identify patients who meet the eligibility criteria and the clinics they attend. Eligibility will be confirmed by direct review of the clinical records by the research nurse and a clinician who are members of the renal care team, on the delegated duty log.

Patients will then be approached at clinic by a member of the renal care team and given patient information sheet to read. Further information about the study will be given and their immediate queries will be addressed. They will be contacted no earlier than 48 hours after the clinic visit to ascertain if they wish to participate in the study.

Clinicians and other professional staff will be recruited by the members of the research team and sent information sheets via email.

Study 1: usability testing

Usability testing will be conducted to evaluate the acceptability of the ePROM system.

Study objectives

- To determine whether patients can easily navigate the ePROM system.
- To determine if patients are able to complete the questionnaires successfully on their own and, if not, how much assistance they require.
- To determine the average length of time required to complete an ePROM report.
- To determine the level of satisfaction with the ePROM interface.
- To identify changes that might be required to improve user performance and satisfaction.

Data collection

Usability testing refers to the appraisal of a product or service by potential service users and involves the observation of such users completing a task within a predetermined scenario.[22 23]

The scenario for this study will be the self-report of a patient's health status between clinic appointments using an electronic device such as a smartphone, tablet or PC/laptop. Each patient will undergo a single one-to-one session with OLA, the project's chief investigator (CI), and attempt to complete the three electronic questionnaires with as little assistance as possible.

The Concurrent Think Aloud (CTA) and Retrospective Probing (RP) moderating techniques will be used for this study.[24] CTA involves the thinking aloud and vocalisation of participants' thoughts during the session while RP refers to interviewing the participant following the completion of the session.[24] The advantage of combining both techniques is that real-time feedback could be obtained for exploration by the CI afterwards.[22]

The CI will take detailed notes of the participants' comments, actions, non-verbal cues and errors and pass minimal comments to encourage them to think aloud during the sessions.

Qualitative data will be collected in the form of a brief audio-recorded interview at the end of each test. The patients will be questioned based on the notes taken by the CI during the session. They will be encouraged to provide any recommendation to improve user experience.

Data analysis

Thematic analysis will be conducted on the qualitative data (see the analysis of patient interviews for more details), which will include the notes taken during the sessions and the transcripts of the post-test interview.

Quantitative data will be summarised using descriptive statistics such as proportions, averages, percentages and rates.[22] Quantitative data will include successful completion rates, error-free rates and average time required for completion.

Sample size

We will recruit up to 30 patients from QEHB for this study based on the recommendations found in literature.[25] The process of improving the usability of any system is an iterative one[22]; therefore, a minimum of two testing cycles will be conducted with the patients. The findings from the first test cycle will guide the process of improving the ePROM system before the second cycle is conducted.

Study 2: qualitative study

This study will explore the views of stakeholders on the optimal use and administration of the CKD ePROM system.

Study 2a: patient interviews

Study objectives

- To determine which symptoms patients with CKD find most bothersome.
- To explore how acceptable ePROMs are.
- To determine how often patients will be willing to complete the ePROM and their preferred method of completion, that is, PC, smartphone, tablet, telephone voice recognition or paper completion.
- To explore the likely factors that may improve or discourage the completion of ePROMs.
- To explore how they would like to receive feedback from the clinical team regarding the ePROMs they provide.

Data collection

Semi-structured face-to-face interviews will be arranged to either coincide with patients' scheduled clinic visit or held on a separate day if preferred. The option of a telephone interview will also be given.

A topic guide will be used to provide a general direction for each interview and ensure that important issues are covered while allowing enough flexibility to capture other relevant themes that may arise during any session.

Interviews will be recorded using an encrypted digital audio recorder and transcribed by a professional transcription company.

Data analysis

The transcripts will be analysed by the CI using the Nvivo 10 software package by QSR International. Thematic analysis of the data will be conducted following the six steps described by Braun and Clarke.[26] The process of analysis will begin with the CI 'actively' reading and engaging with the data set (ie, searching for patterns and meanings). The next phase will be the initial coding of the raw transcript data using the QSR software. Extracts will be coded inclusively (ie, a little surrounding data will be kept to retain contextual meaning).[27]

Phase 3 will involve the analysis and organisation of codes into potential themes. These initial themes will be revised and refined in the fourth phase on two levels. The first is at the level of coded data extracts to ensure they are coherent for each theme. The second level of analysis is to ensure the themes reflect the data set. During this phase, redundant codes and themes may be removed, revised or merged as required. Phase 5 will involve the definition of what each theme is and what it is not, and its importance in relation to the entire data and the research questions. The themes will be considered individually as well as in relation to other themes to ensure overlaps are kept to a minimum. The final phase will be the production of the study report.

The project team (Independent of the CI) will randomly review a sample of transcripts for verification purposes.

Data analysis will be carried out simultaneously with data collection, and both will continue until no new themes emerge from the further analysis, that is, data saturation has been reached.[28]

Respondent validation will be undertaken, whereby a summary of the main points arising from the interview will be sent to each participant for comments.

Sample size

Based on experience from previous similar qualitative studies conducted by the research team, recruitment will continue until a target sample size of between 15 and 30 patient participants is attained or until saturation is achieved.

Study 2b: focus groups with clinicians and other professional staff

Focus group discussions will be held with clinicians who manage patients with CKD at the Renal Unit, QEHB, members of the myHealth team and hospital management staff as required.

Study objectives

- To evaluate and rate the relevance of the items of the ePROM questionnaires with clinical staff (content validation).
- To determine those factors that may improve or discourage the use of ePROM data by clinicians.
- To determine clinicians' preferred method of displaying ePROM data.

Data collection

An independent member of staff will serve as chief moderator and direct the discussions while the CI will act as assistant moderator, making notes and observing the interactions within the groups. The discussions will be allowed to develop with minimum interference following a topic guide to ensure that all the main points are covered. Focus group sessions will be recorded using an encrypted digital audio recorder and transcribed by a professional transcription company.

Face-to-face or telephone interviews with clinicians and other professional staff may be conducted to explore their views on the use of PROMs. These interviews will require a maximum of 1 hour.

Sample size

We will aim to include up to 15 participants in up to two focus groups (seven to eight participants in each group) and if necessary interview the same number of participants.

Data analysis

Thematic analysis will be conducted (see analysis of patient interviews for details).

Study 3: ePROM validation

The purpose of this study is to evaluate the measurement properties of the electronic versions of the KDQOL-36 and the IPOS-Renal against the EQ-5D-5L and clinical data. At the end of this study, the most suitable questionnaire(s) would be taken forward for formal feasibility testing.

Study objectives

- To determine the reliability and validity of the ePROM questionnaires.
- To determine the ability of the ePROM questionnaires to detect change in a patient's health over a period of time.
- To determine which of the two questionnaires is most suitable to take forward for formal feasibility testing.

Data collection

Patients will be registered on 'myhealth@QEHB' in order to access the ePROM system. There will be the option of completing paper versions if preferred. Time for questionnaire completion may be influenced by patients' symptoms but should require no more than 1 hour.

All participants will be asked to complete the two questionnaires three times: at study entry, at 2 weeks after initial completion and at 6 months after initial completion. Completing the questionnaires at these time points will facilitate the comprehensive assessment of psychometric properties.[8] For definitions of measurement properties, see table 2.

Data analysis

Quantitative data will be analysed using statistical software such as STATA. Where appropriate, analysis will be conducted separately for patients with CKD stages 4 and 5 and patients on dialysis.

A disclaimer statement will be included in the patient information sheets for the validation study informing the patients that the questionnaires will not be assessed until the end of the study; therefore, patients should inform their clinician (eg, general practice or renal services) of any healthcare needs for management.

Psychometric evaluation

Classical test theory (CTT) and the Rasch measurement model[29] will be used to evaluate the psychometric properties of the ePROM versions of the KDQOL-36 and IPOS-Renal. CTT is a traditional approach to questionnaire development, which postulates that a person's observed score consists of their true score plus an additional measurement error score, and the underlying assumption is that this relationship is additive.[30 31] This additive model involves the summation of item ratings on a Likert-type scale to obtain a total score; however, the values of the true score and error score cannot be determined and CTT does not describe the hierarchy of the items.[30–32]

Rasch analysis[29] is one method of evaluating measurement tools to ensure they deliver reliable and valid measurement and is increasingly being used in clinical research and practice for refinement and development of PROMs.[33] The Rasch model operationalises axioms of additive conjoint measurement and tests the extent to which PROMs are unidimensional.[34] Fit to the Rasch model is established through a number of fit statistics.[35 36] Analysis is an iterative process identifying and studying anomalies in the data and the extent to which data conforms to the Rasch model. The degree of fit achieved will identify the extent to which KDQOL-36 and IPOS-Renal demonstrate construct validity, unidimensionality and reliability.[37] When data fit the Rasch model, it confirms that the PROMs are unidimensional and summation of scores from the KDQOL-36 and IPOS-Renal is legitimate.[35]

CTT will be used to evaluate the reliability, construct validity and responsiveness while Rasch analysis will be done to complement the CTT assessment of structural validity of the two questionnaires.

CTT methods

Factor analysis

Structural validity

Exploratory factor analysis will be used to evaluate the factor structure of the KDQOL-36 and the IPOS-Renal.[38 39] This will be conducted using principal component analysis (PCA) with orthogonal Varimax rotation of quadrants.[38] Factors will be identified based on the Scree test and the percentage of variance accounted for by a particular factor.[40 41] Eigenvalues measure the amount of variation and factors will be required to have a minimum eigenvalue of 1.0.[40] Subsequently, a confirmatory factor analysis will be conducted to test whether the hypothesised factor models of the questionnaires are supported by actual data.[38 42]

Reliability

Internal consistency

Using baseline scores, Cronbach's alpha[43] will be calculated for the total scale and subscale scores of the KDQOL-36 and IPOS-Renal. Alpha values 0.70–0.90 will be deemed acceptable[44–46] An 'if item deleted' analysis will be conducted to identify whether any items should be dropped from the scale.[47]

Test–retest reliability

The completion of the electronic questionnaires 2 weeks after the initial completion will allow assessment

Table 2 Definitions of domains, measurement properties and aspects of measurement properties

Domain	Measurement property	Aspect of measurement property	Definition
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: for example, using different sets of items from the same HR-PROs (internal consistency), over time (test–retest) by different persons on the same occasion (inter-rater) or by the same persons (ie, raters or responders) on different occasions (intrarater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is because of 'true'* differences among patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses (for instance with regard to internal relationships, relationships to scores of other instruments or differences between relevant groups) based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'
Responsiveness			The ability of an HR-PRO instrument to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability†			The degree to which one can assign qualitative meaning—that is, clinical or commonly understood connotations—to an instrument's quantitative scores or change in scores

Reproduced with permission from Caroline Terwee, COSMIN.
*The word 'true' must be seen in the context of the CTT, which states that any observation is composed of two components—a true score and error associated with the observation. 'True' is the average score that would be obtained if the scale were given an infinite number of times. It refers only to the consistency of the score and not to its accuracy.[12]
†Interpretability is not considered a measurement property but an important characteristic of a measurement instrument.
CTT, classical test theory; HR-PRO, health-related patient-reported outcome.

of the stability of the questionnaires.[48] Intraclass correlation coefficients (ICCs) for agreement will be calculated on subscale and total scores using a two-way random effects model.[49 50] ICC values >0.75 will indicate excellent test–retest reliability, values 0.40–0.75 will be considered good, while values <0.4 will indicate weak agreement.[30 48 51]

Measurement error

As we are not aware of values for measurement error and minimally important clinical change (MIC) for our population of interest, these will be calculated in this study. Measurement error will be calculated using the SE of measurement.[49 52] The MIC will be determined for patients who commence dialysis within the study period using a patient-reported anchor-based method. This is regarded as the ideal method for calculating the MID as it captures the patients' values directly.[53] We will compare changes in measurement scores with a patient-reported global rating of change scale as our reference 'anchor'.[53–55] The measurement error and the MIC will assist with the assessment and interpretation of treatment outcomes and effects.[56]

Construct validity

Convergent validity (hypothesis testing)

We have formulated the following hypotheses to test in order to establish convergent validity. Pearson's or Spearman's correlation coefficients will be calculated for correlations as appropriate. Pearson's or Spearman's correlation coefficients >0.40 will be considered acceptable for scales that are theoretically related.[57 58]

Hypothesis 1

Each item of the KDQOL-36 and the IPOS-Renal will have a positive correlation ≥0.40 with its own hypothesised subscale after correction for overlap.[59] IPOS-Renal items will be correlated with their subscales once these have been established by factor analysis.

Hypothesis 2

The generic (SF-12) and the disease-specific domains of the KDQOL-36 will have weak to moderate positive correlations with each other as they are designed to assess different aspects of health-related quality of life.[60 61]

Hypothesis 3

Each subscale score of the KDQOL-36 will have positive correlations with the overall health rating score (question 1 of KDQOL-36).[62]

Hypothesis 4

The generic (SF-12) subscales, the physical component summary and the mental component summary of the KDQOL-36 will have higher positive correlations with the utility scores of the EQ-5D-5L than the kidney-specific subscales of the KDQOL-36 and the symptom-specific scales of the IPOS-Renal.

Hypothesis 5

Clinical parameters specific to kidney disease such as the eGFR will correlate better with dialysis-targeted dimensions of the KDQOL-36 and the IPOS-Renal than with generic dimensions of KDQOL-36.[63]

Hypothesis 6

The utility scores of the EuroQol Visual Analogue Scale will have a high positive correlation with the overall health rating scores of the KDQOL-36.

Hypothesis 7

The comparisons of the means of the lowest scoring 25th percentile and the higher scoring 75th percentile for each disease-specific subscale of the KDQOL-36 will be statistically significant (p values <0.05, using the Mann-Whitney U test).[64]

Responsiveness

The questionnaires will be administered 6 months after the initial completion in order to assess the ability of the questionnaires in detecting changes in patients' condition. Using Pearson's correlation coefficient, we will test three hypotheses for responsiveness based on changes in scores as recommended by the Consensus-based Standards for the selection of health Measurement INstruments (COSMIN) group.[65]

Hypothesis 1

There will be significant changes in the QOL scores of patients who switch from conservative management to RRT within this period. Therefore, the QOL scores before and after commencing RRT will be compared using the Wilcoxon signed-rank test.[66]

Hypothesis 2

Changes in KDQOL-36 scores for patients who switch to RRT from conservative care will correlate negatively with changes in their creatinine values and correlate positively with changes in residual renal function and serum albumin.[63 65]

Hypothesis 3

There will be positive correlations between the global rating scale and the changes in KDQOL-36 scores for patients who switch to RRT from conservative care.[65]

In addition to these hypotheses, effect sizes (ES) and standardised response mean (SRM) will be calculated for patients with CKD stages 4 and 5 who were initially managed conservatively but progressed to renal failure (on RRT) during the study period.[67 68]

Higher ES or SRM indicate greater responsiveness and values up to 0.2 will be regarded as small, 0.5 moderate and 0.8 as substantial according to Cohen's criteria.[69] Receiver operating characteristic curves will be used to establish a cut point for predicting transition to RRT.

Application of Rasch analysis

The underlying assumption with the Rasch model is that individual items capture a single underlying trait, and therefore the summation of items from the KDQOL-36 or IPOS-Renal forms unidimensional scales. Rasch analysis is an iterative process that identifies and studies anomalies in the data and the extent to which KDQOL-36 or IPOS-Renal data conform to the Rasch model and therefore the extent to which the instrument is unidimensional. Fit will be established using a variety of indicators and fit statistics.[70] The Rasch Unidimensional Measurement Model software (RUMM2030)[71] will be used to analyse KDQOL-36 and IPOS-Renal data.

KDQOL-36 has a mixture of dichotomous and polytomous responses whereas the IPOS-Renal items are all polytomous using a Likert response format (see table 1). Affirmation of response categories by respondents should follow a logical sequence. As their perceived level of health improves/deteriorates, then responses should reflect this by affirming higher or lower scoring categories, moving from a score of 1, to 2, then 3, etc, on any item.[72] Rasch-Andrich thresholds are the points between adjoining categories where the probability of affirming either category is 50/50, when responders' perceived level of health is equidistantly captured by adjoining categories.[73] Where there is agreement with this expected response hierarchy, thresholds appear ordered, and disordered thresholds are observable as a lack of consistency.[73] Disordered thresholds can suggest poorly defined or redundant scoring categories, and therefore conceptual distinctions between categories may be imprecise. Responders then find it difficult to assign a category to their perceived health status or QOL.

Targeting will be established by examining the extent to which distributions of participants perceived QOL/health status and levels of health identified by KDQOL-36 and IPOS-Renal items are analogous. The position of each responder and item on the underlying construct is defined as the person's ability and the item's difficulty.[31] Therefore, responders with low levels of perceived health or QOL should only affirm items and scoring categories which capture low levels of health. The item which captures the average level of ability will be identified as having zero logits by RUMM2030. Therefore, when person and items are appropriately targeted, person mean location scores should approximate the zero value of the item locations. A positive mean value for responders' estimated locations will suggest that responders' average levels of health are higher than the average on the KDQOL-36. Conversely, negative person locations will confirm the opposite to be true.[35]

The person separation index (PSI) uses the logit values to estimate the internal reliability of the KDQOL-36 and IPOS-Renal and is conceptually equivalent to Cronbach's alpha. It identifies the extent to which the instrument is able to discriminate between groups with different health states and the precision of the estimate for each person.[74] Minimum PSI value suggested for group use is 0.70 and for individual use 0.85.[35]

Individual tests of fit for each person and item will reflect the difference between responders observed and expected responses if data fit the Rasch model. RUMM2030 will automatically cluster responders into equivalent size groups (class intervals) according to their overall ability. A number of statistics use these class intervals including χ^2 statistics and residual values.[75] Residuals are summations of individual item or person deviations from expected fit to the Rasch model, standardised as a z-score. Residual scores between ±2.5 indicate adequate fit to the Rasch model.[75]

A χ^2 statistic will compare this difference, with a summed χ^2 for each class interval contributing to the overall χ^2 for that item. The χ^2 for all items will then be summed to demonstrate the overall item-interaction statistic. A non-significant χ^2 interaction statistic will indicate theoretical fit to the Rasch model.[75] A significant χ^2 will indicate the need for further evaluation to establish potential causes of misfit.

Differential item functioning (DIF) is a form of item bias that can affect fit to the model. DIF manifests itself as responses to individual items by sample subgroups (eg, gender or age group) being inconsistent with their overall perceived level of health.[76] DIF will be identified using ANOVA and statistically significant probability ($p<0.05$, or a Bonferroni-corrected level). DIF for gender, age group and kidney disease stage will be examined.

Item independence is an underlying principle of the Rasch model.[37] Response dependency occurs when a person's response to one item determines the response to another item, and therefore responses are not independent of each other.[35] Residual correlation metrics (<0.3) will identify if response dependency is an issue.

Once the 'Rasch factor' is extracted, leftover residuals should not contain any patterns in the data.[35] A PCA of the residuals will detect if multi-dimensionality is an issue.[37] Subsets of items identified by the negative and positive correlations from the PCA will be used to compare estimates of responders' health states on the two subsets. If no significant difference in the estimates is identified using independent t-test, then unidimensionality is assured. Tennant and Conaghan[35] also state that the percentage of tests outside the range of −1.96 to 1.96 should not exceed 5%. Confidence intervals for the binomial test of proportions will be used; this is based on the number of significant tests and the lower bound should overlap the 5% expected value for a non-significant test in order to confirm unidimensionality.[35] If responders' health state estimates are found to be significantly different in more than 5% of cases, this will indicate that the subtests are measuring different but related aspects of health states.[35] Where the scale is being used to measure changes over time, then using different but related subscales might be more appropriate.[75]

Finally, if the data fit the model, patient and item parameter estimates will be positioned on the same log-odds

units (logits) scale, although as independent parameters allowing a linear transformation of the raw scores to be used.[75] Therefore, estimates of a patient's level of health can be derived from the KDQOL-36 and IPOS-Renal data with confidence.

Sample size

There are various schools of thought regarding sample size requirements for validation studies.[77–80] We aim to recruit at least 180 patients based on the recommendations found in contemporary literature outlining psychometric best practice (5 to 10 times the number of variables in any given multivariate statistical model).[77 78]

Reimbursement and withdrawal

Reimbursement

All patients will be reimbursed for their time with a £20 gift card.[81 82] Light refreshments will be provided for focus group participants.

Withdrawal

Patient participants will be informed they have the right to freely withdraw from the study, for any reason, at any time prior to their data being integrated into the data set. They will not be required to supply a reason for their withdrawal and the decision will have no effect on their future medical care.

As focus group participants (clinicians and other professional staff) will be audio recorded as a group, they will be informed it will be impossible to withdraw their data during or after a focus group discussion. They will not be required to supply a reason for their withdrawal before a focus group discussion and the decision will have no effect on their employment.

DISCUSSION

PROMs can be completed electronically, making it possible to remotely monitor symptoms in patients with CKD and generate 'real time' data, which may assist clinicians with the tailoring of treatment to the needs of each patient.[12] The use of ePROMs could potentially foster patient–clinician communication and further support shared-decision between patients with CKD and their clinicians.[13 14] The regular completion of ePROMs may decrease the need for stable patients with acceptable ePROM data to attend clinical appointments, thus sparing them the financial burden and physical stress of travelling. This might free up appointment times for patients who actually need to be seen in clinic.[12] In this manner, the use of PROMs may significantly improve the quality of life of patients with CKD.

At each stage of the project, when necessary, PAG meetings will be convened for their input and drafts of publication manuscripts reviewed by members of the group before submission to journals.

By employing a mixed-methods approach, the PRO-trACK project will provide evidence of the feasibility and validity of the ePROM system in patients with advanced CKD.

ETHICS AND DISSEMINATION

Ethics and data management

This project was approved by the West Midlands Edgbaston Research Ethics Committee (Reference 17/WM/0010) and received Health Research Authority (HRA) Approval on 24 February 2017. It has also been included in the National Institute for Health Research (NIHR) Clinical Research Network (CRN) Portfolio (ID 33117).

Participant data (whether in electronic or paper format) will be acquired, anonymised, transferred and stored according to the Data Protection Act 1998[83]; the Confidentiality—NHS Code of Practice[84]; the Caldicott principles[85]; the University of Birmingham Code of practice for research[86]; and the University of Birmingham Guidance on Out of Hours Activities and Lone Working.[87]

Only members of the research team will have access to the project data. The exception will be permissions given to authorised regulatory personnel in order to conduct audits and inspections on behalf of the ethics committee.

Dissemination

The findings of the project will be provided to the Informatics Team and the Nephrology Unit at the QEHB as required. Participants will be given a summary of the findings, with reference to the full reports if desired.

Research article(s) based on the findings of the studies will be written and submitted for publication to peer-reviewed journals, and all contributors and their contributions to the study will be acknowledged. We will also disseminate our findings at seminars and conferences both nationally and internationally.

Author affiliations
[1]Centre for Patient-Reported Outcomes Research, University of Birmingham, Birmingham, UK
[2]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[3]Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Birmingham, UK
[4]Patient Advisory Group Member, Centre for Patient-Reported Outcomes Research, University of Birmingham, Birmingham, UK

Acknowledgements The authors thank all the members of the Patient Advisory Group (PAG), Centre for Patient-Reported Outcomes Research (CPROR), University of Birmingham, for their comments and suggestions during the development of this project.

Contributors MC is the guarantor for this project. The project was conceived by MC, DK, PC and TM and designed by MC, DK, PC, TM, OLA, MD and AS. OLA drafted the protocol manuscript. The manuscript was reviewed by MC, DK, PC, TM, OLA, MD, AS, NM, GP, RV, JW and KS. The final draft was approved by all authors.

Funding This project is funded as part of the Health Foundation's PhD Awards for Improvement Science. The Health Foundation is an independent charity working to improve the quality of healthcare in the UK. The Health Foundation was not involved in any other aspect of the project. TM is partly funded by the National Institute for Health Research (NIHR) through the Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC/WM). This paper presents independent research and the views expressed in this publication are not necessarily those of the NIHR, the Department of Health, NHS Partner Trusts, University of Birmingham or the CLAHRC/WM Management Group.

Competing interests None declared.

Ethics approval West Midlands Edgbaston Research Ethics Committee (Reference 17/WM/0010).

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

1. Levey AS, Coresh J. Chronic kidney disease. *Lancet* 2012;379:165–80.
2. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002;39:S1–266.
3. Meyer T, Hostetter TT. Medical progress: uremia. *N Engl J Med* 2007;13:1316–25.
4. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39:140–50.
5. Gapstur RL. Symptom burden: a concept analysis and implications for oncology nurses. *Oncol Nurs Forum* 2007;34:673–80.
6. Gansevoort RT, Matsushita K, van der Velde M, *et al*. Lower estimated GFR and higher albuminuria are associated with adverse kidney outcomes. A collaborative meta-analysis of general and high-risk population cohorts. *Kidney Int* 2011;80:93–104.
7. Stringer S, Sharma P, Dutton M, *et al*. The natural history of, and risk factors for, progressive chronic kidney disease (CKD): the Renal Impairment in Secondary care (RIISC) study; rationale and protocol. *BMC Nephrol* 2013;14:95.
8. FDA. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Guidance for industry. 2009. https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf (accessed Feb 2017).
9. Go AS, Chertow GM, Fan D, *et al*. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med* 2004;351:1296–305.
10. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. *J Natl Cancer Inst* 2011;103:1808–10.
11. Bren L. 'The importance of patient-reported outcomes…It's all about the patients' *FDA Consumer*. 2006. https://permanent.access.gpo.gov/lps1609/www.fda.gov/fdac/features/2006/606_patients.html (accessed Feb 2017).
12. Hjollund NH, Larsen LP, Biering K, *et al*. Use of patient-reported outcome (PRO) measures at group and patient levels: experiences from the Generic Integrated PRO System, WestChronic. *Interact J Med Res* 2014;3:e5.
13. Velikova G, Booth L, Smith AB, *et al*. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol* 2004;22:714–24.
14. Gilbody SM, Whitty PM, Grimshaw JM, *et al*. Improving the detection and management of depression in primary care. *Qual Saf Health Care* 2003;12:149–55.
15. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. *CA Cancer J Clin* 2012;62:336–47.
16. McCann L, Maguire R, Miller M, *et al*. Patients' perceptions and experiences of using a mobile phone-based advanced symptom management system (ASyMS) to monitor and manage chemotherapy related toxicity. *Eur J Cancer Care* 2009;18:156–64.
17. Aiyegbusi OL, Kyte D, Cockwell P, *et al*. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol. *BMJ Open* 2016;6:e012014.
18. myhealth @ QEHB University Hospitals Birmingham NHS Foundation Trust. 2016 http://www.uhb.nhs.uk/myhealth-at-QEHB.htm (accessed Feb 2017).
19. EHI 2014 Awards. Digital Health 2014. http://www.ehilive.co.uk/ehi-awards/past-winners/2014-awards/best-use-of-technology-to-share-information-with-patients-and-carers/
20. Tangri N, Stevens LA, Griffith J, *et al*. A predictive model for progression of chronic kidney disease to kidney failure. *JAMA* 2011;305:1553–9.
21. Jesky MD, Dutton M, Dasgupta I, *et al*. Health-related quality of life impacts mortality but not progression to end-stage renal disease in pre-dialysis chronic kidney disease: a prospective observational study. *PLoS One* 2016;11:e0165675.
22. Usability.gov. https://www.usability.gov/ (accessed Feb 2017).
23. Nielsen J. *Usability engineering*. Academic Press Inc, 1994:165.
24. van den Haak M, De Jong M, Jan Schellens P. Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. *Behav Inf Technol* 2003;22:339–51.
25. Terwee CB, Mokkink LB, Knol DL, *et al*. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–7.
26. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.
27. Bryman A. *Social research methods*. Oxford: Oxford University Press, 2001.
28. Fusch Patricia NL. Are we there yet? Data saturation in qualitative research. *The Qualitative Report* 2015;20.
29. Rasch G. *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research, 1960.
30. Nunnally J. *Psychometric theory*. New York: McGraw-Hill, 1978.
31. Prieto L, Alonso J, Lamarca R. Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health Qual Life Outcomes* 2003;1:27.
32. Tennant A, Penta M, Tesio L, *et al*. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care* 2004;42(1 Suppl):I–37.
33. Yorke J, Horton M, Jones PW. A critique of Rasch analysis using the Dyspnoea-12 as an illustrative example. *J Adv Nurs* 2012;68:191–8.
34. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol* 2007;46:1–18.
35. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism* 2007;57:1358–62.
36. Wright BD. Solving measurement problems with the Rasch model. *J Educ Meas* 1977;14:97–116.
37. Smith EV. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 2002;3:205–31.
38. Vet HCWde, Ad?r HJ, Terwee CB, *et al*. Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Quality of Life Research* 2005;14:1203–18. dicussion 19-21, 23-4.
39. Dexter PR, Stump TE, Tierney WM, *et al*. The psychometric properties of the SF-36 health survey among older adults in a clinical setting. *J Clin Geropsychol* 1996;2:225–37.
40. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas* 1960;20:141–51.
41. Cattell RB. The Scree test for the number of factors. *Multivariate Behav Res* 1966;1:245–76.
42. Wolinsky FD, Stump TE. A measurement model of the Medical Outcomes Study 36-Item Short-Form Health Survey in a clinical sample of disadvantaged, older, black, and white men and women. *Med Care* 1996;34:537–48.
43. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
44. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess* 2003;80:99–103.
45. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53–5.
46. Nunnally JC, Bernstein IH. *Psychometric theory*. 3rd edition. New York: McGraw-Hill , 1994.
47. Gliem JA, Gliem RR. *Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales*, 2003.
48. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
49. Terwee CB, Bot SD, de Boer MR, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
50. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996;1:30–46.
51. Bataclan RP, Dial MA. Cultural adaptation and validation of the Filipino version of Kidney Disease Quality of Life–Short Form (KDQOL-SF version 1.3). *Nephrology* 2009;14:663–8.
52. Stratford P. Reliability: consistency or differentiating among subjects? *Phys Ther* 1989;69:299–300.

53. Guyatt GH, Osoba D, Wu AW, Aw W, *et al*. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002;77:371–83.
54. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
55. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009;17:163–70.
56. Johnston BC, Ebrahim S, Carrasco-Labra A, *et al*. Minimally important difference estimates and methods: a protocol. *BMJ Open* 2015;5:e007953.
57. Norman GR, Sd L. *Biostatistics: the bare essentials*. St Louis, MO: :Mosby, 1994.
58. Pakpour AH, Nourozi S, Molsted S, *et al*. Validity and reliability of Short Form-12 questionnaire in Iranian hemodialysis patients. *Iran J Kidney Dis* 2011;5:175–81.
59. Ware JE, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol* 1998;51:945–52.
60. Yang F, Wang VW, Joshi VD, *et al*. Validation of the English version of the Kidney Disease Quality of Life Questionnaire (KDQOL-36) in haemodialysis patients in Singapore. *Patient* 2013;6:135–41.
61. Park HJ, Kim S, Yong JS, *et al*. Reliability and validity of the Korean version of Kidney Disease Quality of Life instrument (KDQOL-SF). *Tohoku J Exp Med* 2007;211:321–9.
62. Tao X, Chow SK, Wong FK. Determining the validity and reliability of the Chinese version of the Kidney Disease Quality of Life Questionnaire (KDQOL-36™). *BMC Nephrol* 2014;15:115.
63. Korevaar JC, Merkus MP, Jansen MA, *et al*. Validation of the KDQOL-SF: a dialysis-targeted health measure. *Qual Life Res* 2002;11:437–47.
64. Thaweethamcharoen T, Srimongkol W, Noparatayaporn P, *et al*. Validity and reliability of KDQOL-36 in Thai kidney disease patient. *Value Health Reg Issues* 2013;2:98–102.
65. Mokkink LB, Terwee CB, Knol DL, *et al*. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010;10:22.
66. Cheung YB, Seow YY, Qu LM, Lm Q, *et al*. Measurement properties of the Chinese version of the Kidney Disease Quality of Life-Short Form (KDQOL-SF™) in end-stage renal disease patients with poor prognosis in Singapore. *J Pain Symptom Manage* 2012;44:923–32.
67. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27(3 Suppl):S178–S189.
68. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;28:632–42.
69. Cohen J. *Statistical power analysis for behavioral sciences*. New York: Academic Press, 1977.
70. Miller KJ, Slade AL, Pallant JF, *et al*. Evaluation of the psychometric properties of the upper limb subscales of the Motor Assessment Scale using a Rasch analysis model. *J Rehabil Med* 2010;42:315–22.
71. RUMM2030 (program). Perth, Australia: RUMM Laboratory Pty Ltd 2010.
72. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002;3:85–106.
73. Andrich D. An expanded derivation of the threshold structure of the polytomous Rasch model that dispels any "threshold disorder controversy" *Educ Psychol Meas* 2013;73:78–124.
74. Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: an introduction and illustrative example. *Int J Nurs Stud* 2009;46:380–93.
75. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol* 2007;46:1–18.
76. Tennant A, Penta M, Tesio L, *et al*. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care* 2004;42:I–37.
77. Tabachnik BJ, Fidel LS. *Using multivariate statistics*. London: Harper & Row , 1993.
78. Terwee CB. Protocol for systematic reviews of measurement properties. *COSMIN* 2011 http://www.cosmin.nl/images/upload/files/Protocol%20klinimetrische%20review%20version%20nov%202011.pdf (accessed Feb 2017).
79. Velicer WF, Fava JL. Effects of variable and subject sampling on factor pattern recovery. *Psychol Methods* 1998;3:231–51.
80. MacCallum RC, Widaman KF, Zhang S, *et al*. Sample size in factor analysis. *Psychol Methods* 1999;4:84–99.
81. INVOLVE. Payment for involvement—a guide for making payments to members of the public actively involved in NHS, public health and social care research. NIHR 2010. accessed Feb 2017http://www.invo.org.uk/wp-content/uploads/2012/11/INVOLVEPayment-Guiderev2012.pdf
82. Draper H, Wilson S, Flanagan S, *et al*. Offering payments, reimbursement and incentives to patients and family doctors to encourage participation in research. *Fam Pract* 2009;26:231–8.
83. Parliament of the United Kingdom of Great Britain and Northern Ireland. *Data Protection Act 1998* 2005;92 http://www.legislation.gov.uk/ukpga/1998/29/pdfs/ukpga_19980029_en.pdf
84. DOH. Confidentiality—NHS Code of Practice. *Department of Health* 2003 https://www.gov.uk/government/publications/confidentiality-nhs-code-of-practice (accessed Feb 2017).
85. DOH. The Caldicott Committee. *Report on the Review of Patient-Identifiable Information Department of Health* 1997 http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationspolicyandGuidance/DH_4068403 (accessed Feb 2017).
86. UoB. Code of practice for research University of Birmingham 2016. 2015 http://www.birmingham.ac.uk/Documents/university/legal/research.pdf
87. UoB. Guidance on out of hours activities and lone working. *University of Birmingham* 2012 https://intranet.birmingham.ac.uk/hr/documents/public/hsu/hsuguidance/31ohalw.pdf (accessed Feb 2017).
88. Chao S, Yen M, Lin TC, *et al*. Psychometric properties of the Kidney Disease Quality of Life-36 Questionnaire (KDQOL-36™). *West J Nurs Res* 2016;38:1067–82.
89. Ricardo AC, Hacker E, Lora CM, *et al*. Validation of the Kidney Disease Quality of Life Short Form 36 (KDQOL-36) US Spanish and English versions in a cohort of Hispanics with chronic kidney disease. *Ethn Dis* 2013;23:202–9.
90. POS. Palliative Care Outcome Scale Website. https://pos-pal.org/ (accessed Feb 2017).
91. Johnson JA, Coons SJ, Ergo A, *et al*. Valuation of EuroQOL (EQ-5D) health states in an adult US sample. *Pharmacoeconomics* 1998;13:421–33.
92. Herdman M, Gudex C, Lloyd A, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.

Chapter 3: Psychometric properties of PROMs used in adult patients with chronic kidney disease - systematic review results

Psychometric properties of PROMs used in adult patients with chronic kidney disease – a systematic review

Selecting appropriate PROMs for any purpose can be quite challenging given the numerous measures available.[1] It is crucial that the psychometric (measurement) properties of PROMs are considered to ensure that they actually measure the intended domains of HRQOL and the data generated serve a useful purpose.[2] It is essential that such decisions are supported by the most recent evidence available. A scoping search found two relevant reviews of PROMs used in patients with CKD which were published over 5 years ago.[3 4] As a number of important methodological guidelines and recommendations[5 6] have been published since, it was necessary to conduct an up-to-date review utilising these newer methodological advances.
One of the methodological guidelines employed for this review was the COSMIN checklist which was designed to assess the quality of the eligible development/validation studies.[6] A quality criteria was subsequently used for evidence synthesis.[5]

References

1. Bryan S, Davis J, Broesch J, et al. Choosing your partner for the PROM: a review of evidence on patient-reported outcome measures for use in primary and community care. Healthcare Policy 2014;**10**(2):38-51.
2. Snyder CF, Watson ME, Jackson JD, et al. Patient-reported outcome instrument selection: designing a measurement strategy. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2007;**10 Suppl 2**:S76-85.
3. Gibbons E, Fitzpatrick , R. A structured review of patient-reported outcome measures for people with chronic kidney disease. Department of Public Health University of Oxford 2010.
4. Wyld M, Morton RL, Hayen A, et al. A systematic review and meta-analysis of utility-based quality of life in chronic kidney disease treatments. PLoS medicine 2012;**9**(9):e1001307.

5. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology 2007;**60**(1):34-42.
6. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2012;**21**(4):651-7.

This Chapter presents the findings of the systematic review which has been published in PLoS One (Impact Factor (IF) 2.76) as:

Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Gheorghe A, Keeley T, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PloS one. 2017;12(6):e0179733. doi: 10.1371/journal.pone.0179733.

The work outlined in this Chapter has been disseminated at the following conferences:

- British Renal Society Conference, Nottingham, UK, April 2017 *(Oral presentation)*
- 2nd PROM Conference, London, UK, June 2017 *(Poster presentation)*
- UK Kidney Week, Liverpool, UK, June 2017 *(Poster presentation)*
- 23rd International Society for Quality of Life Research (ISOQOL) Annual Conference, Copenhagen, Denmark, Oct 2016 *(Poster presentations)*

PLOS ONE

RESEARCH ARTICLE

Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review

Olalekan Lee Aiyegbusi[1,2], Derek Kyte[1,2]*, Paul Cockwell[1,3], Tom Marshall[1,2], Adrian Gheorghe[4], Thomas Keeley[5], Anita Slade[1,2], Melanie Calvert[1,2]

1 Centre for Patient Reported Outcomes Research, University of Birmingham, Edgbaston, Birmingham, United Kingdom, **2** Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham, United Kingdom, **3** Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Edgbaston, Birmingham, United Kingdom, **4** Oxford Policy Management Ltd, Oxford, United Kingdom, **5** PAREXEL International, London, United Kingdom

* d.g.kyte@bham.ac.uk

Check for updates

OPEN ACCESS

Citation: Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Gheorghe A, Keeley T, et al. (2017) Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PLoS ONE 12(6): e0179733. https://doi.org/10.1371/journal.pone.0179733

Editor: Kathrin Eller, Medizinische Universitat Graz, AUSTRIA

Received: April 5, 2017

Accepted: June 2, 2017

Published: June 21, 2017

Copyright: © 2017 Aiyegbusi et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This project was funded as part of the Health Foundation's Improvement Science Programme. Website: http://www.health.org.uk/ OLA is the recipient of the Health Foundation's PhD Awards for Improvement Science (Ref: 7452). The Health Foundation is an independent charity committed to bringing about better health and

Abstract

Background

Patient-reported outcome measures (PROMs) can provide valuable information which may assist with the care of patients with chronic kidney disease (CKD). However, given the large number of measures available, it is unclear which PROMs are suitable for use in research or clinical practice. To address this we comprehensively evaluated studies that assessed the measurement properties of PROMs in adults with CKD.

Methods

Four databases were searched; reference list and citation searching of included studies was also conducted. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist was used to appraise the methodological quality of the included studies and to inform a best evidence synthesis for each PROM.

Results

The search strategy retrieved 3,702 titles/abstracts. After 288 duplicates were removed, 3,414 abstracts were screened and 71 full-text articles were retrieved for further review. Of these, 24 full-text articles were excluded as they did not meet the eligibility criteria. Following reference list and citation searching, 19 articles were retrieved bringing the total number of papers included in the final analysis to 66. There was strong evidence supporting internal consistency and moderate evidence supporting construct validity for the Kidney Disease Quality of Life-36 (KDQOL-36) in pre-dialysis patients. In the dialysis population, the KDQOL-Short Form (KDQOL-SF) had strong evidence for internal consistency and structural validity and moderate evidence for test-retest reliability and construct validity while the KDQOL-36 had moderate evidence of internal consistency, test-retest reliability and construct validity. The End Stage Renal Disease-Symptom Checklist Transplantation Module

health care for people in the UK. The funder did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. AG is employed by Oxford Policy Management Ltd. Oxford Policy Management Ltd provides support in the form of salary for Dr. Gheorghe, but did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. TK is employed by PAREXEL International. PAREXEL International provides support in the form of salary for Dr. Keeley, but did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. MC receives consultancy fees from Ferring and Astellas Pharma. Ferring and Astellas Pharma did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of all the authors are articulated in the 'author contributions' section.

Competing interests: OLA is funded by the Health Foundation's PhD Awards for Improvement Science. DK is funded by the National Institute for Health Research (NIHR) Fellowship. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, Health Education England or the Department of Health. The funder had no role with respect to study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication. TM is partly funded by the NIHR through the Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC-WM). This paper presents independent research and the views expressed in this publication are not necessarily those of the NIHR, the Department of Health, NHS Partner Trusts, University of Birmingham or the CLAHRC-WM Management Group. AG is employed by Oxford Policy Management Ltd. TK is employed by PAREXEL International. MC reports grants from the NIHR, Macmillan Cancer Support, personal fees from Ferring and Astellas Pharma, outside the submitted work; and co-Chairs the International Society for Quality of Life Research Best Practices for PROs in Trials Taskforce. These affiliations do not alter our adherence to PLOS ONE policies on sharing data and materials.

(ESRD-SCLTM) demonstrated strong evidence for internal consistency and moderate evidence for test-retest reliability, structural and construct validity in renal transplant recipients.

Conclusions

We suggest considering the KDQOL-36 for use in pre-dialysis patients; the KDQOL-SF or KDQOL-36 for dialysis patients and the ESRD-SCLTM for use in transplant recipients. However, further research is required to evaluate the measurement error, structural validity, responsiveness and patient acceptability of PROMs used in CKD.

Introduction

Chronic kidney disease (CKD) is a global health issue [1]. It affects up to 16% of the adult population in the developed world and is associated with increased morbidity and mortality that is directly related to severity [2, 3]. CKD also has major healthcare economic costs [4, 5]. For the National Health Service (NHS) in England, the total estimated expenditure attributable to CKD between 2009 and 2010 was over a billion pounds; renal replacement therapy (RRT) accounted for approximately half of this expenditure [6]. In the US, it is estimated that managing stages 3 and 4 CKD cost the Medicare fee-for-service (FFS) approximately $44.4 billion annually [7].

Early symptoms of CKD, such as fatigue, are usually non-specific [8]. However patients with more advanced CKD often report multiple 'clusters' of symptoms including drowsiness, pain, pruritus and dry skin [9]. This overall symptom burden may have a negative impact on the perceived health-related quality of life (HRQOL) of patients with end-stage renal disease [10, 11].

HRQOL can be assessed using self-administered, validated questionnaires known as patient-reported outcome measures (PROMs) [12]. PROMs have a wide variety of applications ranging from clinical trials [13] to product labelling [14] and routine clinical care [15, 16]. There is an increasing awareness that PROMs may have a future role in the management of patients with kidney disease, including integration into routine practice; for example, through monitoring patients for symptoms or changes in HRQOL that may require an intervention [17].

If any benefit is to be derived from the use of PROMs, it is important that they are well validated to ensure that they actually measure what they are supposed to measure, produce consistent results and capture all aspects of the construct(s) under investigation that matter to the target population if any benefit is to be derived from their use [14].

A systematic review by Gibbons and Fitzpatrick [18] evaluated the measurement properties of PROMs used in the CKD population, but this was conducted over six years ago. There have been methodological advances since then [19, 20] and it is reasonable to assume that new research has been published [21]. The review was restricted to studies published in English which might have excluded potentially relevant papers [18]. In addition, the review did not report evaluating the methodological quality of the selected studies. It is vital that the methodological quality of studies evaluating the measurement properties of PROMs is assessed to ensure that conclusions about the reliability and validity of the measures are dependable [22] as these have a potential impact on clinical practice and health policy [23].

Therefore, we have evaluated the methodological quality of the selected studies using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)

checklist [22] and used the findings to inform our evidence synthesis. Thus providing the best evidence possible, to inform the selection of PROMs for monitoring the symptoms of CKD and its treatment effects in pre-dialysis, dialysis and renal transplant patients.

Methods

Design

This systematic review was conducted and reported according to a registered and published protocol (PROSPERO registration number: CRD42016035554) (See S1 Text. Review Protocol) [24] and written in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (See S2 Text. PRISMA Checklist) [25]. We also considered the findings of the review of systematic reviews by Terwee et al. [26].

Search strategy

Relevant databases including MEDLINE (Ovid), EMBASE (Ovid), PsycINFO (Ovid) and CINAHL Plus (EBSCO) were systematically searched from inception to 21st December 2015 without language restrictions [24].

The search strategy was initially developed for MEDLINE and subsequently adapted for the other databases (See S3 Text. Search Strategy). Two existing search filters [27, 28] were combined with key terms generated by the review team for renal disease and its treatment modalities. An information specialist at the Institute of Applied Health Research, University of Birmingham, was consulted during the process.

Search records were downloaded into Endnote X7 and duplicates removed. In addition, the UK Renal Registry website was searched to 17th May 2016.

Screening process

All titles and abstracts were screened independently by two reviewers (OLA and TK/AG). Full-text articles were obtained for studies potentially meeting the eligibility criteria and were independently reviewed by the same reviewers. Reasons for exclusion were documented. Hand searching of reference lists and citation searching of the included papers was also conducted. At all stages, disagreements regarding eligibility were resolved through discussion and, if necessary, consultation with a third reviewer (MC/DK).

Selection of studies

Studies were included if they: (1) focused on PROMs used specifically for measuring HRQOL and/or CKD symptoms (the constructs of interest) in any CKD population; and (2) reported either the development or evaluation of one or more psychometric properties of a PROM [24].

Articles excluded were clinical trial reports, editorials, reviews and conference abstracts. In addition, studies that focussed on clinician-assessed instruments, PROMs developed for use in patients with acute kidney injury or in patients below 18 years of age were excluded.

Data extraction

Data from selected studies were extracted independently by two reviewers (OLA and AS) using a pre-designed data collection form and cross-checked for accuracy.

The following data were extracted where available

1. Characteristics of study populations
2. Questionnaire characteristics

3. Evidence regarding measurement properties as defined by Mokkink et al.[29] namely: reliability (test-retest reliability, internal consistency, measurement error); validity (content validity, construct validity (including hypothesis testing, structural validity and cross-cultural validity); responsiveness of questionnaires to changes over time; the setting and purpose for which the questionnaires were administered and details regarding their interpretability; operational characteristics including patient acceptability and mode and feasibility of administration; and details regarding patient involvement in the PROM development or validation process.

Appraisal of the methodological quality of selected studies

Following the selection process, the quality of the included papers was assessed by two reviewers (OLA/AS) using a validated critical appraisal tool for studies of health measurement instruments: the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist [20, 22]. The COSMIN checklist is designed to evaluate the methodological quality of studies of psychometric properties [22]. It comprises of mini-checklists A to I (otherwise known as 'boxes') which correspond to each measurement property (See Table 1). Some measurement properties are named and defined differently by different authors, therefore the COSMIN definitions [29] were used to ascertain which measurement properties were evaluated by the studies. The COSMIN checklist is intended for use as a modular tool meaning that the mini-checklists (boxes) to be completed for each study will be determined by the measurement properties evaluated by the study [29]. Each mini-checklist has a set of quality items/questions which were rated individually using the COSMIN 4-point scale as 'excellent', 'good', 'fair' or 'poor'.

An item is rated 'excellent' when there is evidence that the methodological quality of the study in relation to the item is adequate [20]. An item is rated 'good' when relevant information is not reported in an article, but it can be assumed that the methodological quality is adequate [20]. An item is rated 'fair' if there is doubt about the adequacy of the study's methodological quality in relation to that item [20]. Finally, an item is rated 'poor' when there is evidence that the methodological quality of the study in relation to that particular item is inadequate [20]. For example, a small sample size was considered poor methodological quality in all the mini-checklists. A sample size $\geq$100 was considered 'excellent', 50–99 'good', 30–49 'fair', and $<$30 'poor' [20].

The 'worst score counts' method was used to determine the methodological quality of each paper per measurement property [20]. This meant taking as the overall score for each measurement property, the lowest rating given to any item within the respective mini-checklist [20].

Reviewers consulted a third author (MC/DK) if they were unable to reach a consensus at any point during the assessment.

Data synthesis

The quality criteria developed by Terwee et al. [30] was used to rate the results for each measurement property per study as either 'positive' (+), 'indeterminate' (?) or 'negative' (-). For example, structural validity was rated as positive, if the factors identified after performing a factor analysis were reported to explain at least 50% of variance. If the factors explained $<$50% of variance structural validity was rated as negative. The indeterminate rating was given if the percentage of variance explained was not reported. (See Table 1) [30].

An evidence synthesis across studies was then conducted for measurement properties reported for each PROM using another set of criteria (See Table 2) [19]. At this stage, the

Table 1. Quality criteria for measurement properties.

Property	Rating †	Quality Criteria
Reliability		
Internal consistency	+	Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha not determined or dimensionality unknown
	-	Cronbach's alpha(s) < 0.70
Reliability	+	ICC / weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80
	?	Neither ICC / weighted Kappa, nor Pearson's r determined
	-	ICC / weighted Kappa < 0.70 OR Pearson's r < 0.80
Measurement error	+	MIC > SDC OR MIC outside the LOA
	?	MIC not defined
	-	MIC ≤ SDC OR MIC equals or inside LOA
Validity		
Content validity	+	All items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement AND the questionnaire is considered to be comprehensive
	?	Not enough information available
	-	Not all items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement OR the questionnaire is considered not to be comprehensive
Structural validity	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	-	Factors explain < 50% of the variance
Hypothesis testing	+	Correlations with instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses AND correlations with related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with instruments measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR correlations with related constructs are lower than with unrelated constructs
Cross-cultural validity	+	No differences in factor structure OR no important DIF between language versions
	?	Multiple group factor analysis not applied AND DIF not assessed
	-	Differences in factor structure OR important DIF between language versions
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥ 0.70
	?	No convincing arguments that gold standard is "gold"
	-	Correlation with gold standard < 0.70
Responsiveness		
Responsiveness	+	Correlation with changes on instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70 AND correlations with changes in related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with changes on instruments measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlations with changes in related constructs are lower than with unrelated constructs

MIC = minimal important change, SDC = smallest detectable change, LoA = limits of agreement

ICC = intraclass correlation coefficient, DIF = differential item functioning, AUC = area under the curve

† + = positive rating,? = indeterminate rating

- = negative rating

(Reproduced with permission from Caroline Terwee, COSMIN)

https://doi.org/10.1371/journal.pone.0179733.t001

overall level of evidence for each PROM was provided by one or more studies, taking into account their methodological quality [20]. The overall level of evidence for each measurement property was graded as 'strong', 'moderate', 'limited' 'unknown' or 'conflicting' [19]. For example, a measurement property was graded as 'strong' if at least one study had 'excellent'

Table 2. Levels of evidence for the quality of the measurement property.[19, 30].

Level of Evidence	Rating†	Criteria
Strong	+++ or —-	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or —	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	+/-	Conflicting findings
unknown	?	Only studies of poor methodological quality

† + = positive rating

? = unknown rating

- = negative rating

(Reproduced with permission from Caroline Terwee, COSMIN)

https://doi.org/10.1371/journal.pone.0179733.t002

methodological quality or at least two studies had 'good' methodological qualities (See Table 2).

Results

The search strategy retrieved 3,702 titles/abstracts. After 288 duplicates were removed, 3,414 abstracts were screened and 71 full-text articles were retrieved for further review. Of these, 24 full-text articles were excluded for various reasons (Fig 1). Following reference list and citation searching, 19 articles were retrieved bringing the total number of papers included in the final analysis to 66. Strength of agreement between the reviewers, calculated using Cohen's Kappa Statistic [31], was good (OLA/AG = 0.889, OLA/TK = 0.863).

Description of tables

Table 1 presents the quality criteria used to evaluate the measurement properties for each PROM while the criteria used for the evidence synthesis are in Table 2. A brief description of the PROMs evaluated in this review is presented in Table 3. The methodological qualities of these studies are summarized in Table 4. Table 5 presents the synthesis of the overall level of evidence for each PROM. The results reported by the included studies were extracted and are summarized in S1 Table. Summary of study results. The characteristics of the included studies are presented in S2 Table. Characteristics of included studies. The PROMs evaluated in this review are summarized in S3 Table. Characteristics of included PROMs. S1–S3 Tables have been submitted as 'supporting information' due to their size.

Evidence synthesis

A total of 25 PROMs were identified from the 66 publications; 20 disease-specific, 3 generic and 2 utility PROMs (See Table 3). As the included studies were conducted in pre-dialysis, dialysis and renal transplant populations, the evidence synthesis is described in 3 corresponding sections (See Table 5). The term 'indeterminate' was used when vital information required to assess a measurement property was missing (See Table 1) while 'unknown' was used for measurement properties that were only assessed by studies of poor methodological quality (See Table 2).

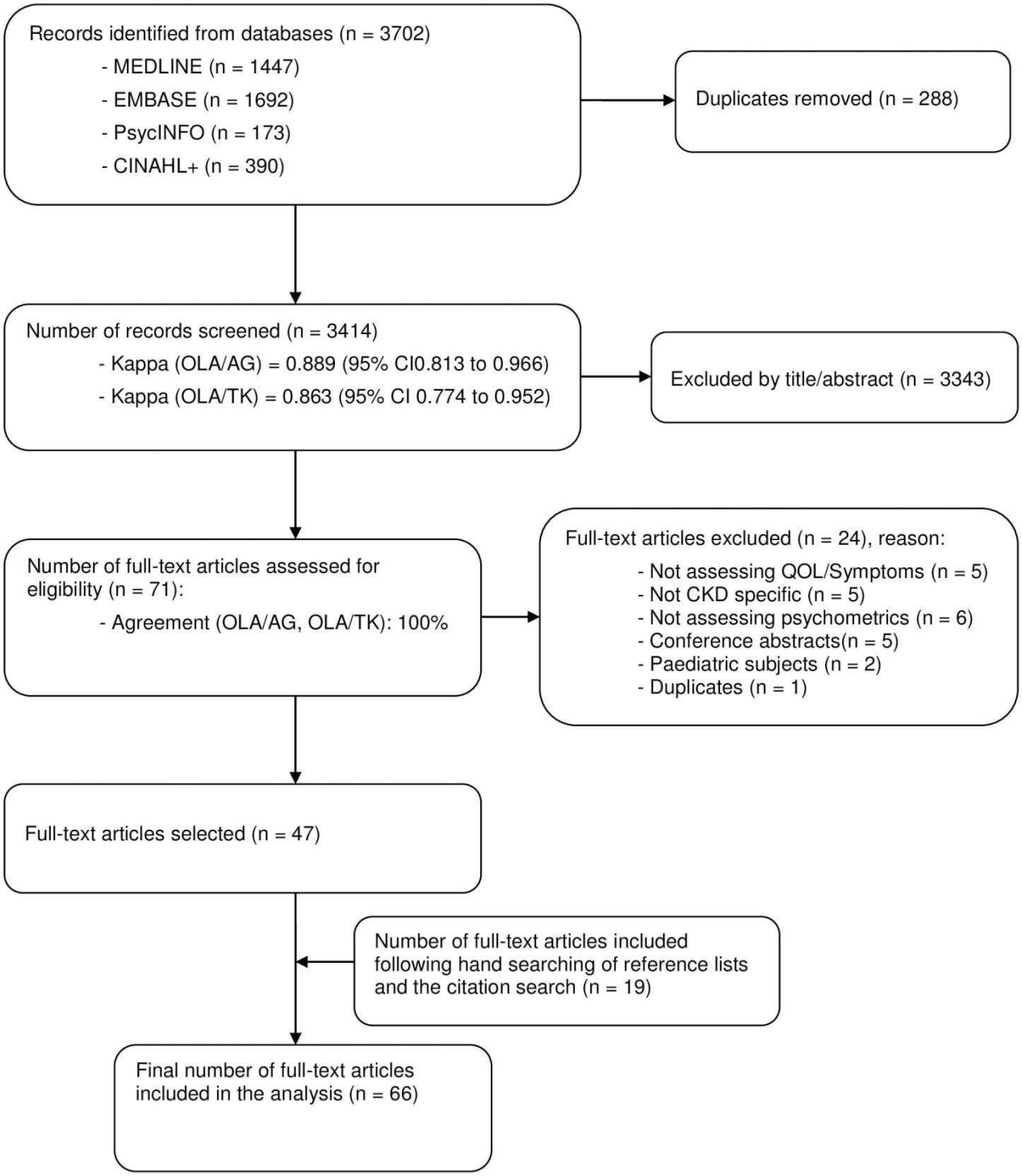


Fig 1. Flow diagram.

https://doi.org/10.1371/journal.pone.0179733.g001

Table 3. Description of PROMs evaluated.

Measure	Description
Disease-specific measures: These measure health in a way that is specific to a particular disease, set of conditions, or part of the body [122].	
Agarwal	A 37-item HRQOL measure for use in non-dialysis patients with mental and physical dimensions [32]. Overall scores range from 0 to 100 with higher scores indicating better HRQOL [32].
Kidney Disease Quality of Life—36 (KDQOL-36)	A 36-item HRQOL measure designed for use in kidney disease patients undergoing dialysis. Derived from the KDQOL-SF [33]. There are 3 specific dimensions namely: symptoms and problems (ii) burden of kidney disease (iii) effects of kidney disease. It also includes two summary scales derived from the SF-12 namely: the physical (PCS) and mental (MCS) scales [34]. Overall scores range from 0 to 100 with higher scores indicating better HRQOL [33].
KDQOL-SF	An 80-item HRQOL measure designed for use in kidney disease patients undergoing dialysis [47]. Derived from the 134-item KDQOL [47, 58]. Version 1.3 differs from version 1.2 by the addition of a screening question for sexual activity [47]. There are 8 generic dimensions from the SF-36 (See below) and 8 disease-specific dimensions namely: (i) symptoms/problems (ii) effects of kidney disease on daily life (iii) burden of kidney disease (iv) work status (v) cognitive function (vi) quality of social interaction (vii) sexual function (viii) sleep. There are 3 additional dimensions namely: (i) social support (ii) dialysis staff encouragement (ii) patient satisfaction. Scores range from 0 to 100 for each dimension and higher scores indicate better HRQOL.
Chinese Dialysis Quality of Life Scale (CDQOL)	A 29-item measure designed to measure the QOL of Chinese dialysis patients. Scored on a 5-point Likert scale. Higher scores indicate better quality of life as perceived by the patient [49].
CHOICE Health Experience Questionnaire (CHEQ)	An 83-item HRQOL measure. Designed to complement the SF-36 and assess the effectiveness of dialysis modalities [50]. Comprises of 8 dimensions from SF36 (see below) combined with 7 supplementary items and 13 specific dimensions namely: (i) freedom (ii) travel restrictions (iii) cognitive functioning (iv) financial (v) restrictions on diet and fluids (vi) recreation (vii) work (viii) body image (ix) symptoms (x) sleep (xi) sexual functioning (xii) access-related problems (xiii) quality of life. An additional 2 person-specific quality of life item. Possible scores range from 0 to 100 with higher scores indicating better HRQOL [50].
Dialysis Symptom Index (DSI)	A 30-item measure designed to assess symptom prevalence and severity in patients on haemodialysis [52]. An overall symptom burden score and a total symptom severity score are calculated. Symptoms are rated on a 5-point Likert scale and total score ranges from 0 to 150 with higher scores indicating greater symptom severity [52].
Modified Edmonton Symptom Assessment System (modified ESAS)	This is a measure of symptom burden for use in dialysis patients. It is a modification of the ESAS. There are 10 symptom-specific items and 10 visual analogue scales with superimposed 0–10 scale [55]. The scale for each symptom is anchored by the words 'No' and 'Severe' at 0 and 10, respectively and the sum of scores range from 0 to 100 with higher scores indicating greater symptom distress and burden [55].

(*Continued*)

Table 3. (*Continued*)

Measure	Description
Kidney Disease Questionnaire (KDQ)	A 26-item measure designed to assess the effects of interventions on the QOL of patients undergoing haemodialysis [56]. There are 5 dimensions namely: (i) physical symptoms (ii) fatigue (iii) depression (iv) relationships with others (v) frustration. Note that the physical symptom dimension is patient-specific, thus the symptoms most important to individual patients are identified and used to evaluate the dimension. Questions are scored on a 7-point Likert scale and higher scores indicate a lower impact of disease on HRQOL [56].
KDQOL	A 134-item QOL measure designed for use in kidney disease patients undergoing dialysis [58]. It consists of SF36 dimensions (see below), 11 kidney disease targeted scales and an item that assesses change in health over a year (overall health rating) [58]. All scale scores are transformed linearly into 0–100 point scales with higher scores indicating better HRQOL [58].
KDQOL (modified)	A 55-item QOL measure derived from the KDQOL [59]. Using affinity mapping, 11 subscales [59] were identified namely: (i) pain (ii) psychological dependency (iii) cognitive functioning (iv) social functioning (v) dialysis-related symptoms (vi) cardiopulmonary symptoms (vii) sleep (viii) energy (ix) cramps (x) diet (xi) appetite. 4 items were ungrouped. The measure is scored on a 0 to 100 scale with higher scores indicating better HRQOL [59].
WHOQOL-BREF (Dialysis)	A 32-item HRQOL measure modified for use in dialysis patients with 4 domains (incorporating 4 dialysis-specific items): (i) physical (ii) psychological (iii) social relationship (iv) environment. The measure also includes two global items (general QOL and general health). A 5-point Likert scale is used and higher scores signify higher HRQOL [63].
Quality of Life Index (QLI) 3.0	A 68-item QOL measure divided into 2 sections. One section measures satisfaction with various domains of life, while the second measures the importance of the domain to the individual [64]. There are 4 domains: (i) health and functioning (ii) Social and economic (iii) psychological/spiritual (iv) family. Each section has 3 additional dialysis-related items. The total QOL score and the four subscale scores range between 0 and 30 with higher scores indicating a better HRQOL [64].
End-Stage Renal Disease Symptom Checklist-Transplantation Module (ESRD-SCL-TM)	A 43-item symptom-specific QOL measure designed for use in renal transplant patients on immunosuppression therapy [70]. There are 6 dimensions: (i) limited physical capacity (ii) limited cognitive capacity (iii) cardiac and renal dysfunction (iv) side effects of corticosteroids (v) increased growth of gum and hair (vi) transplantation-related psychological distress. A 5-point Likert scale is used. Higher scores indicate worse QOL/symptoms [70].
Modified Transplant Symptom Occurrence and Symptom Distress Scale (MTSOSD)	A 29-item measure designed to assess side effects of immunosuppression therapy in renal transplant patients [74]. The symptom occurrence dimension has 20 items while the symptom distress dimension has 9 [74]. Ridit analysis, a statistical method of analysing ordinal data [123] was chosen for data analysis [74]. Higher ridit scores indicate greater symptom distress [74].

(*Continued*)

Table 3. (*Continued*)

Measure	Description
Gastrointestinal Symptom Rating Scale (GSRS)	A 15-item measure that assesses important gastrointestinal side effects of immunosuppressive therapy in renal transplant patients [75]. There are 5 dimensions: (i) reflux (ii) diarrhoea (iii) constipation (iv) abdominal pain (v) indigestion. Each dimension gives an average score ranging from 1 (no discomfort) to 7 (very severe discomfort) with higher scores indicating worse impact [75].
Gastrointestinal Quality of Life Index (GIQLI)	A 36-item measure which focuses on the impact GI complaints on the HRQOL of renal transplant patients [75]. There are 5 dimensions: (i) GI symptoms (ii) emotional status (iii) physical function (iv) social function (v) strain of medical treatment. Total scores range from 0 to 144 with higher scores indicating a better HRQOL [75].
Kidney Transplant Questionnaire (KTQ)	A 25-item HRQOL measure for use in renal transplant patients [78]. It has 5 dimensions namely: (i) physical symptoms (ii) fatigue (iii) uncertainty/fear (iv) appearance (v) emotional. The physical symptom dimension is patient-specific, thus the symptoms most important to individual patients are identified [78]. For each dimension, an average score ranging from 1 to 7 is calculated with higher scores indicating better HRQOL in patients [78].
ReTransQoL (RTQ) version 1	A 45-item measure designed to assess QOL in renal transplant patients [80]. There are 5 dimensions: (i) physical health (ii) mental health (iii) medical care (iv) fear of losing graft (v) treatment. All dimensions are linearly transformed to a 0 to 100 scale and higher scores indicate better HRQOL [80].
ReTransQoL (RTQ) version 2	A 32-item measure designed to assess QOL in renal transplant patients [80]. There are 5 dimensions: (i) physical health (ii) social functioning (iii) medical care (iv) treatment (v) fear of losing graft. All dimensions are linearly transformed to a 0 to 100 scale and higher scores indicate better HRQOL [80].
CKD-Symptom Burden Index (CKD-SBI)	A 32-item measure of symptom burden. Derived from the DSI. The CKD-SBI was developed for use in patients with CKD stages IV and V however it was used in pre-dialysis and dialysis populations in this study [82]. The measure has 4 dimensions namely: (i) prevalence (ii) distress (iii) severity (iv) frequency. Total score ranges from 0 to 100 and higher scores indicate higher symptom burden [82].
Generic measures: These measure health in a general manner and can be used for various health conditions [122].	
Nottingham Health Profile (NHP)	A generic HRQOL questionnaire with 38 yes/no questions [124] grouped into 6 dimensions: (i) pain (ii) energy (iii) physical mobility (iv) sleep (v) emotional reactions (vi) social isolation. There is an optional part. The NHP scores range between 0 (good health status) and 100 (poor health status) [124].
SF-36 version 2	A generic 36-item HRQOL measure with 8 scales [125] namely: (i) physical functioning (ii) physical role (iii) bodily pain (iv) general health (v) vitality (vi) social functioning (vii) emotional role (viii) mental health. An additional 1-item measure of self-evaluated change in health status is available. The Likert rating method is used and raw scores are linearly transformed into 0 to 100 scales with higher transformed scores indicating better HRQOL [125].

(*Continued*)

Table 3. (*Continued*)

Measure	Description
SF-12	A generic 12-item HRQOL measure derived from the SF-36 [126] (see above). The 8 dimensions can be computed into 2 distinct clusters, PCS-12 and MCS-12 with higher values indicating better HRQOL [126].
Utility measures: These provide utilities or values regarding health and can be used for cost-utility analyses of interventions [18].	
EQ-5D	A utility measure with a self-classifier and a visual analogue scale (VAS) which can be used to value health states [127]. The self-classifier includes 5 dimensions: (i) mobility (ii) self-care (iii) usual activities (iv) pain/ discomfort (v) anxiety/depression. Each dimension has 3 levels of severity (no problems, some problems, and severe problems) and it is possible to describe 243 health states between 0 (dead) and 1 (perfect health) [127].
Modified Time Trade-Off (TTO)	This utility measure was used to measure quality of life in dialysis and renal transplant populations and values range from 0 (death) to 1 (full health) [83, 128].

https://doi.org/10.1371/journal.pone.0179733.t003

Pre-dialysis population. Three disease-specific PROMs were used in pre-dialysis populations namely the: Agarwal [32], KDQOL-36 [33, 34] and KDQOL-SF [35, 36]. Although the studies in this section measured estimated glomerular filtration rates (eGFR), there was significant disparity in their description of patients at this stage of CKD. Three studies described patients as 'non-dialysis' [32], 'mild-to-moderate' CKD patients [34] and ESRD patients [36] respectively, while only the studies by Chao et al. [33] and Abd Elhafeez [35] formally categorized patients into CKD stages. None of the studies validated PROMs separately by stage of CKD pre-dialysis.

Agarwal. There was limited evidence for test-retest reliability and content validity. Internal consistency and structural validity were rated as 'unknown' [32].

Kidney Disease Quality of Life—36 (KDQOL-36). Strong evidence for internal consistency was found [33] and there was moderate evidence for hypothesis testing [33, 34]. Structural validity was rated as 'indeterminate' [33].

KDQOL-SF. There was moderate evidence for test-retest reliability [35] and hypothesis testing [35, 36]. Internal consistency, structural validity and responsiveness were rated as 'unknown' [35, 36].

Dialysis population. Fourteen PROMS were used in this group of patients. Among the five studies that evaluated KDQOL-36, two had mixed samples [37, 38]. As the majority of the participants in Tao et al [38] were on dialysis and no significant transplant specific symptoms were elicited in Chow et al. [37] a pragmatic decision was made to analyse them here.

Klersy et al. [39] used a sample which comprised of 85% dialysis and 15% pre-dialysis patients to assess the KDQOL-SF. Again, as the majority were dialysis patients, this study was analysed in this section.

KDQOL-36. Moderate evidence was found for internal consistency [40], test-retest reliability [41, 42] and hypothesis testing [37, 38]. However, structural validity was rated as 'indeterminate' [40].

KDQOL-SF. There was strong evidence of internal consistency and structural validity [43]. Moderate evidence was found for test-retest reliability [44, 45] and hypothesis testing [46] while there was limited evidence for content validity [47]. Responsiveness was rated as 'unknown' [48].

Table 4. Methodological quality of included studies.

Instrument	Study	Language	CKD population	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Translation	Criterion validity	Responsiveness
Agarwal	Agarwal [32]	English	Pre-dialysis	POOR	FAIR		FAIR	POOR	FAIR			
CDQOL	Suet-Ching [49]	Cantonese	Dialysis	POOR	POOR		FAIR		FAIR			
CHEQ	Aiyasanon [51]	Thai	Dialysis	POOR					FAIR	GOOD		
	Wu [50]	English	Dialysis	POOR			GOOD		GOOD			
CKD-SBI	Almutary [82]	Arabic	Pre-dialysis, Dialysis	POOR					FAIR	FAIR		
DSI	Önsoz [53]	Turkish	Dialysis	POOR	FAIR					GOOD		
	Weisbord [52]	English	Dialysis		POOR		FAIR					
QLI 3.0	Dehesh [65]	Persian	Dialysis	POOR				POOR	FAIR	GOOD		
	Ferrans [64]	English	Dialysis	POOR	POOR				FAIR			
	Halabi [66]	Arabic	Dialysis	POOR						FAIR		
	Korkut [67]	Turkish	Dialysis	POOR	FAIR				FAIR	FAIR		
KDQOL (D)	Hays [58]	English	Dialysis	POOR			FAIR	POOR	FAIR			
KDQOL (M)	Rao [59]	English	Dialysis	POOR					FAIR			
KDQOL-36	Chao [33]	Taiwanese	Pre-dialysis	EXCELLENT				EXCELLENT	FAIR	GOOD		
	Chow [37]	Cantonese	Dialysis	POOR	POOR				GOOD			
	Mateti [42]	Kannada	Dialysis	FAIR	FAIR				FAIR	GOOD		
	Ricardo [34]	English, Spanish	Pre-dialysis	POOR					GOOD			
	Tao [38]	Mandarin	Dialysis	FAIR	POOR				FAIR			
	T'charoen [41]	Thai	Dialysis	POOR	FAIR				FAIR			
	Yang [40]	English	Dialysis	GOOD				GOOD	GOOD			
KDQOL-SF	Abd ElHafeez [35]	Arabic	Pre-dialysis	POOR	GOOD			POOR	FAIR	GOOD		
	Barotfi [76]	Hungarian	Transplant	POOR	FAIR				FAIR			
	Bataclan [88]	Filipino	Dialysis	POOR	FAIR				FAIR	GOOD		
	Boini [89]	French	Dialysis	POOR	FAIR				FAIR			
	Bouidida [46]	Moroccan	Dialysis	POOR	POOR				GOOD	GOOD		
	Cheung [36]	Chinese	Pre-dialysis	POOR					FAIR			POOR
	Duarte [96]	Portuguese	Dialysis		FAIR				FAIR	GOOD		
	Fardinmehr [94]	Persian	Dialysis	POOR					FAIR	GOOD		
KDQOL-SF	Green [85]	Japanese	Dialysis	POOR	GOOD				FAIR	GOOD		
	Hays [47]	English	Dialysis	POOR			FAIR					
	Joshi [43]	English, Mandarin, Malay	Dialysis	EXCELLENT				EXCELLENT	GOOD			
	Klersy [39]	Italian	Dialysis	POOR	POOR			POOR	FAIR	GOOD		
	Kontodimopoulos [86]	Greek	Dialysis	POOR					FAIR			
	Kontodimopoulos [87]	Greek	Dialysis	POOR					FAIR			
	Korevaar [48]	Dutch	Dialysis	POOR					GOOD	GOOD		POOR
	Malindretos [44]	Greek	Dialysis	POOR	FAIR				FAIR	GOOD		
	Molsted [92]	Danish	Dialysis	POOR						GOOD		
	Moreira [93]	Portuguese	Dialysis	POOR					FAIR			
	Pakpour [102]	Farsi	Dialysis	FAIR	POOR			FAIR	FAIR	GOOD		
	Park [45]	Korean	Dialysis	POOR	FAIR				POOR	GOOD		
	Perneger [90]	French	Dialysis	POOR					FAIR			
	Vasilieva [95]	Russian	Dialysis	POOR								
	Yildirim [91]	Turkish	Dialysis	POOR					FAIR	FAIR		
KDQ	Alvarez-Ude [57]	Spanish	Dialysis	POOR	FAIR				FAIR	GOOD		

(*Continued*)

Table 4. (*Continued*)

Instrument	Study	Language	CKD population	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Translation	Criterion validity	Responsiveness
	Laupacis [56]	English	Dialysis		FAIR		FAIR	POOR	FAIR			POOR
ESRD-SCL	Franke [70]	German	Transplant	GOOD	POOR		FAIR	GOOD	FAIR			
	Ortega [71]	Spanish	Transplant	GOOD	FAIR			GOOD	FAIR			POOR
	Stavem [72]	Norwegian	Transplant	FAIR	FAIR				GOOD	GOOD		
GIQLI	Kleinman [75]	German, English	Transplant	POOR					GOOD			
GSRS	Kleinman [75]	German, English	Transplant	POOR					GOOD			
KTQ	Chisholm-Burns [84]	English	Transplant	POOR					FAIR			
	Laupacis [78]	English	Transplant	POOR	POOR		FAIR	POOR	FAIR			POOR
	Niu [77]	Chinese	Transplant	GOOD	FAIR			GOOD	FAIR	GOOD		
	Rebollo [79]	Spanish	Transplant	POOR	POOR				FAIR	POOR		POOR
ESAS	Davison [55]	English	Dialysis		FAIR		FAIR		FAIR			
	Davison [54]	English	Dialysis									POOR
MTSOSD	Moons [74]	Dutch	Transplant						GOOD			
RTQ v1	Beauger [81]	French	Transplant					EXCELLENT				
	Gentile [80]	French	Transplant	POOR	POOR		GOOD	POOR	FAIR			POOR
RTQ v2	Beauger [81]	French	Transplant	GOOD				GOOD	FAIR			
EQ-5D	Cleemput [73]	Dutch, French	Transplant						GOOD			
SF-12	Pakpour [62]	Persian	Dialysis	GOOD	GOOD			GOOD	FAIR			
SF-36 v2	Feurer [69]	English	Transplant	POOR					FAIR			
	Mingardi [68]	Italian	Dialysis	GOOD					FAIR			
NHP	Badia [61]	Spanish	Dialysis	FAIR	FAIR							
	Zengin [60]	Turkish	Dialysis	FAIR					FAIR			
TTO (modified)	Churchill [83]	English	Dialysis, Transplant		FAIR				FAIR			
WHOQOL-BREF (D)	Yang [63]	Taiwanese	Dialysis	GOOD	POOR			GOOD	GOOD			

https://doi.org/10.1371/journal.pone.0179733.t004

Table 5. Evidence synthesis of PROMs used in patients with CKD.

Instrument version	Population	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness
Agarwal [32]	Pre-dialysis	?	+		+	?	-		
KDQOL-36 [33, 34, 37, 38, 40–42]	Pre-dialysis	+++				?*	++		
	Dialysis	++	++			?*	++		
KDQOL-SF [35, 36, 39, 43–48, 76, 85–96, 102]	Pre-dialysis	?	++			?	++		?
	Dialysis	+++	++		+	+++	++		?
	Transplant	?	-				+		
CDQOL [49]	Dialysis	?	?		+		+		
CHEQ [50, 51]	Dialysis	?			++		++		
DSI [52, 53]	Dialysis	?	?		+				
ESAS [54, 55]	Dialysis		+		+		+		?
KDQ [56, 57]	Dialysis	?	+		+	?	++		?
KDQOL (D)[58]	Dialysis	?			+	?			
KDQOL (M)[59]	Dialysis	?					+		
NHP [60, 61]	Dialysis	++	?				+		
SF-12 [62]	Dialysis	++	++			+	+		
WHOQOL-BREF (D) [63]	Dialysis	++	?			?	-		
QLI 3.0 [64–67]	Dialysis	?	-			?	++		
SF-36 v2	Dialysis [68]	++					+		
	Transplant [69]	?					-		
ESRD-SCL [70–72]	Transplant	+++	++		+	++	++		?
EQ-5D [73]	Transplant						++		
GIQLI [75]	Transplant	?					++		
GSRS [75]	Transplant	?					++		
KTQ [77–79, 84]	Transplant	++	+		+	?	++		?
MTSOSD [74]	Transplant						++		
RTQ v1 [80, 81]	Transplant	?	?		++	+/-	+		?
RTQ v2 [81]	Transplant	++				++	+		
TTO (modified) [83]	Mixed (D & TX)		+				+		
CKD-SBI [82]	Mixed (D & Pre-D)	?					+		

+ = positive rating

? = unknown rating

- = negative rating

+/- = conflicting findings

?* = indeterminate rating (due to non-reporting of variance explained by factors)

https://doi.org/10.1371/journal.pone.0179733.t005

Chinese Dialysis Quality of Life Scale (CDQOL). Reliability (internal consistency and test-retest) for CDQOL was rated as 'unknown' and there was limited evidence for content validity and hypothesis testing [49].

CHOICE Health Experience Questionnaire (CHEQ). Moderate evidence for hypothesis testing and content validity was found [50] while internal consistency was rated as 'unknown' [50, 51].

Dialysis Symptom Index (DSI). There was limited evidence for content validity [52]. Reliability (internal consistency and test-retest) was rated as 'unknown' [52, 53].

Modified Edmonton Symptom Assessment System (modified ESAS). There was limited evidence of test-retest reliability, content validity and hypothesis testing. Responsiveness was rated as 'unknown' [54, 55].

Kidney Disease Questionnaire (KDQ). There was moderate evidence of hypothesis testing [56, 57] while there was limited evidence for test-retest reliability and content validity [56]. Structural validity, responsiveness [56] and internal consistency [57] were rated as 'unknown'.

KDQOL. There was limited evidence for content validity while internal consistency and structural validity were rated as 'unknown' [58].

KDQOL (Modified). There was limited evidence for hypothesis testing while internal consistency was rated as 'unknown' [59].

Nottingham Health Profile (NHP). There was moderate evidence for internal consistency [60, 61] and limited evidence for hypothesis testing [60]. Test-retest reliability was rated as 'unknown' [61].

SF-12. Moderate evidence was found for internal consistency and test-retest reliability. There was limited evidence for structural validity and hypothesis testing [62].

WHOQOL-BREF (Dialysis). There was moderate evidence for internal consistency while test-retest reliability and structural validity were rated as 'unknown' [63].

Quality of Life Index (QLI) 3.0. There was moderate evidence for hypothesis testing [64, 65]. Internal consistency and structural validity were rated as 'unknown' [64–67] while there was limited evidence against test-retest reliability [67].

SF-36 version 2. Moderate evidence was found for internal consistency while limited evidence was found for hypothesis testing [68].

Renal transplant population. Ten PROMs were evaluated specifically in renal transplant patients and all except the SF-36 and the EQ-5D were disease specific measures.

SF-36 version 2. Internal consistency was rated as 'unknown' [69].

End-Stage Renal Disease Symptom Checklist–Transplantation Module (ESRD-SCL-TM). There was strong evidence for internal consistency [70, 71], moderate evidence for test-retest reliability [71, 72], hypothesis testing [72] and structural validity [71]. There was limited evidence for content validity [70] while responsiveness was rated as 'unknown' [71].

EQ-5D. Moderate evidence for hypothesis testing was found [73].

Modified Transplant Symptom Occurrence and Symptom Distress Scale (MTSOSD). Moderate evidence for hypothesis testing was found for MTSOSD [74].

Gastrointestinal Symptom Rating Scale (GSRS). Moderate evidence for hypothesis testing was found while internal consistency was rated as 'unknown' [75].

Gastrointestinal Quality of Life Index (GIQLI). Moderate evidence for hypothesis testing was found while internal consistency was rated as 'unknown' [75].

KDQOL-SF. There was limited evidence for hypothesis testing and limited evidence against test-retest reliability. Internal consistency was rated as 'unknown' [76].

Kidney Transplant Questionnaire (KTQ). There was moderate evidence for internal consistency [77] and hypothesis testing [78, 79]. Limited evidence was found for test-retest reliability [77] and content validity [78] Structural validity [77] and responsiveness [79] were rated as 'unknown'.

ReTransQoL (RTQ) version 1. There was moderate evidence of content validity [80], conflicting evidence of structural validity [80, 81] and limited evidence of hypothesis testing [80]. Internal consistency, test-retest reliability and responsiveness were rated as 'unknown' [80].

RTQ version 2. This revised version had moderate evidence for internal consistency and structural validity. There was still limited evidence for hypothesis testing [81].

Single studies with mixed samples. The study by Almutary et al. [82] evaluated the CKD-Symptom Burden Index (CKD-SBI) in a mixed sample of pre-dialysis and dialysis patients while Churchill et al. [83] assessed the Time Trade-Off (TTO) in a mixed sample of dialysis and renal transplant recipients.

CKD-Symptom Burden Index (CKD-SBI)

Limited evidence was found for hypothesis testing while internal consistency was rated as 'unknown' [82].

Modified Time Trade-Off (TTO)

There was limited evidence for test-retest reliability and hypothesis testing [83].

Other findings

Reliability. Internal consistency: Of the 58 studies that assessed internal consistency, 42 were scored 'poor' for methodological quality. This was due to one of 3 reasons:

1. Some studies did not conduct a factor analysis and did not reference a relevant study that conducted one [41, 49–51, 53, 61, 64, 75, 82].
2. Some studies conducted a factor analysis but had inadequate sample sizes by COSMIN standards [32, 35, 39, 47, 59, 65, 78, 80].
3. Some studies referenced a relevant study that conducted a factor analysis but the sample size used for the study was inadequate by COSMIN standards [34, 36, 37, 44–48, 57, 66, 67, 69, 76, 79, 84–95].

Test-retest reliability: Thirty-four studies conducted test-retest reliability and most of them reported internal correlation coefficients (ICC). The study by Duarte et al. [96] was the only one that reported inter & intra-observer reliability. The majority scored 'fair' for test-retest reliability and this was largely due to the small sample sizes.

Measurement error: The included studies did not provide adequate information on parameters such as the minimal important change (MIC) [97], the standard error of measurement (SEM) [30, 98] or the limits of agreement (LOA) [99] making it difficult to assess measurement error. Only one study provided an estimate for minimal clinical important difference (MCID) [75].

Validity. Content validity: This was assessed by development studies for RTQ [80], modified ESAS [55], Agarwal [32], CHEQ [50], CDQOL [49], DSI [52], KDQ [56], KTQ [78], KDQOL [58], KDQOL-SF [47]. Six validation studies reported content validity indexes (CVI) [33, 38, 49, 53, 77, 82] and of these only Suet-Ching [49] reported patient involvement in the process of content validation.

Construct validity (Hypothesis testing): Of the 66 studies evaluated, only 13 reported clearly formulated a priori hypotheses or expectations regarding the magnitude and direction of correlations. For this reason, most of the studies were rated as 'fair' for the methodological quality of their hypothesis testing.

The absence of clear a priori hypotheses make it difficult to determine whether any results reported for construct validity and responsiveness was due to chance or not [100, 101].

Structural validity: A number of measures were rated poorly for structural validity due to issues with factor analysis (See internal consistency). Out of the 19 studies that conducted a factor analysis, 8 were scored 'poor' [32, 35, 39, 56, 58, 65, 78, 80] and this can be attributed to

the use of sample sizes inadequate by COSMIN standards (n < 5 times the total number of items and < 100) [100]. It is important that studies perform factor analysis as it verifies scale structure and uni-dimensionality which determines the scoring and interpretation of a measure's internal consistency statistic [100].

Criterion validity: This was not assessed for any study as the COSMIN Delphi panel does not regard any PROM as true 'gold standard' [29] and the FDA holds a similar view [14].

Cross-cultural validity/Translations: Twenty-five studies translated PROM instruments and adapted them to varying degrees for their study population. As none of these studies performed a multi-group confirmatory factor analysis or assessed differential item functioning (DIF) between language groups, the decision was made not to assess cross-cultural validity. Therefore, only the quality of their translations was assessed according to the provisions specified in the COSMIN manual [100].

The translations conducted by 20 studies [33, 35, 39, 42, 44–46, 48, 51, 53, 57, 65, 72, 77, 85, 88, 92, 94, 96, 102] were rated as good, while the translations by 4 studies [66, 67, 82, 91] were rated as fair and 1 translation was rated as poor [79].

Responsiveness. Responsiveness was only assessed by 8 of the included studies [36, 48, 54, 56, 71, 78–80] and all were rated 'poor' as the information provided was inadequate.

Interpretability. A number of the PROMs included in this review had significant floor and ceiling effects which might indicate a reduction in their ability to discriminate between patients with the lowest or highest possible scores and detect changes over time [30].

Ten instruments (KDQOL-SF, Agarwal, KDQ, KDQOL-36, KTQ, QLI, KDQOL, ESRD-SCL, CHEQ and NHP) had floor and ceiling effects > 15% while the RTQ v2 and SF-12 [62, 81] had < 15%.

The 'work status' dimension of the KDQOL-SF had a 70% floor effect [45, 86] while the 'social support', 'patient satisfaction' and 'staff encouragement' domains had ceiling effects of 46%, 45% and 67% [86], respectively. Sexual function had ceiling effects of 53.3% while cognitive function had ceiling effects of 60% [35]. The 'pain' dimension of the SF-36 had ceiling effects as high as 59% [35].

Eight studies namely; RTQ v1 [80], RTQv2 [81], GSRS & GIQLI [75], KDQOL-36 [38, 41], KDQOL-SF [44], CHEQ [51], TTO [83] reported measurement scores for subgroups within their study populations. While some of the differences in scores were statistically significant, it is unclear if any were clinically relevant.

Feasibility and acceptability. This was difficult to assess for the included studies as less than a third reported the average time needed to complete the questionnaires, few reported the recall period used and none reported administrative requirements for collection and analysis of data. However, most of the studies reported good response rates which suggest that patients might find the use of PROMs acceptable.

The fact that majority of the studies failed to report the level and/or the method used for handling missing data meant there might be a risk of bias [101]. Questions relating to sexual activities had the highest levels of missing data with Bataclan and Dial [88] reporting a response rate of <18% in Filipino patients [88].

Patient involvement. Patient involvement in the process of item generation and/or item selection was reported in the development of the RTQ [80], KDQ [56], KTQ [78], ESRD-SCLTM [70], KDQOL-SF [47], Agarwal [32], CHEQ [50], QLI 3.0 [64], KDQOL (Dialysis) [58] and CDQOL [49].

All the studies that translated and adapted measures reported pre-testing their translations in patients to assess a combination of comprehension, cultural relevance and acceptability except Rebollo et al.[79]

Discussion

This is the first review to use the COSMIN checklist [100] to evaluate the measurement properties of PROMs used in patients with CKD. In all, 25 PROMs were evaluated by a total of 66 studies in pre-dialysis, dialysis, and renal transplant patients.

In the pre-dialysis population, the KDQOL-36 exhibited strong evidence for internal consistency and moderate evidence for construct validity (hypothesis testing). It should be noted that the evidence for this measure was obtained from studies conducted in Taiwanese patients [33], and a combination of Hispanic and non-Hispanic white patients [34]. Furthermore, the measurement properties were not reported by CKD stage. Therefore, further validation may be necessary before use outside these study populations and/or where focus is on the use of PROMs in relationship to the severity of CKD.

In dialysis patients, we found evidence to support the use of both the KDQOL-SF and the KDQOL-36. The KDQOL-SF demonstrated strong evidence for internal consistency and structural validity and moderate evidence for test-retest reliability and construct validity (hypothesis testing) while the KDQOL-36 had moderate evidence for internal consistency, test-retest reliability and construct validity (hypothesis testing). Again it should be noted that this evidence was obtained from a significant number of non-English studies, further validation work would be needed before these measures could be confidently utilised in an English speaking population.

In renal transplant patients, the ESRD-SCLTM demonstrated strong evidence for internal consistency and moderate evidence for test-retest reliability, structural validity and construct validity (hypothesis testing).

Consistent with the review by Gibbons and Fitzpatrick [18], much of the evidence we present was derived from cross-sectional studies. However, in contrast to that study, we did not include clinical trial reports as the methodological quality of their PROM evaluations are often unsatisfactory [23] and their PROM analysis often inadequate and insufficiently [103] reported for any meaningful evaluation to be possible [19].

In line with our findings, Gibbons and Fitzpatrick [18] found evidence to support the use of the KDQOL-SF but did not specify which modality of renal replacement therapy (RRT) provided the evidence. We found evidence to support the use of the KDQOL-36 (which was not available at the time Gibbons and Fitzpatrick [18] conducted their review), and the ESRD-SCLTM (which was excluded from their study) [18]. The ERA-EDTA expert panel [104] recommended the KDQOL-36, following a consensus meeting.

There were methodological issues with the majority of the PROMs included in this review. These included: sample sizes smaller than current recommendations, a lack of clearly described a priori hypotheses and inadequate reporting of missing data, and little or no information on measurement error and responsiveness. Similar issues were reported by Gibbons and Fitzpatrick [18] and highlighted in reviews for other health conditions [105–107].

Given that the COSMIN standards only became available within the last decade, it is unsurprising that most of the earlier studies fared poorly when judged against these exacting methodological and reporting standards even though they might meet the minimum standards recommended by the International Society for Quality of Life Research (ISOQOL) [108]. However, this highlights the need to test and revise PROM instruments on a regular basis to ensure they actually perform as intended according to contemporary psychometric standards.

This systematic review provides a basis for identifying PROMs with potential utility in clinical practice. There is evidence that the use of PROMs in clinical practice could enhance communication [109–112] between patients and their clinicians. Basch et al. [113] noted that the use of PROMs in routine clinical care, could facilitate the reporting of serious adverse events due to drug toxicities [113], while a review by Finkelstein et al. [114] suggested that the use of

PROMs could assist renal teams with the development of strategies to improve the HRQOL of the patient with CKD [114]. Calvert et al. [115] suggested that PROM data could potentially facilitate the delivery of tailored healthcare if successfully integrated with routinely collected clinical and laboratory data [115].

Whilst effective management of risk factors can slow CKD progression [116] many patients with severe pre-dialysis CKD progress to end stage renal disease (ESRD) and in this group PROMs may have a significant role [117–119]. For example, PROMs could be used to monitor individual patients for symptoms or changes in HRQOL that may indicate that a medical review or intervention is needed [17, 120].

Whilst a systematic review can provide valuable evidence on psychometric properties of PROMs, clinicians and researchers need to consider a number of issues such as the domains covered by different measures, the available supporting evidence, and the suitability for the target population and use (whether for clinical trials, routine practice, audit or real-time decision making). As no single measure covers all the domains that might be of interest, there might be a need to administer more than one measure. Patient acceptability is also a key issue. Therefore, it is important that patients are involved not just in the development of measures but also in the selection for research and/or practice.

During the course of this review, we became aware of the existence of the IPOS-Renal and contacted its developers [121]. This measure is currently being used within the measurement work stream of the UK Renal Registry and its validation by the Palliative care Outcome Scale (POS) team is on-going [121]. Therefore, evidence to support its use may be available in future.

The key strengths of this review are the use of the COSMIN standards and criteria for evidence synthesis which ensured that our assessments of the included PROMs were robust, the stratification of the review by stage of CKD and the absence of language restrictions which strengthened our findings. The main limitation is the fact that the included studies did not adequately report their assessments for a number of measurement properties thus making it difficult and sometimes impossible to evaluate these properties.

At present, we suggest the use of the KDQOL-36 in pre-dialysis patients though initial validation might be required. We recommend the KDQOL-SF and the KDQOL-36 for use in dialysis patients as we found evidence supporting both of these measures. The shorter 36-item KDQOL-36 may be more practical for use in routine clinic settings, while the longer 80-item KDQOL-SF might be preferred for research purposes where more detailed information may be required. We suggest using the ESRD-SCLTM in renal transplant recipients to assess issues pertaining to renal transplantation and immunosuppression therapy. These measures are recommended based on the fact that they currently possess the best evidence available according to COSMIN standards in these populations and meet the minimum standards recommended by ISOQOL [108]. However, it must be recognised that none of these measures possess evidence underpinning *all* measurement properties and some lack validation in English-speaking populations. Future work should be undertaken to address these gaps. For all measures, it is vital that content validity is established according to FDA guidelines to ensure that they actually measure the concept(s) of interest. This should be conducted before other measurement properties are fully evaluated and adequately reported in order to facilitate their subsequent evaluation. Investigators may use this review to identify the gaps in evidence and design studies to address these issues in future.

Supporting information

S1 Text. Review protocol.
(PDF)

S2 Text. PRISMA checklist.
(DOC)

S3 Text. Search strategy.
(PDF)

S1 Table. Summary of study results.
(DOCX)

S2 Table. Characteristics of included studies.
(DOCX)

S3 Table. Characteristics of included PROMs.
(DOCX)

Acknowledgments

We are grateful to Susan Bayliss who assisted with the development of the search strategy [24]; Dr Lazaros Andronis, and Dr Fouad Al-Baaj who translated some papers; and Dr Harjeet Bhachu who assisted with the tables.

Author Contributions

Conceptualization: MC DK PC TM OLA.

Data curation: OLA TK AG AS.

Formal analysis: OLA AS.

Funding acquisition: MC DK PC TM.

Investigation: OLA TK AG AS.

Methodology: MC DK PC TM OLA.

Project administration: MC DK PC TM OLA.

Supervision: MC DK PC TM.

Visualization: MC DK PC TM OLA.

Writing – original draft: OLA.

Writing – review & editing: MC DK PC TM OLA TK AG AS.

References

1. Couser WG, Remuzzi G, Mendis S, Tonelli M. The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. Kidney international. 2011; 80(12):1258–70. Epub 2011/10/14. https://doi.org/10.1038/ki.2011.368 PMID: 21993585.
2. Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, de Jong PE, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet (London, England). 2010; 375 (9731):2073–81. Epub 2010/05/21. https://doi.org/10.1016/s0140-6736(10)60674-5 PMID: 20483451; PubMed Central PMCID: PMCPmc3993088.
3. Nitsch D, Grams M, Sang Y, Black C, Cirillo M, Djurdjev O, et al. Associations of estimated glomerular filtration rate and albuminuria with mortality and renal failure by sex: a meta-analysis. BMJ (Clinical research ed). 2013; 346:f324. Epub 2013/01/31. https://doi.org/10.1136/bmj.f324 PMID: 23360717; PubMed Central PMCID: PMCPmc3558410.

4. Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al. Global Prevalence of Chronic Kidney Disease—A Systematic Review and Meta-Analysis. PloS one. 2016; 11(7):e0158765. Epub 2016/07/08. https://doi.org/10.1371/journal.pone.0158765 PMID: 27383068; PubMed Central PMCID: PMCPmc4934905.
5. Hamer RA, El Nahas AM. The burden of chronic kidney disease: Is rising rapidly worldwide. BMJ: British Medical Journal. 2006; 332(7541):563–4. PMC1397782. https://doi.org/10.1136/bmj.332.7541.563 PMID: 16528062
6. Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association—European Renal Association. 2012; 27 Suppl 3:iii73–80. Epub 2012/07/21. https://doi.org/10.1093/ndt/gfs269 PMID: 22815543; PubMed Central PMCID: PMCPmc3484716.
7. Honeycutt AA, Segel JE, Zhuo X, Hoerger TJ, Imai K, Williams D. Medical Costs of CKD in the Medicare Population. Journal of the American Society of Nephrology. 2013; 24(9):1478–83. https://doi.org/10.1681/ASN.2012040392 PMID: 23907508
8. Meyer T, Hostetter T. Medical Progress: Uremia. The New England journal of medicine. 2007; 13 (357):1316–25.
9. Jablonski A. The multidimensional characteristics of symptoms reported by patients on hemodialysis. Nephrol Nurs J. 2007; 34(1):29–37; quiz 8. PMID: 17345690.
10. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. Journal of Renal Care. 2013; 39(3):140–50. http://dx.doi.org/10.1111/j.1755-6686.2013.12022.x. PMID: 23826803. https://doi.org/10.1111/j.1755-6686.2013.12022.x
11. Merkus MP, Jager KJ, Dekker FW, de Haan RJ, Boeschoten EW, Krediet RT. Physical symptoms and quality of life in patients on chronic dialysis: results of The Netherlands Cooperative Study on Adequacy of Dialysis (NECOSAD). Nephrology Dialysis Transplantation. 1999; 14(5):1163–70. https://doi.org/10.1093/ndt/14.5.1163
12. Calvert MJ, Freemantle N. Use of health-related quality of life in prescribing research. Part 1: why evaluate health-related quality of life? Journal of clinical pharmacy and therapeutics. 2003; 28(6):513–21. Epub 2003/12/04. PMID: 14651676.
13. Kyte DG, Draper H, Ives J, Liles C, Gheorghe A, Calvert M. Patient reported outcomes (PROs) in clinical trials: is 'in-trial' guidance lacking? a systematic review. PloS one. 2013; 8(4):e60684. Epub 2013/04/06. https://doi.org/10.1371/journal.pone.0060684 PMID: 23560103; PubMed Central PMCID: PMCPmc3613381.
14. FDA. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Guidance for industry. 2009. Available: https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf. Accessed 22 May 2017.
15. Black N. Patient reported outcome measures could help transform healthcare. BMJ (Clinical research ed). 2013; 346:f167. Epub 2013/01/30. https://doi.org/10.1136/bmj.f167 PMID: 23358487.
16. Chris Ham CI, Nick Goodwin, Anna Dixon, Patrick South. Where next for the NHS reforms? The case for integrated care. The King's Fund. 2011.
17. Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. Interactive journal of medical research. 2014; 3(1):e5. Epub 2014/02/13. https://doi.org/10.2196/ijmr.2885 PMID: 24518281; PubMed Central PMCID: PMCPmc3936283.
18. Gibbons E, Fitzpatrick, R. A structured review of patient-reported outcome measures for people with chronic kidney disease. Department of Public Health University of Oxford. 2010.
19. Terwee CB. Protocol for systematic reviews of measurement properties. COSMIN. 2011.
20. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2012; 21(4):651–7. Epub 2011/07/07. https://doi.org/10.1007/s11136-011-9960-1 PMID: 21732199; PubMed Central PMCID: PMCPmc3323819.
21. Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. When and how to update systematic reviews. The Cochrane database of systematic reviews. 2008;(1):Mr000023. Epub 2008/02/07. https://doi.org/10.1002/14651858.MR000023.pub3 PMID: 18254126.
22. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2010; 19(4):539–49. Epub 2010/02/20.

https://doi.org/10.1007/s11136-010-9606-8 PMID: 20169472; PubMed Central PMCID: PMCPmc2852520.

23. Martini C, Gamper EM, Wintner L, Nilica B, Sperner-Unterweger B, Holzner B, et al. Systematic review reveals lack of quality in reporting health-related quality of life in patients with gastroenteropancreatic neuroendocrine tumours. Health and quality of life outcomes. 2016; 14(1):127. Epub 2016/09/12. https://doi.org/10.1186/s12955-016-0527-2 PMID: 27614762; PubMed Central PMCID: PMCPmc5018190.

24. Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Keeley T, Gheorghe A, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol. BMJ Open. 2016; 6(10):e012014. Epub 2016/10/14. https://doi.org/10.1136/bmjopen-2016-012014 PMID: 27733411.

25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ (Clinical research ed). 2009; 339:b2700. Epub 2009/07/23. https://doi.org/10.1136/bmj.b2700 PMID: 19622552; PubMed Central PMCID: PMCPmc2714672.

26. Terwee CB, Prinsen CA, Ricci Garotti MG, Suman A, de Vet HC, Mokkink LB. The quality of systematic reviews of health-related outcome measurement instruments. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2016; 25(4):767–79. Epub 2015/09/09. https://doi.org/10.1007/s11136-015-1122-4 PMID: 26346986; PubMed Central PMCID: PMCPmc4830864.

27. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2009; 18(8):1115–23. Epub 2009/08/28. https://doi.org/10.1007/s11136-009-9528-5 PMID: 19711195; PubMed Central PMCID: PMCPmc2744791.

28. PROM Group AM, Carolina Casañas i Comabella, Monica Hadi, Elizabeth Gibbons, Ray Fitzpatrick, Nia Roberts. PROM GROUP CONSTRUCT & INSTRUMENT TYPE FILTERS. 2010.

29. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of clinical epidemiology. 2010; 63(7):737–45. Epub 2010/05/25. https://doi.org/10.1016/j.jclinepi.2010.02.006 PMID: 20494804.

30. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology. 2007; 60(1):34–42. Epub 2006/12/13. https://doi.org/10.1016/j.jclinepi.2006.03.012 PMID: 17161752.

31. Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement. 1960; 20(1):37–46. https://doi.org/10.1177/001316446002000104

32. Agarwal R. Developing a self-administered CKD symptom assessment instrument. Nephrology Dialysis Transplantation. 2010; 25(1):160–6. http://dx.doi.org/10.1093/ndt/gfp426. PMID: 2010005312.

33. Chao S, Yen M, Lin TC, Sung JM, Wang MC, Hung SY. Psychometric Properties of the Kidney Disease Quality of Life-36 Questionnaire (KDQOL-36). West J Nurs Res. 2016; 38(8):1067–82. Epub 2016/03/31. https://doi.org/10.1177/0193945916640765 PMID: 27026384.

34. Ricardo AC, Hacker E, Lora CM, Ackerson L, De Salvo KB, Go A, et al. Validation of the Kidney Disease Quality of Life Short Form 36 (KDQOL-36TM) US Spanish and English versions in a cohort of Hispanics with chronic kidney disease. Ethnicity and Disease. 2013; 23(2):202–9. PMID: 23530302.

35. Abd Elhafeez S, Sallam SA, Gad ZM, Zoccali C, Torino C, Tripepi G, et al. Cultural adaptation and validation of the "kidney Disease and Quality of Life—Short Form (KDQOL-SFTM) version 1.3" questionnaire in Egypt. BMC nephrology. 2012; 13 (1) (no pagination)(170). http://dx.doi.org/10.1186/1471-2369-13-170. PMID: 2013024551.

36. Cheung YB, Seow YY, Qu LM, Yee ACP. Measurement properties of the Chinese version of the Kidney Disease Quality of Life-Short Form (KDQOL-SFTM) in end-stage renal disease patients with poor prognosis in singapore. Journal of Pain and Symptom Management. 2012; 44(6):923–32. http://dx.doi.org/10.1016/j.jpainsymman.2011.12.282. PMID: 22795902. https://doi.org/10.1016/j.jpainsymman.2011.12.282

37. Chow SK, Tam BM. Is the kidney disease quality of life-36 (KDQOL-36) a valid instrument for Chinese dialysis patients? BMC nephrology. 2014; 15:199. http://dx.doi.org/10.1186/1471-2369-15-199. PMID: 25511462; PubMed Central PMCID: PMCPMC4274701. https://doi.org/10.1186/1471-2369-15-199

38. Tao X, Chow SK, Wong FK. Determining the validity and reliability of the Chinese version of the Kidney Disease Quality of Life Questionnaire (KDQOL-36). BMC nephrology. 2014; 15:115. Epub 2014/07/13. https://doi.org/10.1186/1471-2369-15-115 PMID: 25015224; PubMed Central PMCID: PMCPmc4115482.

39. Klersy C, Callegari A, Giorgi I, Sepe V, Efficace E, Politi P. Italian translation, cultural adaptation and validation of KDQOL-SF, version 1.3, in patients with severe renal failure. J Nephrol. 2007; 20(1):43–51. Epub 2007/03/10. PMID: 17347972.

40. Yang F, Wang VW, Joshi VD, Lau TW, Luo N. Validation of the English version of the Kidney Disease Quality of Life questionnaire (KDQOL-36) in haemodialysis patients in Singapore. Patient. 2013; 6 (2):135–41. http://dx.doi.org/10.1007/s40271-013-0015-2. PMID: 23613285. https://doi.org/10.1007/s40271-013-0015-2

41. Thaweethamcharoen T, Srimongkol W, Noparatayaporn P, Jariyayothin P, Sukthinthai N, Aiyasanon N, et al. Validity and reliability of KDQOL-36 in thai kidney disease patient. Value in Health Regional Issues. 2013; 2(1):98–102. http://dx.doi.org/10.1016/j.vhri.2013.02.011. PMID: 2013289571.

42. Mateti UV, Nagappa AN, Attur RP, Nagaraju SP, Mayya SS, Balkrishnan R. Cross-cultural adaptation, validation and reliability of the South Indian (Kannada) version of the Kidney Disease and Quality of Life (KDQOL-36) instrument. Saudi J Kidney Dis Transpl. 2015; 26(6):1246–52. Epub 2015/11/21. https://doi.org/10.4103/1319-2442.168662 PMID: 26586066.

43. Joshi VD, Mooppil N, Lim JFY. Validation of the kidney disease quality of life-short form: A cross-sectional study of a dialysis-targeted health measure in singapore. BMC nephrology. 2010; 11 (1) (no pagination)(36). http://dx.doi.org/10.1186/1471-2369-11-36. PMID: 2011020886.

44. Malindretos P, Sarafidis P, Spaia S, Sioulis A, Zeggos N, Raptis V, et al. Adaptation and validation of the Kidney Disease Quality of Life-Short Form questionnaire in the Greek language. Am J Nephrol. 2010; 31(1):9–14. Epub 2009/10/30. https://doi.org/10.1159/000252926 PMID: 19864884.

45. Park HJ, Kim S, Yong JS, Han SS, Yang DH, Meguro M, et al. Reliability and validity of the Korean version of kidney disease quality of life instrument (KDQOL-SFTM). Tohoku Journal of Experimental Medicine. 2007; 211(4):321–9. http://dx.doi.org/10.1620/tjem.211.321. PMID: 17409671.

46. Rhou H, Ezaitouni F, Ouzeddoun N, Bayahia R, Elhajji K, Roudies R, et al. Translation, cultural adaptation and validation of the kidney disease quality of life-short form 1.3 in an African country. Transplantation Proceedings. 2014; 46(5):1295–301. http://dx.doi.org/10.1016/j.transproceed.2014.02.011. PMID: 24935292. https://doi.org/10.1016/j.transproceed.2014.02.011

47. Hays R. D. KJD, Mapes D. L., Coons S. J., Amin N., Carter W. B., Kamberg C. Kidney Disease Quality of Life Short Form (KDQOL-SF) Version 1.3: A manual for use and scoring. RAND. 1995.

48. Korevaar JC, Merkus MP, Jansen MAM, Dekker FW, Boeschoten EW, Krediet RT. Validation of the KDQOL-SFTM: A dialysis-targeted health measure. Quality of Life Research. 2002; 11(5):437–47. http://dx.doi.org/10.1023/A:1015631411960. PMID: 12113391.

49. Suet-Ching WL. The psychometric properties of the Chinese Dialysis Quality of Life Scale for Hong Kong dialysis patients. Journal of advanced nursing. 2001; 36(3):441–9. PMID: 11686759.

50. 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: the official journal of the National Kidney Foundation. 2001; 37(1):11–21. Epub 2001/01/03. PMID: 11136162.

51. Aiyasanon N, Premasathian N, Nimmannit A, Jetanavanich P, Sritippayawan S. Validity and reliability of CHOICE health experience questionnaire: Thai version. Journal of the Medical Association of Thailand. 2009; 92(9):1159–66. PMID: 19772174.

52. 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):226–40. http://dx.doi.org/10.1016/j.jpainsymman.2003.07.004. PMID: 15010101. https://doi.org/10.1016/j.jpainsymman.2003.07.004

53. Onsoz H. UYO. Reliability and validity of the Turkish version of the Dialysis Symptom Index in chronic hemodialysis patients. Turkish Nephrology, Dialysis and Tranplantation Journal. 2013; 22(1):60–7.

54. Davison SN, Jhangri GS, Johnson JA. Longitudinal validation of a modified Edmonton symptom assessment system (ESAS) in haemodialysis patients. Nephrology Dialysis Transplantation. 2006; 21 (11):3189–95. http://dx.doi.org/10.1093/ndt/gfl380. PMID: 2006504906.

55. 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. 2006; 69(9):1621–5. PMID: 16672923. https://doi.org/10.1038/sj.ki.5000184

56. Laupacis A, Muirhead N, Keown P, Wong C. A disease-specific questionnaire for assessing quality of life in patients on hemodialysis. Nephron. 1992; 60(3):302–6. Epub 1992/01/01. PMID: 1565182.

57. Alvarez-Ude F, Galan P, Vicente E, Alamo C, Fernandez-Reyes MJ, Badia X. Adaptation and validation of the Spanish version of 'Kidney Disease Questionnaire'. [Spanish]. Nefrologia. 1997; 17(6):486–96. PMID: 1998037857.

58. Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Development of the kidney disease quality of life (KDQOL) instrument. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 1994; 3(5):329–38. Epub 1994/10/01. PMID: 7841967.

59. 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):1099–111. http://dx.doi.org/10.1016/S0149-2918%2800%2980087-9. PMID: 11048907. https://doi.org/10.1016/S0149-2918(00)80087-9

60. Zengin N, Oren B, Gul A, Ustundag H. Assessment of quality of life in haemodialysis patients: A comparison of the Nottingham Health Profile and the Short Form 36. International Journal of Nursing Practice. 2014; 20(2):115–25. http://dx.doi.org/10.1111/ijn.12130. PMID: 24713007. https://doi.org/10.1111/ijn.12130

61. Badia X, Alonso J, Brosa M, Lock P. Reliability of the Spanish version of the Nottingham health profile in patients with stable end-stage renal disease. Social Science and Medicine. 1994; 38(1):153–8. http://dx.doi.org/10.1016/0277-9536%2894%2990310-7. PMID: 8146706.

62. Pakpour AH, Nourozi S, Molsted S, Harrison AP, Nourozi K, Fridlund B. Validity and reliability of short form-12 questionnaire in Iranian hemodialysis patients. Iran J Kidney Dis. 2011; 5(3):175–81. Epub 2011/04/29. PMID: 21525577.

63. Yang SC, Kuo PW, Wang JD, Lin MI, Su S. Development and pyschometric properties of the dialysis module of the WHOQOL-BREF Taiwan version. Journal of the Formosan Medical Association. 2006; 105(4):299–309. PMID: 16618610.

64. Ferrans CE, Powers MJ. Quality of life index: development and psychometric properties. Ans. 1985; Advances in nursing science. 8(1):15–24. 3933411.

65. Dehesh T, Zare N, Jafari P, Sagheb MM. Psychometric assessment of the Persian version of the Ferrans and Powers 3.0 index in hemodialysis patients. International Urology and Nephrology. 2014; 46 (6):1183–9. http://dx.doi.org/10.1007/s11255-013-0537-5. PMID: 23979815. https://doi.org/10.1007/s11255-013-0537-5

66. Halabi JO. Psychometric properties of the Arabic version of Quality of Life Index. Journal of Advanced Nursing. 2006; 55(5):604–10. http://dx.doi.org/10.1111/j.1365-2648.2006.03952_1.x. PMID: 16907792. https://doi.org/10.1111/j.1365-2648.2006.03952_1.x

67. Korkut Y. The reliability and validity study of the Turkish version of Ferrans and Powers' Quality of Life Index for dialysis patients. Archives of Neuropsychiatry. 2007; 44(14).

68. Mingardi G, Cornalba L, Cortinovis E, Ruggiata R, Mosconi P, Apolone G. Health-related quality of life in dialysis patients. A report from an Italian study using the SF-36 Health Survey. Nephrology Dialysis Transplantation. 1999; 14(6):1503–10. http://dx.doi.org/10.1093/ndt/14.6.1503. PMID: 1999183110.

69. Feurer ID, Moore DE, Speroff T, Liu H, Payne J, Harrison C, et al. Refining a health-related quality of life assessment strategy for solid organ transplant patients. Clinical Transplantation, Supplement. 2004; 18(12):39–45. PMID: 2004392633.

70. Franke GH, Reimer J, Kohnle M, Luetkes P, Maehner N, Heemann U. Quality of life in end-stage renal disease patients after successful kidney transplantation: development of the ESRD symptom checklist—transplantation module. Nephron. 1999; 83(1):31–9. Epub 1999/08/26. 45470. PMID: 10461033. https://doi.org/45470

71. Ortega T, Valdes C, Rebollo P, Ortega F. Evaluation of reliability and validity of Spanish version of the end-stage renal disease symptom checklist-transplantation module. Transplantation. 2007; 84 (11):1428–35. http://dx.doi.org/10.1097/01.tp.0000290231.39240.df. PMID: 18091518. https://doi.org/10.1097/01.tp.0000290231.39240.df

72. Stavem K, Ganss R. Reliability and validity of the ESRD Symptom Checklist—Transplantation Module in Norwegian kidney transplant recipients. BMC nephrology. 2006; 7:17. Epub 2006/11/18. https://doi.org/10.1186/1471-2369-7-17 PMID: 17109746; PubMed Central PMCID: PMCPmc1660533.

73. Cleemput I, Kesteloot K, Moons P, Vanrenterghem Y, Van Hooff JP, Squifflet JP, et al. The construct and concurrent validity of the EQ-5D in a renal transplant population. Value in Health. 2004; 7(4):499–509. https://doi.org/10.1111/j.1524-4733.2004.74013.x PMID: 15449642.

74. Moons P, De Geest S, Versteven K, Abraham I, Vlaminck H, Moens G, et al. Psychometric properties of the "Modified Transplant Symptom Occurrence and Symptom Distress Scale". Journal of nursing measurement. 2001; 9(2):115–34. PMID: 11696937.

75. 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):1223–32. http://dx.doi.org/10.1007/s11136-006-0053-5. PMID: 16972162. https://doi.org/10.1007/s11136-006-0053-5

76. Barotfi S, Molnar MZ, Almasi C, Kovacs AZ, Remport A, Szeifert L, et al. Validation of the Kidney Disease Quality of Life-Short Form questionnaire in kidney transplant patients. Journal of Psychosomatic

Research. 2006; 60(5):495–504. http://dx.doi.org/10.1016/j.jpsychores.2005.09.009. PMID: 16650590. https://doi.org/10.1016/j.jpsychores.2005.09.009

77. Niu Y, Zhang W, Mao S, Gao Y, Wang J, Li J, et al. Pilot feasibility research of Chinese version of kidney transplant questionnaire in recipients of living donor kidney transplantation. Int J Clin Exp Med. 2015; 8(12):22570–6. Epub 2016/02/18. PMID: 26885244; PubMed Central PMCID: PMCPmc4730030.

78. Laupacis A, Pus N, Muirhead N, Wong C, Ferguson B, Keown P. Disease-specific questionnaire for patients with a renal transplant. Nephron. 1993; 64(2):226–31. Epub 1993/01/01. PMID: 8321355.

79. Rebollo P, Ortega F, Ortega T, Valdes C, Garcia-Mendoza M, Gomez E. Spanish validation of the "kidney transplant questionnaire": a useful instrument for assessing health related quality of life in kidney transplant patients. Health and quality of life outcomes. 2003; 1:56. Epub 2003/11/14. https://doi.org/10.1186/1477-7525-1-56 PMID: 14613566; PubMed Central PMCID: PMCPmc269994.

80. Gentile S, Jouve E, Dussol B, Moal V, Berland Y, Sambuc R. Development and validation of a French patient-based health-related quality of life instrument in kidney transplant: The ReTransQoL. Health and quality of life outcomes. 2008; 6 (no pagination)(78). http://dx.doi.org/10.1186/1477-7525-6-78. PMID: 2008534167.

81. Beauger D, Gentile S, Jouve E, Dussol B, Jacquelinet C, Briancon S. Analysis, evaluation and adaptation of the ReTransQoL: A specific quality of life questionnaire for renal transplant recipients. Health and quality of life outcomes. 2013; 11 (1) (no pagination)(148). http://dx.doi.org/10.1186/1477-7525-11-148. PMID: 2013572155.

82. Almutary H, Bonner A, Douglas C. Arabic translation, adaptation and modification of the dialysis symptom index for chronic kidney disease stages four and five. BMC nephrology. 2015; 16(1):1–8. https://doi.org/10.1186/s12882-015-0036-2 PMID: 25884303

83. Churchill DN, Torrance GW, Taylor DW, Barnes CC, Ludwin D, Shimizu A, et al. Measurement of quality of life in end-stage renal disease: the time trade-off approach. Clin Invest Med. 1987; 10(1):14–20. Epub 1987/01/01. PMID: 3545580.

84. Chisholm-Burns MA, Erickson SR, Spivey CA, Gruessner RWG, Kaplan B. Concurrent validity of kidney transplant questionnaire in US renal transplant recipients. Patient Preference and Adherence. 2011; 5:517–22. http://dx.doi.org/10.2147/PPA.S24261. https://doi.org/10.2147/PPA.S24261 PMID: 22114465

85. Green J, Fukuhara S, Shinzato T, Miura Y, Wada S, Hays RD, et al. Translation, cultural adaptation, and initial reliability and multitrait testing of the Kidney Disease Quality of Life Instrument for use in Japan. Quality of Life Research. 2001; 10(1):93–100. http://dx.doi.org/10.1023/A:1016630825992. PMID: 11508479.

86. Kontodimopoulos N, Niakas D. Determining the basic psychometric properties of the Greek KDQOL-SFTM. Quality of Life Research. 2005; 14(8):1967–75. http://dx.doi.org/10.1007/s11136-005-3868-6. PMID: 16155785. https://doi.org/10.1007/s11136-005-3868-6

87. Kontodimopoulos N, Niakas D. Measuring health-related quality of life of Greek dialysis patients with the KDQOL-SFTM. [Greek]. Archives of Hellenic Medicine. 2007; 24(6):590–9. PMID: 2008072477.

88. Bataclan RP, Dial MAD. Cultural adaptation and validation of the Filipino version of Kidney Disease Quality of Life—Short Form (KDQOL-SF version 1.3). Nephrology. 2009; 14(7):663–8. http://dx.doi.org/10.1111/j.1440-1797.2009.01130.x. PMID: 19796025. https://doi.org/10.1111/j.1440-1797.2009.01130.x

89. Boini S, Leplege A, Loos Ayav C, Francais P, Ecosse E, Briancon S. Measuring quality of life in end-stage renal disease. Transcultural adaptation and validation of the specific Kidney Disease Quality of Life questionnaire. [French]. Nephrologie et Therapeutique. 2007; 3(6):372–83. http://dx.doi.org/10.1016/j.nephro.2007.05.005. PMID: 17919640. https://doi.org/10.1016/j.nephro.2007.05.005

90. Perneger TV, Leski M, Chopard-Stoermann C, Martin PY. Assessment of health status in chronic hemodialysis patients. Journal of Nephrology. 2003; 16(2):252–9. PMID: 12768073.

91. Yildirim A, Ogutmen B, Bektas G, Isci E, Mete M, Tolgay HI. Translation, cultural adaptation, initial reliability, and validation of the Kidney Disease and Quality of Life-Short Form (KDQOL-SF 1.3) in Turkey. Transplant Proc. 2007; 39(1):51–4. Epub 2007/02/06. 10.1016/j.transproceed.2006.10.196. PMID: 17275473. https://doi.org/10.1016/j.transproceed.2006.10.196

92. Molsted S, Heaf J, Prescott L, Eidemak I. Reliability testing of the Danish version of the Kidney Disease Quality of Life Short Form. Scand J Urol Nephrol. 2005; 39(6):498–502. Epub 2005/11/24. https://doi.org/10.1080/00365590500240253 PMID: 16303727.

93. Moreira CA, Garletti Junior W, Lima LF, Lima CR, Ribeiro JF, Miranda AF. Assessment of the basic psychometric properties for the portuguese version of the KDQOL-SFTM. [Portuguese]. Rev Assoc Med Bras. 2009; 55(1):22–8. PMID: 19360273.

94. Fardinmehr O, Farajzadegan Z, Naini AE, Mortazavi M, Gholamrezaei A. The validity and reliability of the persian version of kidney disease quality of life questionnaire-short form (KDQOL-SF) in Iranian patients. [Persian]. Journal of Isfahan Medical School. 2012; 29(165). PMID: 2014599849.

95. Vasilieva IA. Russian version of the questionnaire The Kidney Disease and Quality of Life Short Form (KDQOL-SF) a valuable diagnostic instrument for assessing quality of life of dialysis patients. Nefrologiya. 2007; 11(1):64–70.

96. Duarte PS, Ciconelli RM, Sesso R. Cultural adaptation and validation of the "Kidney Disease and Quality of Life—Short Form (KDQOL-SFTM 1.3)" in Brazil. Brazilian Journal of Medical and Biological Research. 2005; 38(2):261–70. PMID: 15785838. https://doi.org//S0100-879X2005000200015

97. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury. 2011; 42(3):236–40. Epub 2010/12/15. https://doi.org/10.1016/j.injury.2010.11.042 PMID: 21145544.

98. Terwee CB, Roorda LD, Knol DL, De Boer MR, De Vet HC. Linking measurement error to minimal important change of patient-reported outcomes. Journal of clinical epidemiology. 2009; 62(10):1062–7. Epub 2009/02/24. https://doi.org/10.1016/j.jclinepi.2008.10.011 PMID: 19230609.

99. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England). 1986; 1(8476):307–10. Epub 1986/02/08. PMID: 2868172.

100. Mokkink LB, Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L. et al. The COSMIN checklist manual. 2009.

101. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010; 10:22. Epub 2010/03/20. https://doi.org/10.1186/1471-2288-10-22 PMID: 20298572; PubMed Central PMCID: PMCPmc2848183.

102. Pakpour AH, Yekaninejad M, Molsted S, Harrison AP, Hashemi F, Saffari M. Translation, cultural adaptation assessment, and both validity and reliability testing of the Kidney Disease Quality of Life—Short Form version 1.3 for use with Iranian patients. Nephrology. 2011; 16(1):106–12. http://dx.doi.org/10.1111/j.1440-1797.2010.01389.x. PMID: 21175986. https://doi.org/10.1111/j.1440-1797.2010.01389.x

103. Bjordal K. Impact of quality of life measurement in daily clinical practice. Annals of oncology: official journal of the European Society for Medical Oncology. 2004; 15 Suppl 4:iv279–82. Epub 2004/10/13. 10.1093/annonc/mdh939. PMID: 15477321. https://doi.org/10.1093/annonc/mdh939

104. Breckenridge K, Bekker HL, Gibbons E, van der Veer SN, Abbott D, Briançon S, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. Nephrology Dialysis Transplantation. 2015. https://doi.org/10.1093/ndt/gfv209 PMID: 25982327

105. Elbers RG, Rietberg MB, van Wegen EE, Verhoef J, Kramer SF, Terwee CB, et al. Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2012; 21(6):925–44. Epub 2011/10/21. https://doi.org/10.1007/s11136-011-0009-2 PMID: 22012025; PubMed Central PMCID: PMCPmc3389599.

106. Green A, Liles C, Rushton A, Kyte DG. Measurement properties of patient-reported outcome measures (PROMS) in Patellofemoral Pain Syndrome: a systematic review. Man Ther. 2014; 19(6):517–26. Epub 2014/07/07. https://doi.org/10.1016/j.math.2014.05.013 PMID: 24997774.

107. Janssens A, Rogers M, Gumm R, Jenkinson C, Tennant A, Logan S, et al. Measurement properties of multidimensional patient-reported outcome measures in neurodisability: a systematic review of evaluation studies. Dev Med Child Neurol. 2016; 58(5):437–51. Epub 2015/12/15. https://doi.org/10.1111/dmcn.12982 PMID: 26661333; PubMed Central PMCID: PMCPmc5031226.

108. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2013; 22(8):1889–905. Epub 2013/01/05. https://doi.org/10.1007/s11136-012-0344-y PMID: 23288613.

109. Gilbody SM, Whitty PM, Grimshaw JM, Thomas RE. Improving the detection and management of depression in primary care. Quality & safety in health care. 2003; 12(2):149–55. Epub 2003/04/08. PMID: 12679514; PubMed Central PMCID: PMCPmc1743696. https://doi.org/10.1136/qhc.12.2.149

110. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2004; 22(4):714–24. Epub 2004/02/18. https://doi.org/10.1200/jco.2004.06.078 PMID: 14966096.

Appendices for Chapter 3	Page
Appendix 1 - Systematic review search strategy	209
Appendix 2 - Summary of study results	217
Appendix 3 - Characteristics of included studies	232
Appendix 4 - Characteristics of included PROMs	246

Chapter 4: Relationship between health-related quality of life, clinical outcomes of CKD and clinical variables – analysis of the RIISC dataset

Relationship between health-related quality of life, clinical outcomes of CKD and clinical variables – analysis of the RIISC dataset

One of the postulated benefits of using PROMs/ePROMs in the routine management of patients is its potential to provide early warnings of deterioration. Findings from oncology studies have shown that HRQOL data may signal serious adverse events in patients before clinical and/or laboratory data.[1 2] However, similar studies for CKD are limited.

The analysis of the RIISC dataset reported in this chapter, investigates the ability of HRQOL data (using EQ-5D) to predict CKD outcomes (composite outcome of all-cause mortality and progression to ESRD). The analysis also explores the relationship between HRQOL and clinical and laboratory variables such as eGFR, creatinine and albumin. Evidence of the predictive abilities of HRQOL data for CKD outcomes and its relationship with clinical and laboratory parameters may mean that patients at risk of rapid deterioration may be identified before their clinical data is available.

References

1. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. Journal of the National Cancer Institute 2011;**103**(24):1808-10.
2. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. Journal of the National Cancer Institute 2009;**101**(23):1624-32.

This chapter has been written in paper format for submission to a peer-reviewed journal.

Relationship between health-related quality of life domains and clinical outcomes in chronic kidney disease (CKD): an observational cohort study

Olalekan Lee Aiyegbusi[1, 2*], Derek Kyte[1, 2, 3], Paul Cockwell[1, 4], Tom Marshall[1, 2], Anthony Fenton[4], Stephanie Stringer[4], Melanie Calvert[1, 2, 3]

1. Centre for Patient Reported Outcomes Research, University of Birmingham, Edgbaston, Birmingham, B15 2TT

2. Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham, B15 2TT

3. NIHR Birmingham Biomedical Research Centre, University Hospitals Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK.

4. Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston, Birmingham, B15 2GW

*** Correspondence**: Olalekan Lee Aiyegbusi, oxa238@bham.ac.uk

Abstract

Background

Chronic kidney disease (CKD) is associated with poor health-related quality of life (HRQOL). The association between HRQOL and clinical outcomes (mortality and progression to end-stage renal disease (ESRD)) requires further investigation.

Methods

All 931 participants recruited between 2010 – 2015 into the Renal Impairment In Secondary Care (RIISC) study were included. Outcome data was collected up to December 31st 2017. Baseline demographic, clinical and laboratory data as well as an assessment of HRQOL using the EuroQol EQ-5D-3L were collected. Participants' health states were converted into an EQ-5D index score using UK-specific weighted preferences. Multivariable Cox proportional hazards regression was performed to evaluate the association of HRQOL with a composite primary outcome of progression to ESRD or all-cause mortality. Similar regression analyses were conducted for secondary endpoints of all-cause mortality or progression to ESRD. Results were expressed as a hazard ratio (HR) with 95% confidence interval. Logistic regression analyses were used to explore the relationship between the significant HRQOL dimensions and clinical variables.

Results

In multivariable analyses, the EQ-5D index score and dimensions were not significantly associated with the composite outcome. Problems with self-care and usual activities were associated with a higher risk of all-cause mortality [HR: 2.330, 95% CI: 1.635 - 3.319, p<0.001] and [HR: 1.772, 95% CI: 1.257 - 2.498, p=0.001] respectively. There were no significant associations with risk of progressing to ESRD. A higher level of albumin was significantly associated with lower odds of

reporting problems with self-care [OR 0.944, 95% CI 0.904 – 0.986, p=0.01] and usual activities [OR 0.956, 95% CI 0.924 – 0.989, p=0.01].

Conclusions

Patients with pre-dialysis CKD often report impaired HRQOL. In this study, self-care and usual activities were the dimensions of HRQOL significantly associated with all-cause mortality and progression to ESRD. A number of factors, including serum albumin, influence these two HRQOL dimensions.

INDEX WORDS: Chronic kidney disease; CKD; renal disease; interview; focus group; qualitative research; quality of life; QOL; electronic patient-reported outcome measures; ePROM.

Introduction

Chronic kidney disease (CKD) is a global public health issue[1] affecting up to 14% of the adult population,[2 3] with a recent review highlighting that prevalence may be stabilising.[4] Chronic kidney disease is a major risk factor for end-stage renal disease (ESRD), cardiovascular disease (CVD) and premature death.[5-10] The clinical outcomes of CKD are often poor and most patients die without progressing to ESRD.[11] Patients with CKD often experience a high symptom burden which may impair their health-related quality of life (HRQOL).[12-15]

While biochemical parameters such as estimated glomerular filtration rates (eGFR) and creatinine are traditionally used to assess patients' health status, it is increasingly acknowledged that patients' formal assessment of their HRQOL and/or symptoms should be incorporated in their care.[16] A patient-reported outcome (PRO) has been defined as "...any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else."[17] PROs can be measured using patient-reported outcome measures (PROMS) which may be generic or disease-specific.[18] Although disease-specific PROMs such as the KDQOL-36[19] provide more information on the impact of disease,[20 21] the preference-based EuroQol 5-Dimension (EQ-5D) questionnaire[22 23] is much shorter (thus reducing patient burden) and easier to use for economic evaluation.[24]

Although the association between reduced HRQOL and pre-dialysis CKD has been reported,[12 13 24] quantitative evidence of the relationship between HRQOL and clinical outcomes (all-cause mortality and progression to ESRD) in patients with pre-dialysis CKD is limited.[24-28] I examined the relationship between HRQOL (using EQ-5D data)

and all-cause mortality or the risk of progression to ESRD in pre-dialysis CKD using the Renal Impairment In Secondary Care (RIISC) Study dataset. I also explored the relationship between HRQOL and a number of clinical, laboratory and demographic variables in order to determine which are predictive of HRQOL.

Methods

Full details on the design and methods of the Renal Impairment In Secondary Care (RIISC) Study have been previously reported.[29] A summary of these is provided in this section.

Ethics approval

The RIISC study (ClinicalTrials.gov reference: NCT01722383) was approved by the South Birmingham Local Research Ethics Committee (reference: 10/H1207/6) and the University Hospitals Birmingham (UHB) Research and Development department (reference: RRK3917).

RIISC Overview

RIISC is a prospective, observational cohort study designed to evaluate factors that determine risk of CKD progression.[29] Demographic, clinical and laboratory data from electronic patient records of eligible patients with CKD are collated along with mortality data obtained from the Office of National Statistics (ONS). HRQOL data were obtained at study visits. Previous analysis on the RIISC dataset was conducted by Jesky et al.;[24] the analysis had 745 participants and fewer clinical events had occurred by the censor date than is reported here.

Eligibility criteria

Patients were eligible if they had stage 3 CKD with a decline in eGFR of ≥5 ml/min/year or ≥10 ml/min/5 years or an albumin creatinine ratio (ACR) ≥70 mg/mmol on three consecutive occasions, or stage 4/5 CKD.[29] CKD staging for recruitment was based on eGFR calculated using the Modification of Diet in Renal Disease (MDRD) equation.[30] Patients receiving renal replacement therapy (dialysis

or renal transplant) or immunosuppressive therapy were ineligible for recruitment.[24 29]

Recruitment

A total of 931 patients were recruited from nephrology clinics in two hospitals between October 2010 and December 2015. Written informed consent was provided by all participants and the study is being conducted in compliance with the Declaration of Helsinki.[31]

Data collection

Demographic, clinical, and laboratory data were collected at baseline and at six-month, 18-month and three-year study visits. Outcome data were collected up to 31st December 2017. Socio-economic status (SES) was assessed using the Index of Multiple Deprivation (IMD 2010)[32]; lower scores indicate greater deprivation. Laboratory tests were performed according to current standards.

Quality of life data were collected by administering the EQ-5D-3L at recruitment and follow-up. The EQ-5D-3L is a generic HRQOL questionnaire comprising of five questions each with three response levels and a visual analogue self-rating (VAS) scale.[33] The EQ-5D has been widely validated in numerous patient groups including patients with renal diseases. Responses to the questionnaire may be used to generate a 5-number health state for each respondent. These health states may be converted into index scores ranging from −0.594 to 1.0 (a value of 1 represents full health, 0 represents death while lower values indicate HRQOL worse than death) using a value set specific for many countries including the UK.[23]

Statistical analysis

Analyses were performed using Stata 12 (Statacorp, College Station, Texas, USA). There was little missing data therefore complete case analyses were performed,

unless otherwise stated. A p-value of <0.05 was considered to be statistically significant.

A priori variables

The *a priori* selection of variables was based on their biological plausibility and previously published evidence of significant associations with the outcomes of interest and/or quality of life. The selected covariates included socio-demographic variables (gender, age, ethnicity, educational status, employment, socio-economic status)[34-42]; clinical parameters (CKD stage, co-morbidities)[34 43 44] and laboratory parameters (eGFR, albumin, creatinine, albumin-to-creatinine ratio [ACR], haemoglobin).[13 34 36 42 45-50]

Outcomes of interest

The primary outcome of interest was a composite outcome of all-cause mortality or progression to ESRD. Progression to ESRD and all-cause mortality were analysed separately as secondary endpoints.

Descriptive statistics

Baseline characteristics are presented as frequency and percentage for categorical variables, and as mean with standard deviation (SD) or median with inter-quartile range (IQR) for continuous variables (depending on distribution). Chi-squared tests were used to compare categorical variables. Continuous variables were compared using ANOVA (parametric distribution) or Kruskal-Wallis (non-parametric distribution).

Statistical Modelling

Univariable analyses were initially performed for each of the a priori variables. Variables with a p-value of <0.1 were subsequently included in multivariable analyses. Each significant EQ-5D variable was combined separately with other

significant variables from the univariable analyses. A backward stepwise method was used for the multivariable analyses to determine the optimized models. Statistically significant variables in the multivariable models ($P<0.05$) were kept in the final models.

Survival analyses

Cox proportional hazards analyses were performed to examine the association between the three outcomes of interest (composite outcome, progression to ESRD and all-cause mortality) and the a priori covariates (including the individual EQ-5D dimensions and the EQ-5D index score). The censor date was 31st December 2017. The proportional hazards assumption was assessed graphically. Data are presented using hazard ratios (HR) with 95% confidence intervals (CI), p-values and survival plots.

Competing Risk Analyses

In the standard Cox regression model for the ESRD analysis, censoring patients who die without ESRD makes the incorrect assumption that those patients remain at risk of progression to ESRD. Therefore, a competing risk analysis was performed to assess the association between the a priori variables and risk of progression to ESRD taking into account the competing risk of death. A sub-distribution hazards approach was used as described in literature.[24 51 52]

Predictors of HRQOL

Logistic regressions were performed in order to determine which of the a priori variables predicted problems in the EQ-5D dimensions. Data for the EQ-5D dimensions was dichotomised as "no problems" and "some problem" (which

combined patients with moderate and severe problems). Pseudo R^2 values and odds ratios (with 95% CI and p-values) are reported.

Results

Descriptive statistics

All participants (n=931) recruited by the censor date (31st December 2017) were included in the analyses. Baseline demographic, clinical and laboratory data are provided (Table 1). Median age was 64.2 years (IQR 51.8-76.2 years), 62% were male and 41.9% had stage 4 CKD. 68% were white, 21.6% South Asian, 9.5% Black and 0.9% from other ethnicities. There was no significant difference in ethnicity by CKD stage (p = 0.204). 52.7% of participants were in the most deprived quintile nationally (IMD 2010). There was no significant difference in CKD stage between participants from the least deprived and the most deprived quintile (p = 0.774). The percentage of patients who reported problems for the EQ-5D dimensions were: 58% for mobility, 16% for self-care, 48% for usual activities, 55% for pain/discomfort, and 30% for anxiety/depression. Of the 152 patients who reported problems with self-care, 95% also reported problems with usual activities (p<0.001). There were a total of 479 composite outcome events, 307 (64.1%) had progressed to ESRD (commenced RRT) while 172 (35.9%) had died.

Table 1. Baseline demographic, clinical and laboratory data. Data completeness (DC) and other categorical variables are expressed as percentages (%); and continuous variables as mean (SD) or median (IQR).

Variables	Sample size, *n* = (931)	Data completeness (%)
DEMOGRAPHIC		
Age (median years, IQR)	64.6 (51.8-76.2)	100
Gender – Female (%)	38.5	100
Ethnicity		
White (%)	68.0	100
South Asian (%)	21.6	
Black (%)	9.5	
Other (%)	0.9	
SES (IMD 2010)		
Score (median, IQR)	31.9 (17.6-46.5)	90.9
% in most deprived national quintile	52.7	99.5
Educational Attainment		
No formal qualifications (%)	45.7	98.8
GCSE/O' Level, NVQ, A' Level (%)	37.8	
Undergraduate (%)	10.0	
Postgraduate (%)	5.3	
Current Employment Status		
Not in employment (%)	19	99.6
Employed (%)	27.8	
Retired (%)	52.7	
Job Type (when last employed)		
None (%)	13.2	100
Unskilled manual (%)	17.5	
Skilled manual (%)	33.2	
Clerical (%)	11.2	
Managerial (%)	9.0	
Professional (%)	15.9	
Smoking Status (%)		
Never (%)	47.1	98.2
Current (%)	12.9	
Previous (%)	38.2	
Alcohol Consumption		
None (%)	58.5	100
<10 units (%)	28.6	
11-20 units (%)	8.1	
>20 units (%)	4.8	
BMI (Kg/m^2)	29.7 (14.7-60.0)	95.8
CLINICAL		
CKD Stage		
1 & 2	8.6	96.6
3a	11.4	
3b	29.0	
4	41.9	
5	5.7	
Comorbidities		
Malignancy (%)	14.1	
Diabetes (%)	35.3	

Variables	Sample size, *n* = (931)	Data completeness (%)
COPD (%)	9.1	100
Cerebrovascular disease (%)	10.8	
IHD (%)	21.5	
PVD (%)	9.2	
Blood Pressure (BP)		
Systolic BP (mmHg)	131.0 (20.3)	97.5
Diastolic BP (mmHg)	76.2 (12.5)	97.4
BIOMARKERS		
Creatinine (µmol/L)	234.8 (96.5)	96.6
eGFR MDRD (ml/min/1.73m^2)	27.2 (14.9)	96.6
ACR (mg/mmol)	33.9 (6.5-133.0)	93.7
Haemoglobin (g/L)	123.4 (17.2)	94.3
Bicarbonate (mmol/L)	23.8 (3.6)	97.2
Albumin (g/L)	42.3 (4.6)	96.2
White Cell Count (g/L)	7.6 (2.3)	94.3
EQ-5D INDEX SCORES (means)		
1	0.687 (0.27)	94.8
2	0.758 (0.20)	
3a	0.694 (0.28)	
3b	0.696 (0.27)	
4	0.689 (0.25)	
5	0.702 (0.29)	
Overall score	0.696 (0.26)	

SES - Socio-economic status
IMD - Index of Multiple Deprivation
BMI - Body Mass Index
IHD – Ischemic Heart Disease
PVD – Peripheral Vascular Disease
COPD – Chronic Obstructive Pulmonary Disease
eGFR - estimated Glomerular Filtration Rate
MDRD equation - Modification of Diet in Renal Disease equation
ACR - Albumin creatinine ratio

Relationship between HRQOL and clinical outcomes

Composite outcome (all-cause mortality or progression to ESRD)

Results for the univariable analyses are presented in Table 2. Age, ethnicity, eGFR, ACR, albumin, haemoglobin, co-morbidities (diabetes, CVD, IHD, PVD), reported problems with self-care and usual activities, and EQ-5D index score were significantly associated with the composite outcome.

In the multivariable analyses, reported problems with self-care and usual activities, and EQ-5D index score were not significantly associated with the composite outcome (Table 3). A decrease in eGFR, serum albumin, and haemoglobin; an increase in ACR; and the presence of co-morbidity (IHD) were significantly associated with an increased risk of the composite outcome.

Table 2. Univariable Survival Analyses (Cox regression) for the composite outcome				
	Hazard ratio	95% CI		p-value
		Lower	Higher	
Age (per 10 year increase)	1.099	1.036	1.165	**0.002**
Gender (female as reference)	1.113	0.924	1.340	0.258
Ethnicity (white as reference)				
South Asian	1.228	0.985	1.530	**0.068**
Black/other	1.427	1.082	1.882	**0.012**
Education (some versus none)	1.085	0.905	1.300	0.381
Deprivation (most versus least)	0.947	0.791	1.134	0.554
eGFR (per 5ml/min increase)	0.674	0.634	0.717	**<0.001**
ACR (mg/mmol)	1.003	1.002	1.003	**<0.001**
Albumin (g/L)	0.917	0.899	0.935	**<0.001**
Haemoglobin (g/L)	0.976	0.970	0.981	**<0.001**
Co-morbidity				
Diabetes	1.208	1.006	1.450	**0.043**
CVD	1.268	0.976	1.647	**0.076**
COPD	0.839	0.615	1.144	0.266
IHD	1.610	1.322	1.960	**<0.001**
PVD	1.385	1.059	1.813	**0.018**
Malignancy	1.211	0.956	1.535	0.113
EQ-5D components				
Mobility	0.907	0.757	1.085	0.285
Self-care	1.379	1.090	1.744	**0.007**
Usual activities	1.321	1.103	1.582	**0.002**
Pain/Discomfort	1.097	0.915	1.314	0.317
Anxiety/Depression	1.148	0.942	1.400	0.172
EQ-5D index score	0.710	0.511	0.988	**0.042**

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, COPD - Chronic obstructive pulmonary disease, IHD - Ischemic heart disease, PVD - Peripheral vascular disease, CVD - Cerebro-vascular disease

Table 3. Multivariable Survival Analyses (Cox regression) for the composite outcome				
	Hazard ratio	95% CI		p-value
		Lower	Higher	
Problems with self-care	NS	NS	NS	NS
eGFR (per 5ml/min increase)	0.693	0.650	0.739	<0.001
ACR (mg/mmol)	1.003	1.002	1.003	<0.001
Albumin (g/L)	0.968	0.945	0.991	0.007
Haemoglobin (g/L)	0.988	0.982	0.995	0.001
IHD	1.560	1.244	1.956	<0.001
Problems with usual activities	NS	NS	NS	NS
eGFR (per 5ml/min increase)	0.693	0.650	0.739	<0.001
ACR (mg/mmol)	1.003	1.002	1.003	<0.001
Albumin (g/L)	0.968	0.945	0.991	0.007
Haemoglobin (g/L)	0.988	0.982	0.995	0.001
IHD	1.560	1.244	1.956	<0.001
EQ-5D index score	NS	NS	NS	NS
eGFR (per 5ml/min increase)	0.693	0.650	0.740	<0.001
ACR (mg/mmol)	1.003	1.002	1.003	<0.001
Albumin (g/L)	0.969	0.946	0.992	0.01
Haemoglobin (g/L)	0.988	0.982	0.995	0.001
IHD	1.592	1.267	1.999	<0.001

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, IHD - Ischemic heart disease, NS - not significant

All-cause mortality

Sixteen of the *a priori* variables were associated with all-cause mortality in the univariable analyses (Table 4). In the multivariable analyses (Table 5), reported problems with self-care [HR: 2.330, 95% CI: 1.635 – 3.319, p<0.001] and usual activities [HR: 1.772, 95% CI: 1.257 – 2.498, p=0.001], lower serum albumin, older age and the presence of co-morbidities (IHD and malignancy) were associated with an increased risk of death.

Table 4. Univariable survival analyses (Cox regression) for all-cause mortality				
	Hazard ratio	95% CI		p-value
		Lower	Higher	
Age (per 10 year increase)	2.177	1.887	2.510	**<0.001**
Gender (female as reference)	1.191	0.871	1.627	0.274
Ethnicity (white as reference)				
South Asian	0.431	0.266	0.696	**0.001**
Black/other	0.747	0.445	1.255	0.271
Education (some versus none)	0.450	0.328	0.616	**<0.001**
Deprivation (most versus least)	0.907	0.672	1.222	0.520
eGFR (per 5ml/min increase)	0.906	0.846	0.970	**0.005**
ACR(mg/mmol)	0.999	0.997	1.000	**0.051**
Albumin (g/L)	0.964	0.934	0.995	**0.024**
Haemoglobin (g/L)	0.985	0.976	0.994	**0.001**
Co-morbidity				
Diabetes	1.573	1.167	2.122	**0.003**
CVD	1.593	1.069	2.374	**0.022**
COPD	1.685	1.117	2.541	**0.013**
IHD	2.746	2.028	3.716	**<0.001**
PVD	1.913	1.290	2.836	**0.001**
Malignancy	2.619	1.888	3.634	**<0.001**
EQ-5D components				
Mobility	0.954	0.707	1.286	0.756
Self-care	2.625	1.883	3.660	**<0.001**
Usual activities	2.756	1.993	3.812	**<0.001**
Pain/Discomfort	1.246	0.919	1.689	0.157
Anxiety/Depression	1.314	0.953	1.811	**0.095**
EQ-5D index score	0.464	0.279	0.770	**0.003**

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, COPD - Chronic obstructive pulmonary disease, IHD - Ischemic heart disease, PVD - Peripheral vascular disease, CVD - Cerebro-vascular disease

Table 5. Multivariable survival analyses (Cox regression) for all-cause mortality				
	Hazard ratio	95% CI		p-value
		Lower	Higher	
Problems with self-care	2.330	1.635	3.319	<0.001
Age (per 10 year increase)	2.117	1.802	2.488	<0.001
Albumin (g/L)	0.967	0.934	1.001	0.058
IHD	1.672	1.213	2.305	0.002
Malignancy	1.588	1.109	2.273	0.012
Problems with usual activities	1.772	1.257	2.498	0.001
Age (per 10 year increase)	2.022	1.725	2.371	<0.001
Albumin (g/L)	0.966	0.934	0.999	0.044
IHD	1.676	1.216	2.309	0.002
Malignancy	1.527	1.068	2.185	0.020
EQ-5D index score	NS	NS	NS	NS
Age (per 10 year increase)	2.136	1.817	2.512	<0.001
Albumin (g/L)	0.962	0.929	0.995	0.025
IHD	1.737	1.251	2.411	0.001
Malignancy	1.503	1.045	2.160	0.028

IHD - Ischemic heart disease, NS - not significant

Progression to ESRD

Univariable analyses of age, ethnicity, education, eGFR, ACR, albumin, haemoglobin, co-morbidities (COPD, malignancy), and reported problems with self-care and usual activities had p-values <0.1 (Table 6). These variables were incorporated in the multivariable analyses (Table 7). Reported problems with self-care [HR: 0.609, 95% CI: 0.409 – 0.905, p=0.014] and usual activities [HR: 0.705, 95% CI: 0.542 – 0.915, p=0.009] and increases in age, eGFR, serum albumin, haemoglobin were associated with a reduction in the likelihood of progressing to ESRD. Black/other ethnicity and a higher ACR were associated with an increased risk of progression to ESRD.

Table 6. Univariable survival analyses (Cox regression) for progression to ESRD				
	Hazard ratio	95% CI		p-value
		Lower	Higher	
Age (per 10 year increase)	0.820	0.768	0.875	**<0.001**
Gender (female as reference)	1.016	0.807	1.281	0.890
Ethnicity (white as reference)				
South Asian	1.933	1.494	2.501	**<0.001**
Black/other	1.898	1.363	2.643	**<0.001**
Education (some versus none)	1.472	1.166	1.859	**0.001**
Deprivation (most versus least)	0.993	0.793	1.244	0.952
eGFR (per 5ml/min increase)	0.619	0.569	0.672	**<0.001**
ACR(mg/mmol)	1.003	1.002	1.003	**<0.001**
Albumin (g/L)	0.920	0.898	0.942	**<0.001**
Haemoglobin (g/L)	0.979	0.972	0.986	**<0.001**
Co-morbidity				
Diabetes	0.950	0.753	1.199	0.667
CVD	0.964	0.681	1.365	0.836
COPD	0.450	0.276	0.734	**0.001**
IHD	0.917	0.700	1.201	0.528
PVD	0.954	0.658	1.383	0.804
Malignancy	0.578	0.401	0.834	**0.003**
EQ-5D components				
Mobility	0.904	0.722	1.132	0.381
Self-care	0.721	0.509	1.022	**0.066**
Usual activities	0.775	0.616	0.974	**0.029**
Pain/Discomfort	0.977	0.780	1.225	0.843
Anxiety/Depression	0.976	0.758	1.256	0.848
EQ-5D index score	1.142	0.732	1.784	0.558

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, COPD - Chronic obstructive pulmonary disease, IHD - Ischemic heart disease, PVD - Peripheral vascular disease, CVD - Cerebro-vascular disease

Table 7. Multivariable survival analyses (Cox regression and Competing risk) for progression to ESRD								
	Cox analyses				Competing risk analyses			
	HR	95% CI		p-value	SHR	95% CI		p-value
		Lower	Higher			Lower	Higher	
Problems with self-care	0.609	0.409	0.905	0.014	0.584	0.370	0.922	0.021
Age (per 10 year increase)	0.676	0.623	0.733	<0.001	0.697	0.633	0.767	<0.001
eGFR (per 5ml/min increase)	0.612	0.559	0.670	<0.001	0.622	0.549	0.705	<0.001
Albumin (g/L)	0.961	0.934	0.989	0.007	NS	NS	NS	NS
ACR (mg/mmol)	1.002	1.002	1.003	<0.001	1.002	1.001	1.004	<0.001
Haemoglobin (g/L)	0.987	0.978	0.996	0.003	0.987	0.979	0.996	0.004
Problems with usual activities	0.705	0.542	0.915	0.009	0.698	0.515	0.946	0.020
Age (per 10 year increase)	0.687	0.632	0.746	<0.001	0.705	0.640	0.777	<0.001
eGFR (per 5ml/min increase)	0.608	0.556	0.665	<0.001	0.618	0.546	0.698	<0.001
Albumin (g/L)	0.964	0.937	0.992	0.013	NS	NS	NS	NS
ACR (mg/mmol)	1.002	1.002	1.003	<0.001	1.002	1.001	1.004	<0.001
Ethnicity (white as reference)								
South Asian	NS	NS	NS	NS				
Black/other	1.631	1.108	2.401	0.013	NS	NS	NS	NS
Haemoglobin (g/L)	0.988	0.979	0.997	0.007	0.988	0.979	0.997	0.007

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, HR – Hazard ratio, SHR – sub-distribution hazard ratio, NS – Not significant

Competing risk analyses

Findings from the competing risk analyses were also reported in Table 7. Reported problems with self-care [SHR: 0.584, 95% CI: 0.370 – 0.922, p=0.021] and usual activities [SHR: 0.698, 95% CI: 0.515 – 0.946, p=0.020], age, eGFR, ACR and haemoglobin remained significant in the competing risk models.

Predictors of HRQOL

The survival analyses above show that problems with self-care and usual activities are associated with an increased risk of all-cause mortality while problems with usual activities are linked to a decrease in progression to ESRD. Logistic regressions were conducted to explore which demographic, clinical or laboratory variables are predictive of these two EQ-5D components.

Self-care

Table 8 reports the findings from the logistic regression for self-care. In the final multivariable model, unemployment; retirement; co-morbidities (diabetes, COPD); and a higher BMI were significantly associated with an increased likelihood of reporting problems with self-care [unemployed: OR 11.608, 95% CI 4.389 – 30.700, p<0.001; retired: OR 9.102, 95% CI 3.594 – 23.051, p<0.001; diabetes: OR 1.806, 95% CI 1.182 – 2.759, p=0.006; COPD: OR 2.233, 95% CI 1.301 – 3.832, p=0.004; BMI: OR 1.044, 1.013 – 1.076, p=0.005]. A higher level of albumin was significantly associated with a slightly lower odds of reporting problems with self-care [OR 0.944, 95% CI 0.904 – 0.986, p=0.01]. The model explains 19.6% of the variability in the EQ-5D self-care dimension (Nagelkerke R^2) and correctly classifies 85.1% of cases.

Table 8. Logistic regression analyses for self-care								
	Univariable analyses				Multivariable analyses			
	Odds ratio	95% CI		p-value	Odds ratio	95% CI		p-value
		Lower	Higher			Lower	Higher	
Age (per 10 year increase)	1.188	1.057	1.335	**0.004**				
Gender (female as reference)	0.919	0.644	1.311	0.642				
Ethnicity (white as reference)								
South Asian	1.838	1.234	2.738	**0.003**				
Black/other	1.411	0.807	2.468	0.227				
Education (some versus none)	0.482	0.337	0.689	**<0.001**				
Deprivation (most versus least)	0.650	0.457	0.923	**0.016**				
Employment (currently employed as reference)								
Unemployed	11.552	5.058	26.383	**<0.001**	11.608	4.389	30.700	<0.001
Retired	9.323	4.264	20.384	**<0.001**	9.102	3.594	23.051	<0.001
eGFR (per 5ml/min increase)	1.002	0.944	1.063	0.959				
ACR (mg/mmol)	1.000	0.999	1.001	0.532				
Albumin (g/L)	0.935	0.901	0.970	**<0.001**	0.944	0.904	0.986	0.010
Haemoglobin (g/L)	0.988	0.978	0.999	**0.026**				
Co-morbidity								
Diabetes	2.604	1.826	3.712	**<0.001**	1.806	1.182	2.759	0.006
CVD	1.312	0.793	2.169	0.290				
COPD	2.747	1.710	4.414	**<0.001**	2.233	1.301	3.832	0.004
IHD	1.947	1.339	2.832	**<0.001**				
PVD	1.422	0.840	2.409	0.190				
Malignancy	1.081	0.675	1.731	0.746				
BMI	1.068	1.041	1.096	**<0.001**	1.044	1.013	1.076	0.005

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, CVD - Cerebro-vascular disease, COPD - Chronic obstructive pulmonary disease, IHD - Ischemic heart disease, PVD - Peripheral vascular disease, BMI - Body Mass Index

Usual activities

Table 9 reports the findings from the logistic regression for usual activities. The final multivariable model found that unemployment, retirement and co-morbidities (Diabetes, CVD, IHD) were significantly associated with an increased likelihood of reporting problems with usual activities [unemployed: OR 6.475, 95% CI 4.040 – 10.377, p<0.001; retirement: OR 4.419, 95% CI 2.948 – 6.623, p<0.001; diabetes: OR 1.951, 95% CI 1.422 – 2.678, p<0.001; CVD: OR 1.789, 95% CI 1.116 – 2.868, p=0.016, IHD: OR 1.558, 95% CI 1.076 – 2.257, p=0.019]. Men were less likely to report problems with usual activities [OR 0.636, 95% CI 0.465 – 0.871, p=0.005] and a higher level of serum albumin was associated with a reduced likelihood of reporting problems with usual activities [OR 0.956, 95% CI 0.924 – 0.989, p=0.01]. The model explains 26% of the variance in EQ-5D usual activities dimension (Nagelkerke R^2) and correctly classifies 67.5% of cases.

Table 9. Logistic regression analyses for Usual Activities								
	Univariable analyses				Multivariable analyses			
	Odds ratio	95% CI		p-value	Odds ratio	95% CI		p-value
		Lower	Higher			Lower	Higher	
Age (per 10 year increase)	1.002	0.983	1.022	0.832				
Gender (female as reference)	0.772	0.654	0.911	**0.002**	0.636	0.465	0.871	0.005
Ethnicity (white as reference)								
South Asian	0.980	0.743	1.293	0.888				
Black/other	0.846	0.566	1.264	0.415				
Education (some versus none)	1.250	1.032	1.514	**0.023**				
Deprivation (most versus least)	0.766	0.640	0.917	**0.004**				
Employment (currently employed as reference)								
Unemployed	1.667	1.228	2.261	**0.001**	6.475	4.040	10.377	<0.001
Retired	1.330	1.112	1.592	**0.002**	4.419	2.948	6.623	<0.001
eGFR (per 5ml/min increase)	0.973	0.952	0.994	**0.012**				
ACR (mg/mmol)	1.000	0.999	1.001	0.597				
Albumin (g/L)	0.997	0.994	1.000	**0.035**	0.956	0.924	0.989	0.010
Haemoglobin (g/L)	0.999	0.998	1.000	**0.011**				
Co-morbidity								
Diabetes	1.514	1.227	1.869	**<0.001**	1.951	1.422	2.678	<0.001
CVD	2.000	1.348	2.968	**0.001**	1.789	1.116	2.868	0.016
COPD	1.735	1.138	2.646	**0.010**				
IHD	1.821	1.381	2.400	**<0.001**	1.558	1.076	2.257	0.019
PVD	1.794	1.179	2.729	**0.006**				
Malignancy	1.072	0.773	1.489	0.676				
BMI	0.998	0.994	1.003	0.417				

eGFR - Estimated glomerular filtration rate, ACR - Albumin creatinine ratio, CVD - Cerebro-vascular disease, COPD - Chronic obstructive pulmonary disease, IHD - Ischemic heart disease, PVD - Peripheral vascular disease, BMI - Body Mass Index

Discussion

Summary of findings

This study confirms that impairment of HRQOL is associated with the risk of death and progression to ESRD. The EQ-5D index score or any of the dimensions was not significantly associated with the composite outcome (progression to ESRD or all-cause mortality). However, a decrease in eGFR, serum albumin, and haemoglobin; an increase in ACR; and the presence of co-morbidity (IHD) were significantly associated with an increased risk of the composite outcome.

Problems with self-care and usual activities, older age, a lower serum albumin and the presence of co-morbidities (IHD and malignancy) were significantly associated with an increased risk of all-cause mortality. The risk of progression to ESRD was found to be lower in participants who reported problems with self-care and usual activities, and older patients. A lower eGFR, serum albumin, haemoglobin, and Black/other ethnicity were associated with a significantly higher risk of progression to ESRD. Reported problems with self-care and usual activities, age, eGFR, ACR and haemoglobin remained significant in the competing risk analyses for progression to ESRD (with all-cause mortality as the competing risk). It should be noted that commencement of RRT was our indicator for progression to ESRD.

Further analyses were conducted to explore the relationships between the EQ-5D dimensions (self-care and usual activities) and *a priori* demographic, clinical and laboratory variables. Unemployment, retirement, the presence of co-morbidities (diabetes & COPD) and higher BMI were associated with reported problems with self-care. An increase in serum albumin was found to be associated with a minimal reduction in the likelihood of reported problems with self-care. Unemployment,

retirement and the presence of co-morbidities (diabetes, CVD and IHD) were also associated with an increased likelihood of problems with usual activities. Being male and having a higher level of serum albumin was associated with a lower likelihood of reporting problems with usual activities.

Findings in context of existing literature

We found a significant relationship between impaired HRQOL and death. Reported problems for the EQ-5D dimensions of self-care and usual activities were associated with an increased risk of all-cause mortality. Similar findings have been reported by previous studies.[24 26 28 53]

We found a statistically significant association between impaired HRQOL and progression to ESRD. From our analyses the EQ-5D dimensions of self-care and usual activities were associated with a reduced likelihood of progression to ESRD (commencement of RRT). Both dimensions remained significant in the competing risk analyses. A possible explanation for this unexpected finding is that some patients, with eGFR <9 (ml/min/1.73m^2), who reported problems for these HRQOL dimensions, were unable to cope with the additional burden of RRT and so chose conservative management instead. This might also explain why the analyses showed older patients have a reduced risk of progressing to ESRD. Older patients may experience frailty issues which might make them unsuitable for RRT.

The evidence on the relationship between impaired HRQOL and progression to ESRD in existing literature is by no means straightforward. Tsai et al. found that adjusted regression analysis of the total, physical and psychological domain scores of the WHOQOL-BREF(TW) predicted dialysis (ESRD).[28] Porter et al. reported that their fully adjusted regression model found an association was found between lower *time-varying* and *cumulative* mental health composite scores of the KDQOL-36 and

CKD progression and the composite outcome of CKD progression/death.[27] However, recent studies by Jesky et al. and Grove et al. (both of which utilised the EQ-5D) reported no association between HRQOL and progression to ESRD.[24 25] Grove et al. found a link between self-rated health (using a single item from SF-36) and the initiation of dialysis.[25]

It should be noted that the studies by Tsai et al. and Porter et al. were conducted specifically within Taiwanese[28] and African-American populations[27] respectively, limiting their generalizability. The study by Jesky et al. was limited by its shorter follow-up which meant that fewer events had occurred by the censor date[24] while the study by Grove et al. was limited by overall sample size and number of events as reflected in the wide confidence intervals reported.[25]

Our findings on the socio-demographic predictors of HRQOL dimensions of self-care and usual activities were similar to those reported by Jesky et al.[24] However, we found significant associations between serum albumin and both dimensions of HRQOL.

Evidence from oncology research suggests that patient reporting could detect potentially serious adverse events before clinician reporting.[54 55] Therefore the existence of significant relationships between patient-reported HRQOL data and clinical parameters could mean that patients with advanced CKD at risk of rapid deterioration may be identified through their PROM data before clinical and/or laboratory data are available.

Strengths and limitations

Strengths

This study employs a relatively large, ethnically diverse sample which is fairly representative of the general population of Birmingham and other major UK cities,

thus improving the generalizability of its findings. A variety of statistical methods were used to explore the data thus strengthening the validity of the results. The longer follow up period meant that the robustness of the analyses were higher as more events had occurred by the censor date. It can be logically assumed that the use of the brief EQ-5D reduced patient burden and positively influenced data completion and retention rates.

Limitations

This study like all observational studies explored association and not causation. As the study employs a self-reported measure (EQ-5D) to assess HRQOL, we cannot rule out the presence of social desirability bias.[53] The commencement of RRT was our indicator for progression to ESRD. However, some patients with eGFR <9 (ml/min/1.73m^2) might opt for conservative management instead of RRT. The lack of data on conservative management means that we could not confirm our explanation for the unexpected association between self-care and usual activities and progression to ESRD. Finally, the analyses did not include other factors which have been reported to influence HRQOL such as frailty and nutritional status.[24 56-58]

Conclusion

This study explored the relationship between HRQOL and clinical outcomes in patients at risk of CKD progression. Our analyses show that patients with CKD often experience impaired HRQOL which has a significant relationship with all-cause mortality and progression to ESRD. A number of demographic, clinical and laboratory variables were also found to be predictive of impaired HRQOL. These findings strengthen and extend the existing evidence around the relevance and the potential role for PROMs in routine renal care.

Contributors

MC is the guarantor for this study which was conceived and designed by OLA, MC, DK, PC, TM and AF. The analyses were conducted by OLA and reviewed by AF. OLA drafted the manuscript. OLA, MC, DK, PC, TM, AF and SS revised the manuscript. All authors approved the final draft.

Funding

This project is funded as part of the Health Foundation's Improvement Science Programme (Grant reference: 7452). The Health Foundation is an independent charity working to improve the quality of healthcare in the UK. The Health Foundation was not involved in any other aspect of the project.

Tom Marshall is partly funded by the National Institute for Health Research (NIHR) through the Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC-WM). Derek Kyte is funded by the NIHR post-doctoral research fellowship programme. Melanie Calvert is partly funded by the NIHR Birmingham Biomedical Research Centre and the NIHR Surgical Reconstruction and Microbiology Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. This paper presents independent research and the views expressed in this publication are not necessarily those of the NIHR, the Department of Health, NHS Partner Trusts, University of Birmingham or the CLAHRC WM Management Group.

Competing interests: Calvert reports personal fees from Astellas and Takeda; grants from Innovate UK, NIHR, and Macmillan Cancer Support, outside the submitted work.

References

1. Jha V, Garcia-Garcia G, Iseki K, et al. Chronic kidney disease: global dimension and perspectives. The Lancet 2013;**382**(9888):260-72.
2. Mills KT, Xu Y, Zhang W, et al. A systematic analysis of worldwide population-based data on the global burden of chronic kidney disease in 2010. Kidney international 2015;**88**(5):950-7.
3. Hill NR, Fatoba ST, Oke JL, et al. Global Prevalence of Chronic Kidney Disease - A Systematic Review and Meta-Analysis. PloS one 2016;**11**(7):e0158765.
4. Hsu RK, Powe NR. Recent trends in the prevalence of chronic kidney disease: not the same old song. Current Opinion in Nephrology and Hypertension 2017;**26**(3):187-96.
5. Go AS, Chertow GM, Fan D, et al. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. The New England journal of medicine 2004;**351**(13):1296-305.
6. Said S, Hernandez GT. The link between chronic kidney disease and cardiovascular disease. Journal of Nephropathology 2014;**3**(3):99-104.
7. Subbiah AK, Chhabra YK, Mahajan S. Cardiovascular disease in patients with chronic kidney disease: a neglected subgroup. Heart Asia 2016;**8**(2):56-61.
8. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, et al. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. The Lancet 2013;**382**(9889):339-52.
9. Wright J, Hutchison A. Cardiovascular disease in patients with chronic kidney disease. Vascular Health and Risk Management 2009;**5**:713-22.
10. Di Angelantonio E, Chowdhury R, Sarwar N, et al. Chronic kidney disease and risk of major cardiovascular disease and non-vascular mortality: prospective population based cohort study. BMJ (Clinical research ed) 2010;**341**.
11. Matsushita K, van der Velde M, Astor BC, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet (London, England) 2010;**375**(9731):2073-81.
12. Pagels AA, Soderkvist BK, Medin C, et al. Health-related quality of life in different stages of chronic kidney disease and at initiation of dialysis treatment. Health and quality of life outcomes 2012;**10** (71).
13. Gorodetskaya I, Zenios S, McCulloch CE, et al. Health-related quality of life and estimates of utility in chronic kidney disease. Kidney international 2005;**68**(6):2801-8.
14. Meyer T, Hostetter, T. Medical Progress: Uremia. The New England journal of medicine 2007;**13**(357):1316-25.
15. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. Journal of Renal Care 2013;**39**(3):140-50.
16. Rotenstein LS, Huckman RS, Wagle NW. Making Patients and Doctors Happier — The Potential of Patient-Reported Outcomes. New England Journal of Medicine 2017;**377**(14):1309-12.
17. FDA. Guidance for industry. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Silver Spring, MD: US Department of Health and Human Services Food and Drug Administration, 2009.
18. Black N. Patient reported outcome measures could help transform healthcare. BMJ (Clinical research ed) 2013;**346**:f167.
19. Aiyegbusi OL, Kyte D, Cockwell P, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PloS one 2017;**12**(6):e0179733.
20. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2012;**21**(8):1305-14.

21. Bryan S, Davis J, Broesch J, et al. Choosing your partner for the PROM: a review of evidence on patient-reported outcome measures for use in primary and community care. Healthcare Policy 2014;**10**(2):38-51.
22. Brooks R. EuroQol: the current state of play. Health Policy 1996;**37**(1):53-72.
23. Dolan P. Modeling valuations for EuroQol health states. Medical care 1997;**35**(11):1095-108.
24. Jesky MD, Dutton M, Dasgupta I, et al. Health-Related Quality of Life Impacts Mortality but Not Progression to End-Stage Renal Disease in Pre-Dialysis Chronic Kidney Disease: A Prospective Observational Study. PloS one 2016;**11**(11):e0165675.
25. Grove BE, Schougaard LM, Hjollund NH, et al. Self-rated health, quality of life and appetite as predictors of initiation of dialysis and mortality in patients with chronic kidney disease stages 4-5: a prospective cohort study. BMC Res Notes 2018;**11**(1):371.
26. Porter AC, Lash JP, Xie D, et al. Predictors and Outcomes of Health–Related Quality of Life in Adults with CKD. Clinical journal of the American Society of Nephrology : CJASN 2016;**11**(7):1154-62.
27. Porter A, Fischer MJ, Wang X, et al. Quality of Life and Outcomes in African Americans with CKD. Journal of the American Society of Nephrology : JASN 2014;**25**(8):1849-55.
28. Tsai Y-C, Hung C-C, Hwang S-J, et al. Quality of life predicts risks of end-stage renal disease and mortality in patients with chronic kidney disease. Nephrology Dialysis Transplantation 2010;**25**(5):1621-26.
29. Stringer S, Sharma P, Dutton M, et al. The natural history of, and risk factors for, progressive chronic kidney disease (CKD): the Renal Impairment in Secondary care (RIISC) study; rationale and protocol. BMC nephrology 2013;**14**:95.
30. Levey AS, Bosch JP, Lewis J, et al. A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation. Annals of internal medicine 1999;**130**(6):461-70.
31. WMA. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. The Journal of the American College of Dentists 2014;**81**(3):14-8.
32. The English Indices of Deprivation 2010: Department for Communities and Local Government, 2011.
33. EQ-5D Value Sets: Inventory, Compratice review and User Guide. Netherlands: Springer 2007.
34. Mujais SK, Story K, Brouillette J, et al. Health-related Quality of Life in CKD Patients: Correlates and Evolution over Time. Clinical journal of the American Society of Nephrology : CJASN 2009;**4**(8):1293-301.
35. Kind P, Dolan P, Gudex C, et al. Variations in Population Health Status: Results from a United Kingdom National Questionnaire Survey. BMJ: British Medical Journal 1998;**316**(7133):736-41.
36. Rocco MV, Gassman JJ, Wang SR, et al. Cross-sectional study of quality of life and symptoms in chronic renal disease patients: the Modification of Diet in Renal Disease Study. American journal of kidney diseases : the official journal of the National Kidney Foundation 1997;**29**(6):888-96.
37. Bjorner JB, Damsgaard MT, Watt T, et al. Tests of Data Quality, Scaling Assumptions, and Reliability of the Danish SF-36. Journal of clinical epidemiology;**51**(11):1001-11.
38. Leplège A, Ecosse E, Verdier A, et al. The French SF-36 Health Survey. Journal of clinical epidemiology;**51**(11):1013-23.
39. Apolone G, Mosconi P. The Italian SF-36 Health Survey. Journal of clinical epidemiology;**51**(11):1025-36.
40. Ricardo AC, Hacker E, Lora CM, et al. Validation of the Kidney Disease Quality of Life Short Form 36 (KDQOL-36TM) US Spanish and English versions in a cohort of Hispanics with chronic kidney disease. Ethnicity and Disease 2013;**23**(2):202-09.

41. Mapes DL, Bragg-Gresham JL, Bommer J, et al. Health-related quality of life in the Dialysis Outcomes and Practice Patterns Study (DOPPS). American journal of kidney diseases : the official journal of the National Kidney Foundation 2004;**44**(5 Suppl 2):54-60.
42. Landray MJ, Emberson JR, Blackwell L, et al. Prediction of ESRD and Death Among People With CKD: The Chronic Renal Impairment in Birmingham (CRIB) Prospective Cohort Study. American Journal of Kidney Diseases 2010;**56**(6):1082-94.
43. Van Manen JG, Korevaar JC, Dekker FW, et al. Adjustment for comorbidity in studies on health status in ESRD patients: which comorbidity index to use? J Am Soc Nephrol 2003;**14**(2):478-85.
44. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J Chronic Dis 1987;**40**(5):373-83.
45. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Annals of internal medicine 2009;**150**(9):604-12.
46. Go AS, Chertow GM, Fan D, et al. Chronic Kidney Disease and the Risks of Death, Cardiovascular Events, and Hospitalization. New England Journal of Medicine 2004;**351**(13):1296-305.
47. Hallan SI, Ritz E, Lydersen S, et al. Combining GFR and Albuminuria to Classify CKD Improves Prediction of ESRD. Journal of the American Society of Nephrology : JASN 2009;**20**(5):1069-77.
48. Kelley K, Aricak OT, Light RP, et al. Proteinuria is a determinant of quality of life in diabetic nephropathy: modeling lagged effects with path analysis. Am J Nephrol 2007;**27**(5):488-94.
49. Revicki DA, Brown RE, Feeny DH, et al. Health-related quality of life associated with recombinant human erythropoietin therapy for predialysis chronic renal disease patients. American journal of kidney diseases : the official journal of the National Kidney Foundation 1995;**25**(4):548-54.
50. Perlman RL, Finkelstein FO, Liu L, et al. Quality of life in chronic kidney disease (CKD): a cross-sectional analysis in the Renal Research Institute-CKD study. American journal of kidney diseases : the official journal of the National Kidney Foundation 2005;**45**(4):658-66.
51. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association 1999;**94**(446):496-509.
52. Noordzij M, Leffondré K, van Stralen KJ, et al. When do we need competing risks methods for survival analysis in nephrology? Nephrology Dialysis Transplantation 2013;**28**(11):2670-77.
53. Berg SK, Thorup CB, Borregaard B, et al. Patient-reported outcomes are independent predictors of one-year mortality and cardiac events across cardiac diagnoses: Findings from the national DenHeart survey. Eur J Prev Cardiol 2018:2047487318769766.
54. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. Journal of the National Cancer Institute 2011;**103**(24):1808-10.
55. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. Journal of the National Cancer Institute 2009;**101**(23):1624-32.
56. Abdel-Kader K, Unruh ML, Weisbord SD. Symptom Burden, Depression, and Quality of Life in Chronic and End-Stage Kidney Disease. Clinical journal of the American Society of Nephrology : CJASN 2009;**4**(6):1057-64.
57. Campbell KL, Ash S, Bauer JD. The impact of nutrition intervention on quality of life in pre-dialysis chronic kidney disease patients. Clin Nutr 2008;**27**(4):537-44.
58. Lee SJ, Son H, Shin SK. Influence of frailty on health-related quality of life in pre-dialysis patients with chronic kidney disease in Korea: a cross-sectional study. Health and quality of life outcomes 2015;**13**:70.

Chapter 5: Patient and clinician perspectives on the use of electronic patient-reported outcome measures (ePROMs)

Patient and clinician perspectives on the use of electronic patient-reported outcome measures (ePROMs): a qualitative study

The narrative review (Chapter 1) highlighted the need for clinician buy-in to ensure the successful implementation of ePROM interventions.[1-4] Eliciting pre-implementation views of patients and clinicians could inform the development of ePROM interventions and ensure that the interventions are acceptable to the end-users. This in turn may facilitate the subsequent adoption of the ePROM intervention.[5-9] It is also important that these opinions are explored with stakeholders within the intended setting as recent research shows that the use of PROMs varies widely between medical specialities and depends to a large extent on the setting.[10]

This chapter presents the qualitative study of the views of patients with CKD and renal clinicians from UHB on the use of PROMs/ePROMs for routine management of patients. Saturation tables as suggested by Kerr et al.[11] were used to demonstrate the achievement of data saturation (Appendices 8 & 9).

References

1. Rotenstein LS, Agarwal A, O'Neil K, et al. Implementing patient-reported outcome surveys as part of routine care: lessons from an academic radiation oncology department. J Am Med Inform Assoc 2017;**24**(5):964-68.
2. Antunes B, Harding R, Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: A systematic review of facilitators and barriers. Palliative Medicine 2014;**28**(2):158-75.
3. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. The Lancet;**374**(9687):369-70.
4. Aiyegbusi OL, Kyte D, Cockwell P, et al. A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures. Curr Opin Nephrol Hypertens 2017;**26**(6):442-49.
5. Szajna B. Empirical Evaluation of the Revised Technology Acceptance Model. Management Science 1996;**42**(1):85-92.
6. Holzner B, Giesinger JM, Pinggera J, et al. The Computer-based Health Evaluation Software (CHES): a software for electronic patient-reported outcome monitoring. BMC medical informatics and decision making 2012;**12**:126.

7. Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: challenges and opportunities. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2009;**18**(1):99-107.
8. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(4):419-29.
9. Zbrozek A, Hebert J, Gogates G, et al. Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2013;**16**(4):480-9.
10. Greenhalgh J, Gooding K, Gibbons E, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. Journal of Patient-Reported Outcomes 2018;**2**(1):42.

The study has been accepted for publication as:

Aiyegbusi, O.L., Kyte, D., Cockwell, P., Marshall, T., Dutton, M., Walmsley-Allen, N., Slade, A., McMullan C. and Calvert, M. Patient and clinician perspectives on electronic patient-reported outcome measures (ePROMs) in the management of advanced CKD: a qualitative study. *American Journal of Kidney Diseases.*

The study was presented at the following conferences:

- 3rd PROM Conference, Birmingham, UK, June 2018 *(Oral presentation)*
- UK Kidney Week, Harrogate, UK, June 2018 *(Poster presentation)*
- 12th Annual Research Poster Conference, Birmingham, UK, June 2018 *(Poster presentation)*
- College of Medical and Dental Sciences (MDS) Research Festival, Birmingham, UK, April 2018 *(Poster presentation)*

Patient and clinician perspectives on electronic patient-reported outcome measures (ePROMs) in the management of advanced CKD: a qualitative study

Olalekan Lee Aiyegbusi, MBChB[1, 2*], Derek Kyte, PhD[1, 2, 3], Paul Cockwell, PhD[1, 4], Tom Marshall, PhD[1, 2], Mary Dutton, MSc[4], Natalie Walmsley-Allen, RN[4], Anita Slade, PhD[1, 2], Christel McMullan, PhD[1, 2], Melanie Calvert, PhD[1, 2]

1. Centre for Patient Reported Outcomes Research, University of Birmingham, Edgbaston, Birmingham, B15 2TT
2. Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham, B15 2TT
3. NIHR Birmingham Biomedical Research Centre, University Hospitals, Birmingham NHS Foundation Trust and University of Birmingham, Birmingham, UK.
4. Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston, Birmingham, B15 2GW

*** Correspondence**: Olalekan Lee Aiyegbusi, oxa238@bham.ac.uk

Centre for Patient Reported Outcome Research, Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham, B15 2TT

Abstract

Rationale and Objective: Chronic kidney disease (CKD) can significantly affect patients' health-related quality of life (HRQOL). Electronic patient-reported outcome measures (ePROMs) may capture symptoms and HRQOL, and assist in the management of CKD. This study explored patient and clinician views on the use of a renal ePROM system.

Study Design: Qualitative study.

Setting & Participants: Twelve patients with stage 4 or 5 CKD (pre-dialysis); 22 clinicians (6 CKD community nurses, 1 clinical psychologist, 10 nephrologists, 3 specialist registrars and 2 renal surgeons) in the United Kingdom.

Analytical Approach: Semi-structured interviews and focus group discussion during which patients received paper versions of the Kidney Disease Quality of Life-36 (KDQOL-36) and the Integrated Patient Outcome Scale-Renal (IPOS-Renal) to exemplify the type of content that could be included in an ePROM. Thematic analysis of interview transcripts.

Results: Four themes were identified: (i) general opinions of PROMs; (ii) potential benefits and applications of an ePROM system; (iii) practical considerations for the implementation of ePROMs; and (iv) concerns, barriers and facilitators. Patients were willing to complete ePROMs on a regular basis as part of their care despite clinician concerns about patient burden. Patients assessed the questionnaires favourably. Clinicians suggested that the extent of adoption of renal ePROM systems in routine clinical settings should be based on evidence of significant impact on patient outcomes. Clinicians were concerned that an ePROM system may raise patient expectations to unrealistic levels and expose clinicians to the risk of litigation.

Patients and clinicians identified potential benefits, highlighted issues and concerns that need to be addressed to ensure the successful implementation of the renal ePROM system.

Limitations: Transferability of the findings may be limited as only English-speaking participants were recruited to the study.

Conclusions: A renal ePROM system may play a supportive role in the routine clinical management of patients with advanced CKD if the concerns of clinicians and patients can be sufficiently addressed.

INDEX WORDS: Chronic kidney disease; CKD; renal disease; chronic renal failure; interview; focus group; qualitative research; health-related quality of life; HRQOL; QOL; patient-reported outcome; PRO; electronic patient-reported outcome measures; ePROM.

Plain language summary

Patients with advanced chronic kidney disease often experience clusters of symptoms and impaired health-related quality of life (HRQOL) as their kidney function deteriorates. Remote electronic monitoring of patients' symptoms may facilitate communication between patients and the clinical team, and allow early interventions to manage symptoms; potentially delaying deterioration in kidney function and improving HRQOL. To ensure that our ePROM system met stakeholder needs we interviewed patients and clinicians to seek their opinions regarding system design as well as barriers and facilitators to implementation. Clinicians and patients were supportive of the proposed system suggesting that a renal ePROM system may play a supportive role in the clinical management of patients with advanced kidney disease.

Introduction

Patients with chronic kidney disease (CKD) commonly report 'clusters' of non-specific symptoms such as pain, fatigue and pruritus[1 2] , which may adversely affect their physical, emotional and psychological well-being[3] and significantly impair health-related quality of life (HRQOL).[4]

While 'hard' parameters such as estimated glomerular filtration rate (eGFR) are established routine indicators of health status,[5] these may not fully capture the impact of disease on patients' symptoms and HRQOL.[6-8] Thus, there has been a growing move towards patient-led provision of HRQL and symptom data using validated self-reported questionnaires known as patient-reported outcome measures (PROMs).[8 9] In recent years, there has been increasing interest in the use of PROMs in routine renal practice.[10 11]

With rapid information technology developments , collection of electronic PROMs (ePROMs) via digital platforms – e.g. personal computers, smartphones and tablets[8] – is increasingly common. This 'real-time' data could assist with timely tailoring of treatment to individual patient needs and may reduce unnecessary clinical appointments in stable patients.[8 12] In addition, the use of ePROMs may promote patient-centred care by aiding identification of health-related issues of priority to patients, facilitating patient-clinician communication and shared-decision making.[8 13-15]

The aim of this qualitative study was to explore the perspectives of patients and clinicians on the use of a renal ePROM being developed by the Centre for Patient Reported Outcomes Research, University of Birmingham and University Hospitals Birmingham NHS Foundation Trust (UHB), in the UK. The objective was to gain

valuable insights to inform the design, implementation and delivery of such a system within routine clinical practice.

Methods

Participant selection

Participants at the host site (UHB) were recruited and data collected and analysed simultaneously between August 2017 and May 2018 according to the published protocol (ethical approval [13/02/2017] ref: 17/WM/0010).[16] Consenting, English-speaking, adult patients with stage 4 and 5 CKD (pre-dialysis) were purposively recruited as we hypothesised that a cohort, with high symptom burden and risk of rapid progression to end-stage renal disease (ESRD), would benefit most from the ePROM system.[16] A third of patients were recruited from minority ethnic groups to reflect the diversity of the patient catchment area, as highlighted in the recent Renal Impairment In Secondary Care (RIISC) cohort study where 32% of the patients were from minority ethnic backgrounds.[17] UHB Renal clinicians were recruited through a combination of purposive and snowball sampling:[18] where clinicians known to the research team were contacted directly and those who agreed to be interviewed were asked to suggest other potentially eligible colleagues.

Data collection

All participants were provided information sheets which explained the study aim/objectives and also outlined the UHB ePROM system under development. They had the opportunity to ask questions before signing the consent forms prior to any interview/focus group.

All interviewees were provided with advance copies of the Kidney Disease Quality of Life-36 (KDQOL-36) and Integrated Patient Outcome Scale-Renal (IPOS-Renal) questionnaires to aid discussion. These questionnaires were selected, following review of the literature[19 20] and discussions with the project patient advisory group, to provide participants with an idea of the type of content that could be included in the ePROM system. The KDQOL-36 has strong evidence for internal consistency and moderate evidence for construct validity[19] while the IPOS-Renal was reported to have good evidence for test-retest reliability, internal consistency, and construct validity.[20] Interviews/focus groups were conducted in accordance with relevant topic guides, developed based on findings from existing literature and discussions within the research team (Boxes 1 & 2). OLA conducted audio-recorded semi-structured face-to-face or telephone interviews with patients according to their preference. A focus group and semi-structured interviews were held with clinicians. Focus groups were not conducted with patients as it was felt some might have found it distressing to discuss their condition within a group setting.[21]

Data analysis

Thematic data analysis was conducted by OLA[22] alongside data collection, providing an opportunity to refine the topic guide in response to findings.[23] Initial coding was reviewed by DK and MC prior to further analysis supported by the Nvivo 11 Plus software package by QSR International. As data analysis progressed, OLA grouped the codes into categories, subthemes and themes. These were regularly discussed with members of the research team and modified where deemed appropriate. Data collection and analysis continued until data saturation was achieved for both participant groups.[24 25] At this point, no new concept-relevant information was elicited

from analysing additional interview data.[26 27] Saturation tables, which documented the initial and subsequent occurrences of codes, were used to assess data saturation.[26]

Results

Study participants

Participant characteristics are summarized in Tables 1-2 and interview characteristics/saturation data in Table 3. Focus group members were predominantly female nurses. For this reason, we invited medical doctors for interviews and targeted male doctors in order to minimise the potential risk of non-response bias and increase the diversity of our sample. Examination of the saturation data (Supplementary tables S1 & S2) throughout the study suggested that (i) saturation was reached at the 10th patient and 12th clinician interviews (ii) there were no appreciable differences in the views held by nurses and doctors, and (iii) no gender differences in the views held by participants.

Themes

Four themes were identified in the interviews/focus group: (i) general opinions of PROMs; (ii) potential benefits and applications of ePROMs; (iii) Practical considerations; and (iv) concerns, barriers and facilitators. Themes and sub-themes are presented below, illustrative quotations are provided in Table 4 and supplementary Table S3. Figure 1 illustrates the conceptual relationships between the themes. Specific comments KDQOL-36 and the IPOS-Renal questionnaires are provided in supplementary file S4.

General opinions of PROMs

The role of PROMs in healthcare (clinician-generated)

Most clinicians welcomed the idea of incorporating symptoms/HRQOL assessment into clinical practice as they believed that clinical parameters may not fully capture the impact of CKD on patients nor reflect outcomes regarded as important by patients (q1). Some clinicians felt that the use of ePROMs would play an important role in the continued digitisation of healthcare (q2).

Knowledge and experience of PROMs (clinician-generated)

Knowledge and experience of PROMs varied among clinicians. Only four doctors stated that they had used PROMs for research and only one stated that they had used a PROM for routine clinical practice. Majority of clinicians expressed a keen interest PROMs research and the findings of this present study (q3).

Cautious optimism (clinician-generated)

Although majority of clinicians acknowledged the importance of assessing HRQOL, some believed that PROMs were of limited benefit in routine clinical practice (q4). They perceived PROMs as a useful adjunct to traditional clinical management but insufficient on their own for obtaining research funding or changing health policy (q5).

Criticism of PROMs (clinician-generated)

One doctor argued that PROMs do not capture HRQOL as they are: (i) incapable of measuring the gap between patient health experience and aspirations; and (ii) not truly ‘patient-reported’ as the questions are not generated by the individual patient and so may not be important to them (q6). The individual went on to speculate that

PROMs gained prominence due to time pressures and reluctance or inability of clinicians to communicate effectively with patients (q7).

Comparison with history taking (clinician-generated)

Some doctors compared the administration of PROMs with traditional history taking, maintaining that the latter remains the model for patient-clinician interactions. They described PROMs as a structured way of asking the same questions they would ask during a clinical consultation (q8). The key difference was that "you *[the doctor]* ask them a series of questions that *you* select". In other words, the questions asked during a clinical consultation are questions the doctor deems important.

Potential benefits and applications of ePROMs

Potential benefits (patient and clinician-generated)

Patients and clinicians asserted that employing an ePROM system could improve or open lines of communication between patients and clinicians as well as providing clinicians with an insight into patient experiences and care priorities (q9). They believed that ePROMs could generate early warnings of deterioration in patients' condition and detect psychological distress in patients (q10, 11).

Patients with stable CKD suggested that the use of an ePROM system could potentially reduce the frequency of their hospital appointments and in turn the need to take time off work (q12). Clinicians agreed with patients on this point with the proviso that such patients were made aware of the symptoms of deterioration and had mandatory laboratory tests done (q13).

Some patients felt this would facilitate self-monitoring and management; aspects of care they believed were currently neglected (q14). Patients and clinicians believed

that the regular completion of an ePROM would raise the level of CKD awareness and self-reflection among patients (q14, 15).

Some clinicians proposed that ePROM use could help focus clinical consultations on issues prioritised by patients and reduce time pressures (q16).

Determinants of benefits (patient and clinician-generated)

Clinicians concluded that the utility of an ePROM system would depend on the manner and purpose of its use, the consultation style of individual doctors, and the personalities of patients (q17). They felt that ePROMs may facilitate communication with patients and promote rapport in situations where the clinician was unfamiliar with the patient (q18).

Potential Applications (clinician-generated)

Clinicians suggested ePROM systems could assist with the remote monitoring of patient symptoms and response to treatments, and provide additional information to review during the regular Multidisciplinary Team (MDT) meetings (q19, 20). They believed ePROMs could play a vital role in the conservative/palliative management of patients where symptoms/HRQOL is the focus rather than clinical outcomes (q21).

Practical considerations

Administration of ePROMs (patient and clinician-generated)

While there was no consensus on the optimal frequency of administration, patients and clinicians generally felt that it would become burdensome if an ePROM was administered more frequently than on a monthly basis. Some participants reasoned that the frequency of administration would depend on disease trajectory in individual

patients with stable patients requiring less frequent administration. However, they pointed out that deteriorating patients, who may require closer monitoring, might be unwilling to complete ePROMs due to the burden of illness (q22, 23). There were suggestions that patients could complete an ad-hoc ePROM if they felt unwell rather than wait for the next fixed time interval (q24). Some participants felt the frequency could coincide with patients' follow-up so that clinicians have the results to review before hospital appointments. Patients and clinicians agreed that the best time to complete an ePROM was at home prior to a clinical appointment away from the pressures of busy clinics and interference from medical staff (q25). The general opinion of patients and clinicians was that everyone involved in a patient's care, including the patient, should have access to ePROM data (q26).

Participants pointed out that some patients might not be able or willing to complete ePROMs due to age, computer literacy, access to the internet and electronic devices, and language difficulties. While most patients expressed a strong preference for ePROMs, a few preferred paper questionnaires but were willing to complete ePROMs (q27). Some participants speculated that family members who assist patients with the completion of ePROMs may influence the results obtained (q28).

Feedback (patient and clinician-generated)

Patients felt that receiving feedback on their ePROM reports would help them understand and monitor their condition better, while clinicians would like feedback on the effectiveness of their response to ePROMs data (q29, 30).

Interpretation of ePROM data (clinician-generated)

Many clinicians discussed the interpretation of ePROMs data giving examples of how they would interpret responses to certain questions (q31). A clinician mentioned the difficulty with interpreting and understanding the clinical significance of summary scores and score changes, particularly if clinicians were unfamiliar with PROMs (q32).

Presentation of ePROM data (clinician-generated)

Clinicians believed that ePROMs data should be provided within the electronic health care record to encourage clinician uptake. They suggested pictorial representation specifically; graphical and colour-coded ‘traffic light’ formats which were seen as quick and easy to review. A ‘traffic light’ format could have the additional benefit of providing triggers or alerts if there was a significant change in a patient’s condition (q33-35).

Concerns, potential barriers and facilitators

Concerns (patient and clinician-generated)

One patient was convinced that patients generally ‘want to keep the doctor happy’ and so may provide very positive responses which they believe will please their doctor and protect their existing relationship (q36). This was echoed by some doctors who stressed the importance of patients completing their ePROMs in the absence of their doctor to ensure that they provide ‘truthful’ responses.

Although patients identified data security and patient confidentiality as potential issues, they did not appear unduly concerned by them (q37).

Clinicians worried about the potential burden on very ill patients who may struggle to complete ePROMs (q38). One clinician argued that improving patient’s HRQOL (e.g.

reduced frequency of dialysis) might worsen patients' clinical parameters which would have a negative effect on the monthly assessment of clinicians' performance (q39). There were concerns that ePROMs might generate more work for clinicians and lead to an increase in psychological/psychiatric referrals (q40, 41). Some clinicians worried about charges of clinical negligence and the risk of litigation if ePROM results were not acted upon and patients experience adverse outcomes (q42).

Some clinicians suggested that an increased focus on patients' emotions and psychological wellbeing through the use of ePROMs could lead to the medicalization of what they believed to be normal reactions to illness (q43). Some feared that using ePROMs may unwittingly raise patients' expectations beyond what they considered to be achievable (q44).

Potential barriers (patient and clinician-generated)

A lack of interest in the ePROM system, a dislike of information technology, limited capabilities or access to the internet/electronic devices were identified by patients as potential barriers (q45).

Some patients speculated that individuals with relatively stable disease would be less inclined to complete an ePROM on a regular basis (q46). This contradicted the suggestion by patients with lived experience of 'stable' CKD that they would rather complete an ePROM report on a regular basis than take time off work to attend clinic appointments (q12).

Clinicians felt that the current 'alert fatigue' as a result of numerous alerts being provided in Clinical Portal may cause clinicians to ignore ePROM results and alerts (q47). Some believed that clinician discomfort broaching issues outside of physical symptoms may be a potential barrier (q48). The perception was that there are tight

limits to what is achievable; limited financial resources and time pressures during clinical consultations were considered potential barriers (q49, 50).

Potential facilitators (patient and clinician-generated)

Patients felt that the provision of explanations of the significance of questions and results would encourage them to complete ePROMs on a regular basis (q51). Clinicians believed that evidence of impact on patient outcomes would be the strongest facilitator for the adoption of ePROMs. It was acknowledged that such hard evidence would take time to generate but anecdotal or preliminary evidence would suffice in the meantime (q52).

The provision of clear guidelines on how to interpret ePROM results and address issues raised was also suggested as a facilitator (q53, 54). Nurses asserted that buy-in by doctors and their involvement in the design and implementation process was crucial (q55). Easy access to ePROMs results and providing patients with paper option were also cited as potential facilitators (q56, 57).

Discussion

This study explores the views of patients and clinicians on the use of an ePROM system in the clinical management of patients with advanced CKD. Consistent with previous literature, patients welcomed the idea of completing ePROMs on a regular basis as part of their care.[28] They believed ePROMs could help clinicians manage their care more efficiently and effectively. A notable finding was the interest expressed by some patients in using ePROM data to improve their knowledge of CKD to facilitate self-monitoring and management.

Although, as reported by other studies, most patients expressed a preference for ePROMs over paper questionnaires,[29 30] patients and clinicians stressed the need to provide a paper option for patients unable/unwilling to complete ePROMs. Studies have shown that response rates can be optimised when electronic and paper formats are provided and reminders are utilised.[12 31 32]

Whilst all the clinicians interviewed had research experience, their overall clinician knowledge and experience of PROMs in clinical research or routine practice was limited. Whilst most clinicians agreed on the potential benefits of ePROMs, they were cautious about their real life performance in practice. The consensus was that renal ePROM systems could become a vital adjunct in the clinical management of patients, providing there was compelling evidence of their usefulness. This demonstration of value has been recommended as an important facilitator for the adoption of ePROM systems.[33] Clinicians in the study believed there was little evidence to support the use of PROMs in routine renal practice. However, this is not the case, as recent research suggests that ePROMs may detect clinically relevant changes in patients receiving dialysis.[8 34] In Denmark the use of an ePROM system (Ambuflex), led to a 50% decrease in follow-up hospital visits.[31] In other areas, a

recent ePROM trial demonstrated improved survival, HRQOL, and reduction in hospitalization in patients with cancer undergoing chemotherapy.[6]

Some clinicians were concerned that ePROMs could potentially raise patient expectations beyond what is achievable in daily practice. However, most of the patients interviewed understood, through first-hand experience, the chronic and progressive nature of their condition, and the challenges facing the health system.

Some clinicians were concerned that evaluating psychological domains in patients might lead to a 'medicalization' of what *they* consider as normal responses to illness. The assumption that clinicians are able to distinguish between 'normal' and 'abnormal' responses has been questioned;[35] patients often see different doctors at each clinic appointment and so may be unable to forge relationships that could enable the detection of deviations from 'normal' responses.[35] Some clinicians felt that the evaluation of these domains would be helpful as they believed that renal doctors were either unskilled at detecting and addressing psychological problems or reluctant to explore these issues due to factors such as time constraints and personal discomfort. While medicalization of normal emotional responses may be an unintended consequence of evaluating patients' emotional and psychological status, this may be unlikely if appropriate thresholds are set for the ePROM system and suitable training provided.[36 37] Moreover, it has been shown that the emotional self-awareness gained from self-monitoring may reduce depression in patients.[38]

Clinicians pondered the issue of providing appropriate response to ePROMs data particularly results relating to psychological domains. While making referrals to appropriate specialists appeared as the obvious solution, some felt this could lead to excessive referrals being made. This is a valid concern as a substantial increase in referrals would place greater burden on limited resources. However, this may be

unlikely if appropriate thresholds are set. Others worried about the potential risk of litigation if ePROMs data is not acted on and patients suffer unforeseen adverse outcomes such as suicide. The worry about litigation and charges of medical negligence is an important consideration which may influence clinician adoption of the ePROM system. The risk of litigation could be minimised if prior to full system implementation appropriate care protocols are developed for clinicians and patients are provided information for emergency situations. Unsurprisingly, clinicians requested clear guidance on meaning of and the appropriate response to ePROMs data as recommended in literature.[39 40]

In order to be fit for purpose the ePROM system needs to be efficient, effective and satisfactory for stakeholders.[41 42] If appropriately addressed, these study findings could assist with the design and implementation of the renal ePROM system, improve its usability, facilitate stakeholder adoption.[43] Although the study focuses on ePROMs, majority of the key findings are applicable to the use of PROMs in routine clinical practice regardless of format. While we have selected the KDQOL-36 and IPOS-Renal for this project, there are other renal-specific PROMs in use such as the CHOICE Health Experience Questionnaire (CHEQ) and the Edmonton Symptom Assessment System (ESAS).[19]

Patients with stages 4 and 5 (pre-dialysis) CKD were interviewed as they represented our target population for the intervention. Further research is necessary to see if similar views are held by patients with stages 1-3 CKD, or patients receiving renal replacement therapy for ESRD. In addition, this study was conducted during the early stages of system development therefore, participant opinions may differ post-implementation and may not reflect those held in settings where ePROMs are already implemented. While the focus group was all-female and the majority were

nurses, their opinions on the key issues were confirmed by the mostly male doctors interviewed. We therefore believe that the composition of the focus group does not affect the transferability of our findings in terms of gender and clinician profession.

At present, there is limited information on the use of ePROM systems in the routine renal practice and the majority of the articles we came across described systems at development or pilot stages.[28] The RePROM system being developed by researchers at the Centre for Patient Reported Outcome Research (CPROR), UHB, and Birmingham Clinical Trials Unit (BCTU) has a focus on patient symptoms. Recruitment has commenced for a randomised pilot trial of usual care supplemented with the RePROM system in patients with advanced CKD.[44]

The use of a renal ePROM system has the potential to enhance routine clinical practice by facilitating patient engagement and involvement in their care and providing clinicians with timely information which may guide clinical management. The rapid developments in information technology may also assist with the integration of ePROM data with other routinely collected electronic health data, thus facilitating its impact.[45] However, patients and clinicians need to be involved at every stage in the development of ePROM systems. Patient and clinician views should be sought, considered and appropriately utilised in other to facilitate their subsequent engagement with ePROM interventions. The degree of patient and clinician engagement may crucially influence the usefulness of ePROMs post-implementation.

Article Information

Contributions

This study was conceived and designed by OLA, MC, DK, PC, TM, MD and NWA. AS moderated the focus group. OLA conducted the interviews and analysed the transcripts. OLA, MC, DK, PC, TM, CM reviewed the data analysis. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

Funding

This project is funded as part of the Health Foundation's Improvement Science PhD Awards Programme. The Health Foundation is an independent charity working to improve the quality of healthcare in the UK. The Health Foundation was not involved in any other aspect of the project.

Tom Marshall is partly funded by the National Institute for Health Research (NIHR) through the Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC-WM). Derek Kyte is funded by the NIHR post-doctoral research fellowship programme (Ref: PDF-2016-09-009). Melanie Calvert and Anita Slade are funded by the NIHR Birmingham Biomedical Research Centre and the NIHR Surgical Reconstruction and Microbiology Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. This paper presents independent research and the views expressed in this publication are those of the author(s) and not necessarily those of the NHS, NIHR, the Department of Health, University of Birmingham or the CLAHRC WM Management Group.

The research team acknowledges the support of the National Institute for Health Research Clinical Research Network (NIHR CRN).

Competing interests: MC has received consultancy fees outside this work from Ferring and Takeda. The remaining authors declare that they have no relevant financial interests.

Peer review: Received August 22, 2018. Evaluated by 2 external peer reviewers and a methods reviewer, with direct editorial input from an Associate Editor, and the Editor-in-Chief. Accepted in revised form February 6, 2019.

References

1. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. Journal of Renal Care 2013;**39**(3):140-50.
2. Jablonski A. The multidimensional characteristics of symptoms reported by patients on hemodialysis. Nephrol Nurs J 2007;**34**(1):29-37.
3. Gapstur RL. Symptom burden: a concept analysis and implications for oncology nurses. Oncology nursing forum 2007;**34**(3):673-80.
4. Fairclough DL. *Design and analysis of quality of life studies in clinical trials*. New York: Chapman & Hall/CRC Press, 2002.
5. Tangri N, Stevens LA, Griffith J, et al. A predictive model for progression of chronic kidney disease to kidney failure. Jama 2011;**305**(15):1553-9.
6. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. Journal of Clinical Oncology 2016;**34**(6):557-65.
7. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. Journal of the National Cancer Institute 2011;**103**(24):1808-10.
8. Aiyegbusi OL, Kyte D, Cockwell P, et al. A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures. Curr Opin Nephrol Hypertens 2017;**26**(6):442-49.
9. FDA. Guidance for industry. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Silver Spring, MD: US Department of Health and Human Services Food and Drug Administration, 2009.
10. Finkelstein F. #NephMadness 2018: Patient-Centered Care Medicine – The Future. AJKD Blog, 2018.
11. Dwyer J. #NephMadness 2018: Trial Outcomes Region. AJKD Blog, 2018.
12. Hjollund NH, Larsen LP, Biering K, et al. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. Interactive journal of medical research 2014;**3**(1):e5.
13. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2004;**22**(4):714-24.
14. Gilbody SM, Whitty PM, Grimshaw JM, et al. Improving the detection and management of depression in primary care. Quality & safety in health care 2003;**12**(2):149-55.
15. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA: a cancer journal for clinicians 2012;**62**(5):337-47.
16. Aiyegbusi OL, Kyte D, Cockwell P, et al. Using Patient-Reported Outcome Measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney disease (PRO-trACK project): a mixed-methods project protocol. BMJ Open 2017;**7**(6):e016687.
17. Jesky MD, Dutton M, Dasgupta I, et al. Health-Related Quality of Life Impacts Mortality but Not Progression to End-Stage Renal Disease in Pre-Dialysis Chronic Kidney Disease: A Prospective Observational Study. PloS one 2016;**11**(11):e0165675.
18. Patton MQ. *Qualitative Research and Evaluation Methods*. 3rd ed. Thousand Oaks: Sage, 2002.
19. Aiyegbusi OL, Kyte D, Cockwell P, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PloS one 2017;**12**(6):e0179733.
20. Raj R, Ahuja K, Frandsen M, et al. Validation of the IPOS-Renal Symptom Survey in advanced kidney disease: a cross-sectional study. Journal of Pain and Symptom Management;**56**(2):281-87.

21. Tonkiss F. Using focus groups. In: Seale C, ed. Researching Society and Culture. 2 ed. London: Sage Publications, 2004.
22. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology 2006;**3**(2):77-101.
23. Pope C, Ziebland S, Mays N. Analysing qualitative data. BMJ : British Medical Journal 2000;**320**(7227):114-16.
24. Fusch P, Ness L. Are We There Yet? Data Saturation in Qualitative Research. The Qualitative Report 2015;**20**(9):1408-16.
25. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. Field Methods 2006;**18**(1):59-82.
26. Kerr C, Nixon A, Wild D. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. Expert rev 2010;**10**(3):269-81.
27. Rothman M, Burke L, Erickson P, et al. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(8):1075-83.
28. Schick-Makaroff K, Molzahn A. Brief communication: patient satisfaction with the use of tablet computers: a pilot study in two outpatient home dialysis clinics. Canadian Journal of Kidney Health and Disease 2014;**1**:22.
29. Crawley JA, Kleinman L, Dominitz J. User Preferences for Computer Administration of Quality of Life Instruments. Drug Information Journal 2000;**34**(1):137-44.
30. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(4):419-29.
31. Schougaard LM, Larsen LP, Jessen A, et al. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2016;**25**(3):525-34.
32. Kongsved SM, Basnov M, Holm-Christensen K, et al. Response rate and completeness of questionnaires: a randomized study of Internet versus paper-and-pencil versions. Journal of medical Internet research 2007;**9**(3):e25.
33. Schwartzberg L. Electronic Patient-Reported Outcomes: The Time Is Ripe for Integration Into Patient Care and Clinical Research. American Society of Clinical Oncology educational book American Society of Clinical Oncology Meeting 2016;**35**:e89-96.
34. Pittman ZCL, John SG, McIntyre CW. Collection of daily patient reported outcomes is feasible and demonstrates differential patient experience in chronic kidney disease. Hemodialysis International 2017;**21**(2):265-73.
35. Baik S-Y, Bowers BJ, Oakley LD, et al. The recognition of depression: the primary care clinician's perspective. Annals of family medicine 2005;**3**(1):31-37.
36. Thangadurai P, Jacob KS. Medicalizing Distress, Ignoring Public Health Strategies. Indian Journal of Psychological Medicine 2014;**36**(4):351-54.
37. Jacob JD, Gagnon M, McCabe J. From distress to illness: a critical analysis of medicalization and its effects in clinical practice. Journal of Psychiatric and Mental Health Nursing 2013;**21**(3):257-63.
38. Kauer SD, Reid SC, Crooke AH, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. Journal of medical Internet research 2012;**14**(3):e67.
39. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Quality of

life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2012;**21**(8):1305-14.
40. Aaronson N, Elliott T, Greenhalgh J, et al. User's Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice. International Society for Quality of Life Research 2015.
41. ISO. ISO 9241-11:2018(en). Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts, 2018.
42. Aiyegbusi OL, Kyte D, Cockwell P, et al. Development and usability testing of an electronic patient-reported outcome measure (ePROM) system for patients with advanced chronic kidney disease. Computers in Biology and Medicine 2018;**101**:120-27.
43. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. Journal of Comparative Effectiveness Research 2016;**5**(5):507-19.
44. Kyte D, Bishop J, Brettell E, et al. Use of an electronic patient-reported outcome measure in the management of patients with advanced chronic kidney disease: the RePROM pilot trial protocol. BMJ Open 2018;**8**(10):e026080.
45. Scalise K, Allen DD. Use of open-source software for adaptive measurement: Concerto as an R-based computer adaptive development and delivery platform. British Journal of Mathematical and Statistical Psychology 2015;**68**(3):478-96.

Box 1. Topic guide for patient interviews

- Was the purpose for the study clear from reading the patient information sheets?
- Do you think the study will benefit patients? If so why and how?
- Views on the content of the questionnaire(s)
    1. What do you think of the lengths?
    2. Are the questions easy to understand? Are the questions (and their response options) worded in a way that would be easily understood by patients? How easy do you think it is to complete the questionnaires?
    3. Do individual questionnaires cover all aspects that might be important to patients with advanced CKD, or are any elements missing?
    4. Are there questions that might be seen as unacceptable? If so why?
- How often would you be prepared to complete the questionnaire(s) as part of your care? Do you think that patients with advanced CKD would be willing to complete the questionnaire on a regular basis in-between their clinic appointments? If so, how frequently do you think they would be willing to do so?
- How would you prefer to complete the questionnaires - by PC, smartphone, tablet, telephone voice recognition or paper completion?
- What might discourage you from completing the questionnaires?
- What might encourage you to complete the questionnaires?
- Do you think the benefits you mentioned earlier are sufficient to motivate other patients to fill the questionnaire on a regular basis?
- If not mentioned, ask if feedback is desired. What kind of feedback would you prefer?
- Are there any issues about completing these questionnaires that worry you?
- Any final comments?

Box 2. Topic guide for clinician focus group and interviews

- What are your thoughts on the use of ePROMs in the management of CKD patients?
- Do you think the use of ePROMs will benefit patients? If so why and how? If not why?
- Do you think that patients with advanced CKD would be willing to complete the questionnaire on a regular basis in-between their clinic appointments?
- How often would you like patients to complete questionnaires like these?
- How would you like the data to be displayed?
- Are there any issues about using these questionnaires that you would like to mention?
- What factors may improve the use of ePROM data by clinicians?
- What factors may discourage the use of ePROM data by clinicians?
- Is there anything else anyone would like to say on the use of PROMs for the management of patients with CKD?
- Views on KDQOL-36 and IPOS-Renal
    5. What are your thoughts on the lengths of questionnaires?
    6. What are your opinions of the wording of the questions and their response options?
    7. What about the scope of the questionnaires? Do individual questionnaires cover all aspects that might be important to patients with advanced CKD, or are any elements missing?
    8. Are there questions that might be seen as unacceptable to patients? If so which questions and why?

Table 1. Patient characteristics (n = 12)	
Variable	n
Age (years)	
below 50	1
50 and over	11
Gender	
Female	5
Ethnicity	
British-White	7
British-Asian	4
Irish-White	1
Occupation	
Retired	7
Employed (Part-time & full time)	4
Unemployed	1

n – Sample size

Table 2. Clinician characteristics	
Focus group (n = 8)	
Variable	n
Gender	
Female	8
Ethnicity	
White	6
Asian	1
Black	1
Occupation	
CKD community nurses	6
Clinical psychologist	1
Nephrologist	1
Interviews (n = 14)	
Variable	n
Gender	
Female	6
Ethnicity	
White	11
Asian	3
Position	
Specialist registrars	3
Renal surgeons	2
Nephrologists	9

n – Sample size

CKD – chronic kidney disease

Table 3. Description of interviews and saturation tables	
Interviews	
Number of patient participants (number invited) Reasons for non-participation	12 (15) • Involvement in other research studies • Lack of time due to numerous clinical appointments • Lack of interest in research
Duration of patient interviews (minutes) Range Mean ± S.D	 22 - 64 37.4 ± 15.2
Number of focus group participants (number invited) Reason for non-participation	8 (14) Competing clinical commitments
Duration of focus group (minutes)	45
Number of interviews with doctors (number invited) Reason for non-participation	14 (17) Competing clinical commitments
Duration of interviews with doctors (minutes) Range Mean ± S.D	 21 - 59 32.8 ± 10.8
Saturation tables	
Patient saturation table Number of codes generated Point of data saturation	 77 10th interview
Clinician saturation table Total number of codes generated Number of codes identified in focus group Percentage of codes from focus group that recurred in subsequent interviews with doctors (%) Point of data saturation	 141 56 86* 12th interview

S.D - Standard deviation

* - non-recurring codes from the focus group pertained to discussions about the importance of involving doctors in the ePROMs project

Table 4. Illustrative Quotations by Theme and sub-themes	
Sub-themes	**Illustrative Quotations**
	General opinion of PROMs
The role of PROMs in healthcare	“It would be great if we could start incorporating more quality of life aspect..it is an aspect partly neglected at the moment.” (Doctor 8, q1) “It's the natural direction of healthcare.” (Doctor 3, q2)
Knowledge of PROMs	“I’m definitely looking forward to seeing the outcomes of the research.” (Doctor 8, q3)
Cautious optimism	“So things like level of kidney function, etc, and age, they feel at the moment more like hard risk factors. Whereas quality of life at the moment doesn’t feel like a hard risk factor that I can really latch on to and use that to stratify a patient’s risk.” (Doctor 8, q4) “I’m not sure on its own we would be able to use to change policy or to get funding or permission for various interventions but it’s an extremely useful adjunct.” (Doctor 11, q5)
Criticism of PROMs	“I agree with the concept that quality of life is a reflection of the gap between where you are and where you want to be. Those questionnaires don’t ask that…It’s just not individualised….they’re not coming up with it themselves. In that sense, it’s not patient reported.” (Doctor 5, q6) “You sit the patient and you say, ‘How are you?’ and you remain silent and the patient will tell you everything you need in two minutes. I wonder whether the whole business of PROMs has arisen through time pressure and a reluctance to do exactly that...” (Doctor 5, q7)
Comparison with history taking	“When I was trained in medicine that's not the way that you have clinical consultations. Basically history and examinations and that's still maintained as a model. I suppose you could say it is a bit like a questionnaire because it’s a formulaic you have in your mind.” (Doctor 3, q8)
	Potential benefits and applications of ePROMs
Potential benefits	“It raises issues the clinician, or the patient has never felt that they can broach before and that is the heart of clinical care...” (Doctor 6, q9) “I’m hoping that if the PROMs comes in it would raise alerts and that it would flag up if anything was wrong....” (Patient 8, q10)

	"..,the advantage is you might more systematically detect or detect sooner physical or psychological symptoms." (Doctor 3 q11) "…if I can do these sort of interviews at home I don't have to take time off work. It makes things easier for everybody." (Patient 2, q12) "I think to cut down clinic and distance for patients you know." (Doctor 4, q13) "I'm looking after my own health by doing it...it's nice to know how you're getting on....or things I should be changing? …" (Patient 8, q14) "I think it will make them more reflective." (Doctor 5, q15) "It helps them decide what they like to discuss…having the patient narrow it down, might help you in clinic as well then." (FG, q16)
Determinants of benefits	"I think the utility depends on what it's used for… people operate, perform differently, have a different style in clinic" (Doctor 3, q17) "If you're seeing someone regularly…I'm not sure that there is as much utility because you develop that feel for the patient." (Doctor 9, q18)
Potential applications	"Using more global assessments of patient wellbeing is often a very useful marker of whether a treatment is working….or at least to capture other benefits of our interventions that can't be succinctly summarise in a blood test." (Doctor 11, q19) "I think these questionnaires might be part of the wider multidisciplinary management of patients with complex diseases." (Doctor 9, q20) "Another area where it would be especially useful would be in conservative management of chronic kidney disease or palliative care in which the focus really is on managing the symptoms of their chronic kidney disease." (Doctor 8, q21)
	Practical considerations
Administration of ePROMs	"...it depends how I feel. I suppose if you're at stage four or five, you wouldn't really want something that's monthly." (Patient 11, q22) "If you've got advanced CKD, you probably don't want to be filling in a questionnaire every 6 weeks. If they're reasonably stable, they should be okay filling out a questionnaire every 3 to 6 months. I think it depends on the patient...on the nature of the disease." (Doctor 9, q23) "I suppose the other way to look at this would be to say, 'Okay, if you feel less well, maybe you should complete this questionnaire'…. rather than being very didactic and say every two months, every six months or every four months." (Doctor 9, q24) "I think filling it in before they come (but bringing it with them) so they're not rushed." (Doctor 7, q25) "I think generally the medical people that are looking after me so my GP, my consultant and anyone else..." (Patient 10, q26)

	“Probably my tablet or my phone ‘cause it’s quick and handy to just pick up, isn’t it? Yeah, rather than a paper or a laptop.” (Patient 8, q27) “It might be interpreted by someone else which would, create a bit of a struggle transferring what the patients actually think.” (Doctor 5, q28)
Feedback	“I want an explanation so that I understand what’s going on. I don’t see the point in doing it and me telling you how I’m feeling unless I get some sort of feedback from it as to why that’s happening to me.” (Patient 4, q29) “...some sort of absolute feedback for the clinician so they know where they are is probably going to be useful.” (Doctor 9, q30)
Interpretation of ePROMs	“… If they’re struggling to climb several flights of stairs, that might mean fluid overloaded…” (FG, q31) “I think with a lot of these questionnaires, when the numbers change, you don’t quite know what that means in real life.” (Doctor 6, q32)
Presentation of ePROMs	“…. I think having a separate tab where you can view it probably useful.” (Doctor 9, q33) “…a graph, could be a bar chart, the score against time or something, something visual that's easy, quick to see.” (Doctor 3, q34) “...we need trigger points don’t we? ...nstead of sifting through quite a lot of information, a traffic light system or something.” (FG, q35)
	Concerns, potential barriers and facilitators
Concerns	“They want to keep the doctor happy.....let’s be true it’s 1 but to keep the doctor happy I’m going to say 4.5….” (Patient 5, q36) “...so much of our details are online these days... *[shrugs shoulders]* I don’t think that really bothers me *[laughter]*...” (Patient 4, q37) “They do but they feel obliged some of them…They do feel obliged…” (FG, q38) “If I want to make my numbers better and make my patients live longer I dialyse them more. If you look at the one thing that impairs their quality of life in the questionnaire it’s having more dialysis.” (Doctor 2, q39) “…I mean I think we are at risk of just generating more work for ourselves.” (Doctor 10, q40) “The potential is that we could then be making lots psychology referrals or getting lots and lots of outpatient psychiatry referrals.” (FG, q41) “..once you start to collect this information…if somebody committed suicide and have ticked certain boxes, if you’re standing up in a Coroner’s Court, how do you justify that you haven’t done anything about it, it starts to raise ethical issues…(FG, q42) “…I would like to see this used to focus on very practical issues that can be dealt with in a clinic environment but I wouldn’t want to

	see this phenomenon of medicalising what is really a spectrum of sort of normal response to stress….” (Doctor 10, q43) “There is a danger of raising expectations.....” (Doctor 2, q44)
Potential barriers	“The biggest problem is people who don’t like computers, don’t have one or don’t understand them but have one.” (Patient 1, q45) “...as long as you’re feeling well, you might not be inclined to do it....” (Patient 4, q46) “There are lots of alerts already in PICS and it’s well known that people get alert fatigue and just click ignore, ignore …” (FG, q47) “Clinicians feel comfortable asking about physical health and probably less comfortable asking about mental health.” (Doctor 3, q48) “The resources available to any of us to try and address and support people psychologically are quite limited.” (Doctor 3, q49) “The limitation is going to be time for consultation.” (Doctor 3, q50)
Potential facilitators	“I think what would help is if there was an explanation as to why you’re filling that….in fairly basic language.” (Patient 4, q51) “I want there to be some evidence at the very least, even anecdotal evidence to begin with until full evidence, you can’t just magic up evidence like that, it’s a process and it needs some anecdotal evidence that filling in this questionnaire has helped Patient X, Y, Z.” (Doctor 2, q52) “I think there needs to be some clear guidelines or clear advice about what constitutes an abnormal result….” (Doctor 9, q53) “Simplicity and then expert suggestions of what I can be doing might be useful.” (Doctor 2, q54) “I think it will be really exciting to be honest, if there was enough buy into it. I think you probably need to look at the Consultants as well include them, in this sort of discussion…...” (FG, q55) “I think it needs to be easy to access.” (Doctor 9, q56) “Some of them won’t prefer to use computers, so I think it would be useful for the patient to have a choice.” (Patient 1, q57)

FG – Focus group participant (clinician)

PICS - Prescribing Information and Communications System

Figure 1. Thematic schema

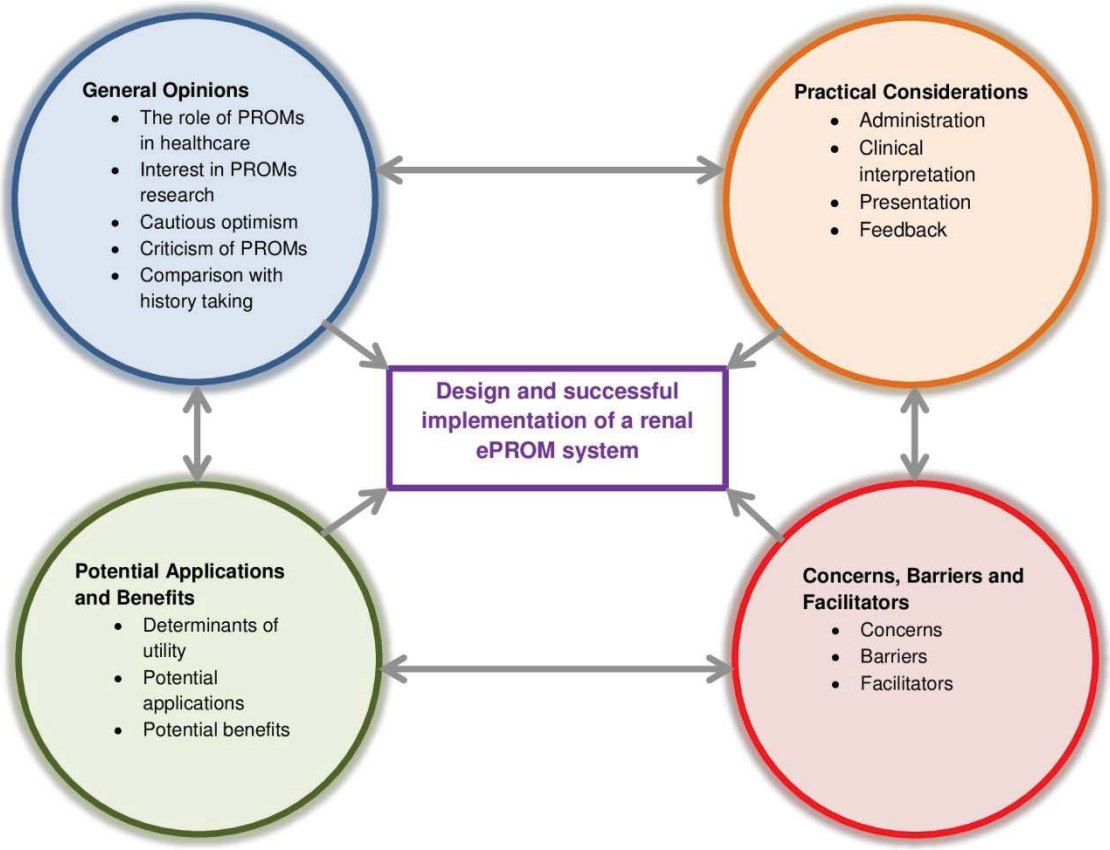


Figure 1. Thematic schema

Supplementary materials

S1. Patient saturation table – Appendix 5

S2. Clinician saturation table – Appendix 6

S3. Supplementary table of quotations – Appendix 7

S4. Comments on selected questionnaires – Appendix 8

Reflexivity

It has been well documented that qualitative researchers often downplay their role and influence on the research process.[1] Phrases such as "emerging themes" are commonly used in a bid to emphasise researcher objectivity and study integrity.[1] However, the reality is that by collecting; curating; analysing; interpreting and reporting qualitative data,[2] researchers are active research participants and have direct and/or indirect influence on the entire process.[1 2]

It is therefore crucial that researchers continuously reflect on their theoretical positions,[1] role and potential influence on the research process.[2] This continuous self-awareness and evaluation has been described as reflexivity.[2 3] Mays and Pope defined reflexivity as "….sensitivity to the ways in which the researcher and the research process have shaped the collected data, including the role of prior assumptions and experience…."[4] It is recommended that reflections on the influence of personal, intellectual biases, personal characteristics on research are reported as this improves the trustworthiness of the findings.[2 4]

As a black, male, medically trained (MBChB) PhD student, I was aware of the potential to unconsciously and indirectly influence the research work. My self-reflections particularly on the qualitative aspects of the project began during their conception and are reported below.

Interaction with patients

Patient perceptions

It is unlikely that my gender, race and age had any substantial influence on the studies given the candid manner the participants responded to my queries (interviews) and/or instructions (usability testing).

My status as an 'outsider' to the patient group did not appear to influence their responses to my queries and none of the patients questioned my ability (directly or

indirectly) to understand their lived experiences of CKD. Similar findings have been reported in literature.[5]

The patient information sheets stated that the research would form part of my doctorate thesis and I introduced myself as a PhD researcher to the patients. This might have had some positive influence on the data collection. There is a strong possibility that the patients' general perception of me as less 'intimidating' than a medical doctor might have facilitated the establishment of rapport with them. This might have encouraged them to share their opinions more freely and in greater detail as they would with their nurses or other non-medical professionals.

Potential influence

I was aware there was risk of unconsciously slipping into a patient-doctor role and made a conscious effort to avoid this situation as it could compromise the study.[6] For this reason, I simply gave my name at the start of each interview or usability session and made no references to my medical background nor gave any indication of my knowledge of the field. I refrained from phrasing questions in the investigative manner often used for taking clinical histories and asked neutral, open-ended questions.

I believe my outsider status made it easier for me to remain as objective as possible in my analysis and reporting of the patients' opinions of the use of PROMs in clinical practice.

Interaction with clinicians

Clinician perceptions

It was unlikely that the clinicians' perceptions of me was influenced my gender or race. All of the clinicians were aware of the fact that I was a PhD student and the majority were aware of my medical background. Therefore with this group of study participants, I might have been perceived as an 'insider'[5] albeit a 'peripheral

member' as I am not a practicing clinician.[7] This meant that, to a degree, I was regarded as a colleague. It is quite possible though that the nurses who participated in the focus group perceived me as an insider to a lesser degree than the doctors given the historical and universal tensions between doctors and nurses.

On the whole, the perception of me as an insider might have aided the establishment of rapport with the clinicians. It also meant that I was expected to understand and sometimes agree with their positions and opinions. Surprisingly, the nurses were not as guarded as I expected while discussing some doctors' disinterest in patients' emotion and psychological problems. They also made it clear that buy-in from doctors is crucial for the successful implementation of an ePROM system.

Potential influence

The focus group was conducted before the one-to-one interviews and provided an initial insight into the opinions of the clinicians. As the focus group also served as a 'comparator' for the subsequent interviews, it was crucial that my influence over the discussions was minimal. The decision was therefore taken to have an independent person moderate the focus group session while I observed the proceedings and made notes. Despite this pre-emptive arrangement, as I was seen as an expert in PRO measurement, some of the focus group participants still asked me for my opinions during the discussions.

For the one-to-one interviews, many of the doctors took the opportunity to ask questions about PROMs. This meant that I had to make a conscious effort to keep the meetings as interviews and not educational sessions for sharing my views and providing information. I consciously kept my questions open-ended and neutral, giving non-committal responses to direct queries during the interviews. I answered the questions that might have been posed by the doctors in the course of the discussions at the end of the interviews.

Having worked 48-hour shifts as a House Officer at a teaching hospital in Nigeria, I understood the pressures and dilemmas the clinicians I interviewed faced practicing in a health system with limited resources. I also understood why some clinicians were concerned that the use of PROMs will increase their work load and potentially raise patient expectations. I realised that without my knowledge of PROMs, gained as a researcher, I would be just as sceptical as some of the clinicians.

I believe being able to view the issues around PROMs as a clinician (insider) and a researcher (outsider) gave me a deeper understanding of the opinions of the clinicians and helped me stay objective throughout the research process.

References

1. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative research in psychology 2006;**3**(2):77-101.
2. Finlay L. “Outing” the Researcher: The Provenance, Process, and Practice of Reflexivity. Qual Health Res 2002;**12**(4):531-45.
3. Mortari L. Reflectivity in Research Practice: An Overview of Different Perspectives. International Journal of Qualitative Methods 2015;**14**(5):1609406915618045.
4. Mays N, Pope C. Assessing quality in qualitative research. BMJ (Clinical research ed) 2000;**320**(7226):50.
5. Dwyer SC, Buckle JL. The Space Between: On Being an Insider-Outsider in Qualitative Research. International Journal of Qualitative Methods 2009;**8**(1):54-63.
6. Britten N. Qualitative Research: Qualitative interviews in medical research. BMJ (Clinical research ed) 1995;**311**(6999):251.
7. Adler P, Adler P. *Membership roles in field research*. Newbury Park, CA: Sage, 1987.

Appendices for Chapter 5	Page
Appendix 5 - Patient saturation table	254
Appendix 6 - Clinician saturation table	263
Appendix 7 - Supplementary table of quotations	281
Appendix 8 - Comments on selected questionnaires	292
Appendix 9 – Illustrative quotes for the selected questionnaires	294
Appendix 10 - HRA approval letter	299
Appendix 11 - Patient invitation letter	307
Appendix 12 - Patient information sheet	308
Appendix 13 - Clinician information sheet	315
Appendix 14 - Patient consent form	322
Appendix 15 - Topic guide for patient interviews	324
Appendix 16 - Topic guide for clinician focus group/interviews	326

Chapter 6: Development and usability testing of a renal ePROM prototype

Development and usability testing of a renal ePROM prototype

Historically, paper-based formats were used for PROMs administration.[1] However, recently, interest in adapting and developing PROMs for electronic administration has surged.[2] The development of ePROM systems has been facilitated by the recent technological advances and an increase in the number of individuals who have access to the internet and electronic devices such as desktop and laptop computers, tablets and smartphones.

The use of ePROMs makes it possible to remotely monitor patients' symptoms/HRQOL and obtain 'real time' data which may provide early warnings of deterioration in patients. This may give clinicians the chance to intervene in a timely manner and potentially delay disease progression.[3-5] Additionally, the use of ePROMs could reduce administrative burden, increase acceptance rates, prevent secondary data entry errors, and lower the incidence of missing data.[1 2 6] However, these benefits could only be realised if ePROM systems are well designed and acceptable to end-users. Therefore, it is essential that the usability of an ePROM system is formally assessed during development by potential end-users to ensure that the system is acceptable and fit for purpose.[2 7] This chapter reports the development and usability testing of a renal ePROM prototype.

References

1. Lee SJ, Kavanaugh A, Lenert L. Electronic and computer-generated patient questionnaires in standard care. Best Practice & Research Clinical Rheumatology 2007;**21**(4):637-47.
2. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(4):419-29.
3. Pittman ZCL, John SG, McIntyre CW. Collection of daily patient reported outcomes is feasible and demonstrates differential patient experience in chronic kidney disease. Hemodialysis International 2017;**21**(2):265-73.

4. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. Journal of Clinical Oncology 2016;**34**(6):557-65.
5. Aiyegbusi OL, Kyte D, Cockwell P, et al. A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures. Curr Opin Nephrol Hypertens 2017;**26**(6):442-49.
6. Velikova G, Wright EP, Smith AB, et al. Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 1999;**17**(3):998-1007.
7. Zbrozek A, Hebert J, Gogates G, et al. Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2013;**16**(4):480-9.

The usability testing reported in this chapter has been published in Computers in Biology and Medicine (Impact Factor (IF) 2.12) as:

Aiyegbusi, O. L., Kyte, D., Cockwell, P., Marshall, T., Dutton, M., Walmsley-Allen, N., Auti, R. and Calvert, M. Development and usability testing of an electronic patient-reported outcome measure (ePROM) system for patients with advanced chronic kidney disease. Computers in Biology and Medicine, 2018. **101**: p. 120-127.

The work reported in this chapter was presented at the following conferences:

- 3rd PROM Conference, Birmingham, UK, June 2018 *(Poster presentation)*
- UK Kidney Week, Harrogate, UK, June 2018 *(Poster presentation)*

Chapter 7: Discussion

7.0 Discussion

This final chapter summarises: the principal findings of the PhD research project; interpretation and implications of the findings; strengths and limitations; and recommendations for future research. The studies presented in this thesis addressed three main aims: (i) evaluation of the evidence to support the use of PROMs in routine renal practice (Chapters 1, 3, 4); (ii) exploration of the opinions of patients and clinicians (Chapter 5); (iii) development and usability testing a renal ePROM prototype with patients (Chapter 6).

7.1 Summary of principal findings

7.1.1. Evidence to support the use of PROMs in patients with CKD

The narrative review[1] (Chapter 1) was conducted to evaluate and summarise the current evidence of 'real life' impacts of PROMs administration on patient outcomes and clinical processes in routine renal practice. The narrative review included studies which demonstrated the acceptability and feasibility of administering ePROMs data in routine renal practice.[2 3] These studies also supported the view that the 'real-time' utilisation of ePROMs data may assist with the detection of clinically relevant changes in patients' conditions by providing early warnings of issues and reduce the need for patient follow-up visits.[3 4] Emerging evidence from an oncology setting suggested that the utilization of ePROMs data could potentially reduce emergency hospitalization rates and improve HRQOL and survival.[5]

The systematic review[6] (Chapter 3) was conducted to provide the most up-to-date psychometric evidence of measurement properties of PROMs used in CKD in order to facilitate the selection of appropriate PROMs for subsequent use in the project. Based on the findings of the systematic review, the KDQOL-36 was supported for

use in patients with CKD (pre-dialysis) as it had adequate evidence for internal consistency and construct validity (though further validation might be necessary). Both the KDQOL-SF and the KDQOL-36 were supported for use in patients receiving dialysis as the KDQOL-SF demonstrated adequate evidence for internal consistency, structural validity, test-retest reliability and construct validity while the KDQOL-36 had adequate evidence for internal consistency, test-retest reliability and construct validity. The ESRD-SCLTM had adequate evidence for internal consistency, test-retest reliability, structural and construct validity. It was therefore recommended for use in renal transplant recipients who may experience symptoms specific to renal transplantation and immunosuppression therapy.[6]

The review, however, also identified methodological issues with many of the included studies. These included: utilization of sample sizes smaller than current recommendations; poor reporting of missing data; inadequate or no analysis of measurement properties such as measurement error and responsiveness; and poorly articulated *a priori* hypotheses.

Primary analyses of the RIISC dataset (Chapter 4) was conducted to explore the predictive capability of HRQOL data for CKD outcomes (progression and mortality) and obtain statistical evidence of the relationship between HRQOL data and clinical and laboratory variables. Survival analyses found that, along with other clinical and laboratory variables, problems with the EQ-5D dimensions of self-care and usual activities significantly increased the risk for all-cause mortality. The likelihood of reporting problems with self-care increased significantly with unemployment, retirement, co-morbidities (diabetes, COPD), and a higher BMI. Similarly, unemployment, retirement and co-morbidities (diabetes, CVD, IHD) were

significantly associated with an increased likelihood of reporting problems with usual activities.

7.1.2. Opinions of patients and clinicians on the use of ePROMs in routine renal practice

Findings from the qualitative research conducted with patients and clinicians were reported in Chapter 5. Patients expressed a willingness to complete ePROMs on a regular basis concurring with the finding reported in the narrative review[1] (Chapter 1). Interviewees also viewed the selected questionnaires (including the KDQOL-36 recommended by the systematic review[6]) as acceptable and relevant. Some patients were interested in using ePROMs data for self-monitoring and management. Clinicians acknowledged that PROMs may have a potentially important role to play in the future care of patients with CKD. However, some were of the opinion that PROMs might only play a limited role in decision-making as they were not convinced there was enough evidence to support the use of ePROMs. It was interesting to note that this view was expressed by clinicians despite admitting to possessing a limited knowledge and experience of PROMs. Very few were familiar with the KDQOL-36 or the IPOS-Renal and only three had previously used a PROM for clinical practice or research purposes. The studies reported in chapters 1, 3, 4 have demonstrated the existence of evidence that (i) there are PROMs available with reasonably robust validity/reliability which may be useful for monitoring patients with CKD. However, further validation is required in the UK population, (ii) PROMs are capable of predicting CKD outcomes (mortality), which could assist care management. However, further research is required to explore this predictive ability in disease-specific PROMs, (iii) patients and clinicians are broadly supportive of ePROM use in

practice. However, clinicians in particular would value a stronger evidence-base to increase engagement.

Some clinicians worried that the use of ePROMs would 'medicalise' what they regarded as normal emotional/psychological response to CKD diagnosis and management. Clinicians also expressed concerns that the use of an ePROM system could expose them to the risk of litigation if patients expressed adverse outcomes. Patients and clinicians identified numerous potential benefits, facilitators and barriers.

7.1.3. Usability testing of a renal ePROM prototype

The usability testing reported in Chapter 6 found that patients regarded the renal ePROM system as acceptable and easy to use thus further concurring with the findings from the narrative review (Chapter 1) and the qualitative interviews (Chapter 5). Patients required a mean time of 10 (± 1.6) minutes to complete the electronic version of the KDQOL-36 and made relatively minimal errors. The average usability and satisfaction score was 4.6 on a 5-point scale (where 5 = very satisfied and 1 = very dissatisfied).

7.2 Interpretation and implications of findings

7.2.1. Assessment of evidence supporting the use of PROMs

The narrative review confirmed that PROM research is progressing from the era of theorising about the potential benefits of PROMs [7-9], and identifying the barriers and facilitators of implementation[10] to providing concrete evidence of these benefits in routine clinical practice.[3-5 11 12]

At the beginning of the PhD project, a review by Gibbons and Fitzpatrick[13] was identified. However, this was conducted in 2009 and important methodological guidelines[14-18] and validation studies have since been published.[19] It was therefore necessary to conduct a systematic review in order to obtain the best evidence to inform PROM selection for this project.

The KDQOL-36, KDQOL-SF and the ESRD-SCLTM were found to have the adequate evidence to support their use and were therefore recommended. These three measures are disease-specific with symptoms-related dimensions which might increase their sensitivity to changes in the relevant CKD populations.[20 21] The systematic review highlighted a number of methodological issues similar to those reported by other authors.[13 22-24] It is quite possible that some measures that performed poorly in the present review might score better if these methodological issues were addressed in subsequent validation studies.

The IPOS-Renal[25] was still undergoing validation when the systematic review was published.[6] The measure was validated in Australian in patients with advanced CKD (on dialysis and supportive care) and nurses. The results which have since been published show that the patient and staff versions of the IPOS-Renal have good test-retest reliability, internal consistency, and construct validity.[25]

Although the systematic review did not find sufficient evidence to support the use of the EQ-5D (as a result of the limited number and methodological quality of the relevant validation studies), results of the RIISC data analysis (Chapter 4) and published studies by Basch et al.,[5] and Jesky et al.[26] show that it could provide useful HRQOL data. The RIISC analyses (Chapter 4) showed significant associations between HRQOL (EQ-5D dimensions of self-care and usual activities) and the risk of death as demonstrated by previous studies.[12 26-28] The socio-

demographic predictors of problems with self-care and usual activities such as employment and co-morbidities were similar to those reported by Jesky et al.[26] The KDQOL-36, (recommended in Chapter 3), includes questions that capture the domain of usual activities which was identified as statistically significant in the RIISC analyses of the EQ-5D (Chapter 4). The existence of significant relationships between PROM data and clinical outcomes and variables mean that patients at risk of rapid deterioration may be identified through PROM data before their clinical and/or laboratory data are available as suggested by findings from oncology studies.[11 29]

7.2.2 Implementation of an ePROM intervention

Damschroder et al. developed the Consolidated Framework For Implementation Research (CFIR), a comprehensive framework based on existing implementation theories, which may be used for the implementation of healthcare interventions.[30] The CFIR has five major domains namely: intervention characteristics; outer setting (socio-economic and political context the organisation is situated);[30] inner setting (structural, political, and cultural contexts for the implementation process);[30] characteristics of the individuals involved; and the process of implementation.[30] Foster et al. used the CFIR to explore potential barriers, facilitators of PROMs implementation and relevant issues such as instrument selection and purpose of use.[31]

The development and implementation of a PROMs intervention in clinical practice commonly involves the selection of suitable existing measure(s).[21 31 32] This task belongs to the process of implementation domain.[30] As most existing PROMs are copyrighted, ePROM developers may need to obtain formal permissions from

original developers and sign agreements before they can begin the process of converting paper PROMs to ePROMs. The selection process for PROMs needs to be considered carefully and backed by up-to-date evidence if any benefit is to be derived from their use.[6 9 31]

The work reported in Chapters 1, 3, and 4, assessed the level of evidence to support the use of PROMs in routine clinical practice, and comes under the intervention characteristics domain of the CFIR.[30] The review (Chapter 3) included 25 PROMs assessing various aspects of HRQOL[6] indicating how daunting the task of PROM selection could be. While it is important that the psychometric properties (i.e. reliability and validity) of PROMs are evaluated during the selection process, a number of practical issues such as the purpose of use (clinical trials, benchmarking, routine practice or product labelling) and target population need to be considered as well.[21 31 32] For instance, although the systematic review[6] found evidence to support the use of both the KDQOL-36 and the KDQOL-SF in dialysis patients; the shorter KDQOL-36 may be more practical for use in routine renal setting. The longer KDQOL-SF might be the preferred option for clinical trials where HRQOL is an endpoint and more information is required. In situations where utility or economic evaluations might be required, the EQ-5D might be the preferred measure of HRQOL. These considerations belong to the inner setting of the CFIR.[30]

Another important issue to consider when selecting PROMs, is the domains of interest. Individual PROM instruments combine a variety of HRQOL domains such as physical function, emotional function and symptoms. This means that clinicians, researchers and patients need to consider which measure best combines the domains of interest.[32] It was observed that there was a tendency for the studies included in the narrative review (Chapter 1) to employ questionnaires which

incorporate symptoms such as the KDQOL-36 recommended in the systematic review[6] and the brief EQ-5D measure. This might be due to the fact that patients and current KDIGO guidelines consider symptom management a care priority.[33 34] External organisations such as KDIGO belongs to the outer setting domain of the CFIR.[30]

As mentioned in the systematic review (Chapter 3), it is quite common for clinicians, researchers and patients to discover that no single PROM covers all the domains that might be of particular interest or relevance.[6] In such situations, there are several options – modify or combine existing measures or develop a new one. The systematic review included both measures developed de novo such as the SF-36 and modified measures such as the KDQOL-36 (which has at its core the SF-12 which is derived from the SF-36). Two of the studies included in the narrative review utilised ePROM systems that combined more than one PROM.[2 3] The qualitative protocol (Chapter 2) and the usability paper (Chapter 6) described the process of developing and testing the renal ePROM prototype which incorporated the KDQOL-36 and the IPOS-Renal. However, the long term objective is to have an efficient, but precise, tool that is non-burdensome and only includes the domains/items of importance to the target population and their care team.

While developing a new PROM might ensure that the final measure meets the needs of its target population and removes the requirement to obtain permissions from a third party, the process requires a huge amount of resources in terms of expertise, finance and time. It should therefore be considered only where there is a strong justification. The less intensive options of adapting or combining existing measures are not without their own issues. Original developers may grant permissions with clauses that restrict the modification of the content or layout of the original measures.

Chapters 5 and 6 explore the opinions of stakeholders which belong to the CFIR domain characteristics of the individuals involved. Using more than one PROM might increase questionnaire burden on patients who are already burdened by their condition. Concerns about questionnaire burden on patients were raised by clinician in the qualitative study (Chapter 5) even though patients did not perceive the individual questionnaires as burdensome. It is quite possible that patients downplayed concerns about questionnaire burden due to feelings of obligation or a desire to please their clinicians as suggested by some of the patients and clinicians interviewed. The CFIR may assist with the successful delivery of the ePROM system in future by highlighting the various domains that may require consideration and intervention.[30 31]

7.2.3. Development of an ePROM intervention - Input from stakeholders

The qualitative research (Chapter 5) provided a deeper understanding of the opinions of patients and renal clinicians on the use of ePROMs in routine renal practice. Insights from this research were intended to inform the design and implementation of the renal ePROM system in order to ensure its success post-implementation.[35] Findings from the qualitative study (Chapter 5) and the usability testing[36] (Chapter 6) show that patients may perceive the use of a renal ePROM system as an acceptable component of their routine clinical management. The qualitative study revealed patient interest in using ePROMs for self-monitoring and management which could have a positive influence on their engagement with health care and vice-versa. This may in turn enhance patient adherence to treatment and potentially clinical outcomes.[37]

The emerging evidence cited in the narrative review (Chapter 1), that the use of ePROMs may reduce the frequency of hospital appointments[3], was supported by patients and clinicians (Chapter 5). They suggested that remote monitoring of stable patients with the aid of ePROMs could potentially free up appointment slots for patients who really need to be seen in clinic. This would also reduce the need for stable patients to take time of work, thus securing their jobs and income. Evidence from other types of remote monitoring suggest benefits to a range of patients for example, a review of remote monitoring of parameters such as blood pressure and medication use in patients with heart failure reported lower admission rates, reduced healthcare costs and no evidence of negative impact on patients' health.[38] The successful interventions were easy to use which made them were widely accepted by patients, including elderly patients.[38] McManus et al. found that the use of remote monitoring of blood pressure for the purpose of titration of antihypertensive medication led to significantly lower blood pressure than titration guided by clinic readings in patients with poorly controlled blood pressure.[39] The trial also found that remote monitoring served as reminders for patients to take their medication thereby improving medication compliance.[39] The simple text-based telemonitoring system was designed to be independent of mobile phone providers or clinical systems so that it could be implemented in other settings.[39] Based on these approaches, remote monitoring interventions need to be easy to use and flexible in design to improve their adoption.[38 39]

The usability testing[36] reported in Chapter 6 provided patient feedback on the renal ePROM prototype. Findings from the study showed that some commonly held beliefs about the technological capabilities of elderly patients may no longer hold.[36] Participants who tested the ePROM prototype had a mean age of 64 years with

varying levels of experience and usage of the internet and electronic devices. All described the system as easy to use regardless of their age and technological experience. The participants required a mean time of 10 min to complete the electronic version of the KDQOL-36 which is shorter than the mean time of 15 min reported in literature for the paper format.[40] Similar acceptance rates and technological capabilities have been reported by other studies for this age group.[2 41] An explanation for these findings is that the digital divide between older and younger populations decreases as individuals take the information technology (IT) skills they gain through work and study with them into retirement.[42 43] This means that the population of elderly patients with sufficient IT capabilities to complete an ePROM grows with each passing decade.[42 43] Digital divide has been described as the gap in access to digital technology between individuals or groups with access and those without.[44] Digital divide can be observed at national or global level. It has been argued that with the rapid growth in IT, the digital divide has led to digital inequality which has important socio-economic implications.[44 45] Predictors of digital divide/inequality are: age, education, income level, gender, race or ethnicity, and rural or urban residence.[43 45] Interventions to bridge the digital divide need to focus on providing access (physical and financial) as well as technical support which will aid the acquisition of IT skills.[44]

The qualitative study (Chapter 5) also revealed a high level of awareness and interest in PROMs and PROMs research among clinicians despite their reservations. This is evidenced by the increase in editorials and opinion articles on PROMs written by medical doctors and published in high impact journals.[35 46-48] Increase in clinician awareness and engagement, identified as crucial for the successful implementation of PROMs in clinical practice[35], might explain the recent success of a number of

ePROMs programmes.[49 50] These programmes are now in a position to provide the much needed 'real-life' evidence to support the use of PROMs in clinical practice as demonstrated by the findings of the narrative review (Chapter 1).[3 5]

While the majority of clinicians acknowledged that ePROMs data relating to HRQOL dimensions of symptoms and physical functioning could be potentially beneficial in patient management, a number of clinicians were less convinced about the usefulness of data relating to emotional/psychological wellbeing. Their main argument was that addressing problems with symptoms/physical functioning could be more straightforward and achievable. This might explain the finding by Salmon et al. that general practice doctors were reluctant to pick up on cues of psychological distress in their patients.[51] Our findings suggest that secondary care clinicians may feel more comfortable focusing on physical aspects of their specialist clinical area rather than broader holistic issues for overall body. Wong et al. found that the physical component of SF-36, unlike the mental component, deteriorated with time in patients with CKD making it a better tool for the assessment of interventions and health service evaluations.[52]

One of the major challenges facing the implementation of ePROM systems in routine practice is its optimal integration with existing care pathways and systems. Findings from the qualitative study (Chapter 5) showed that clinicians are very unlikely to access ePROM data if it is not provided within the existing patient information system (Clinical Portal). The Transforming Participation in Chronic Kidney Disease (TP-CKD) programme reported the various ways ePROMs (combining the EQ-5D, POS-symptom module and Patient Activation Measure) were incorporated into patient care at 14 renal units.[53] In line with the suggestions from the clinicians interviewed for this PhD research, the ePROM data was used by the renal units at

quarterly multidisciplinary meetings to inform individual care planning; included in letters to patients about their care plan; and reviewed by nurses at haemodialysis sessions flagging scores that warranted discussion during subsequent consultation.[53] As suggested by the qualitative study, clinicians who participated in the TP-CKD programme reported difficulties interpreting ePROM data and required on-site workshops to facilitate the process.

Socio-technical approach to system design considers the complex relationships between human, social and technical factors which may determine the functionality and adoption of computer-based systems such as healthcare systems.[54] Human-centred design is one of the most commonly used socio-technical system design methods in health-care.[55] It employs a clear understanding of users, their tasks, and the environments in which those tasks are carried out.[55] Das and Svanæs 2013 reported the use of human-centred design in the development of an e-health weight loss intervention and noted that the involvement of patients throughout the design process is crucial and must consider personal experiences and vulnerabilities.[56]

7.3 Strengths and limitations

Previous research into the use of ePROMs in CKD has been minimal. The research reported in this thesis therefore contributes novel and important findings to the existing body of evidence to support the use of ePROMs in the routine clinical management of patients with advanced CKD. The systematic review, qualitative study and usability testing were conducted according to the published peer-reviewed protocols (Chapter 2). The use of a mixed-methods approach, involving systematic review, quantitative and qualitative methods, has enhanced the robustness of the project ensuring that its findings and conclusions are reliable and valid. The

involvement of patients early in the research guided its design and conduct ensuring that the findings from each study were relevant and beneficial.

The findings from this research have been disseminated as follows:

- Two oral presentations and nine poster presentations at national and international conferences including:
    - The International Society for Quality of Life Research (ISOQOL) Annual Conference, Copenhagen, Denmark, Oct 2016
    - UK Kidney Week, Liverpool 2017, Harrogate 2018
    - PROMs Conference, London 2017, Birmingham 2018

The narrative review (Chapter 1) provided real life evidence of the impact of ePROMs on patient outcomes.[1] However, as it was not systematically conducted, i.e. no search strategy was used or formal screening conducted, it is possible that relevant studies might have been missed. A few CKD-related PROM studies have been published since the narrative (and systematic) reviews (Chapters 1 & 3) were published. However, these articles are cited and discussed in the RIISC data analyses (Chapter 4); the qualitative study (Chapter 5) and the various sections of this concluding chapter.[37 47 52 57-59]

The use of current methodological guidelines and checklists such as the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist[14] and the criteria for evidence synthesis[18] ensured that the systematic review[6] was robust and provided the highest evidence possible (Chapter 3). The study was registered on the PROSPERO website and reported according to PRISMA guidelines[60 61] demonstrating transparency in its conduct and reducing the risk of reporting bias. The review was broad in scope, evaluating PROMs within three distinct groups of patients with CKD: i.e. pre-dialysis, dialysis and post-transplantation. There were no language restrictions and the search strategy was

developed (using validated search filters) and pre-tested to minimise the chances of missing important eligible studies. The main limitation of the systematic review was the fact that a number of measurement properties were not adequately investigated and/or reported by the included studies making it difficult and sometimes impossible to evaluate these elements. In addition, very few studies reported validation work in pre-dialysis patients. Therefore, a stronger recommendation of the KDQOL-36 could not be made for this group of patients. Reporting the results of the systematic review by CKD stage might have been a more useful approach for clinicians and CKD researchers. However, this was impossible as there was a wide variation in the approach taken by the included studies and it was often unclear what classification methods were employed.

The findings of the RIISC data analyses (Chapter 4) was reported according to the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines for transparent reporting.[62] The use of data from a relatively large, ethnically diverse sample improved the generalizability of the findings. Using a variety of statistical methods strengthened the validity of the results. The longer follow up period meant that more outcome events had occurred by the censor date which also improved the robustness of the analyses. The analyses were limited by the fact that causal inferences cannot be made from the results as the study, like all observational studies, explored association and not causation.

The qualitative work with patients and clinicians (Chapter 5) was reported according to the Consolidated criteria for reporting qualitative research (COREQ) guidelines for transparent reporting.[63] Due to the subjective nature of qualitative enquiry it was susceptible to researcher bias. However, undue influence on the data collection,

analysis and interpretation was mitigated by the use of interview guides, regular discussions with members of the research team and triangulation of findings.

The usability testing (Chapter 6) was conducted following the recommended guidelines in literature.[64-66] However, the study was limited by the fact that the tests were conducted on-site in an interviewer-controlled setting. Therefore, participants' usability performance and experience may differ in the absence of interviewer prompts in a ‘real life’ home setting. The small sample size also limited the statistical analysis that could be performed on the data. It should be noted, however, that the primary purpose of conducting the study was to access acceptability and usability and not to explore statistical inferences.

The qualitative study (Chapter 5) and the usability testing (Chapter 6) explored patient and clinician views pre-implementation which may differ once the renal system is implemented and used under ‘real life’ conditions. However, it was important to obtain these early views as research has shown that initial attitudes may influence the subsequent adoption of technological innovations.[64 65 67-69] In addition, the findings from both studies may assist with the development of the renal ePROM prototype and ensure that the final version is acceptable to the stakeholders (patients and clinicians).

7.4 Future research

During the course of conducting the research for this thesis, a number of research gaps were identified. The increasing interest in the use of ePROM systems means that more healthcare providers will develop their own systems in the future. There is therefore a need for further research to guide the development and implementation of ePROM systems. There is also a need for more real-life evidence of the benefits that may be derived from using ePROM systems in routine renal practice.

7.4.1. Development of ePROM systems

Psychometric properties of PROMs

The systematic review reported in Chapter 3 revealed a number of gaps in psychometric evidence to support the measures currently used in CKD. Little or no information on the assessment of measurement error or responsiveness was found for the included measures. When choosing PROMs for patient monitoring, knowing the responsiveness of measures i.e. ability to detect clinically important changes in patients' condition over time,[70] would assist with the selection of the most responsive measure. It is also important to have adequate information on measurement error as this helps to determine whether a measure is capable of detecting true clinically relevant changes in patients.[71] It was also clear from the findings that there is a lack of consensus on how these psychometric properties should be assessed, the definition, measurement and interpretation of parameters such as the minimal important change (MIC), the minimally detectable change (MDC), the minimal clinically meaningful difference (MCMD) or the Minimal Clinically Important Differences (MCID).[72-81] The protocol for a study which aims to review anchor-based methods of estimating minimal important difference (MID) and develop an instrument to evaluate the credibility of MIDs was published three years ago.[82] The results of

this study are yet to be published. Reaching a consensus and having a clear guidance on these methodological issues will assist researchers and clinicians with the design of methodologically sound development/validation studies and improve the quality of evidence for PROMs (electronic and paper). The COSMIN checklist has recently been revised and the new tool places less emphasis on aspects such as reporting quality and a priori hypothesis which the COSMIN steering committee concluded would not necessarily lead to biased results.[83] Future systematic reviews should employ this checklist in the assessment of psychometric properties of PROMs. Improving the psychometric performance of measures may ensure that hypothesised benefits are obtained in real life.

Computer Adaptive Testing (CAT)

As discussed earlier, ePROM developers and clinicians often combine questionnaires in order to cover all the HRQOL domains of interest. Such combinations may significantly increase patient burden especially if a number of the questions are not relevant to the individual patient. Questionnaire burden on patients has been a persistent concern for clinicians as demonstrated by the findings from the qualitative study reported in Chapter 5. Computer adaptive testing (CAT) which utilises algorithms to match the most relevant questions to respondents has been suggested as a possible solution.[84-86] The narrative review reported in Chapter 1 suggested further research into CAT. While a few exploratory studies have been published since the narrative review was conducted[58 84-87], there is very limited evidence on the development of CAT interventions and use in routine renal practice.

PPI in PROMs research

Patient and public involvement (PPI) in research is defined as "...research being carried out '**with**' or '**by**' patients and members of the public rather than '**to**', '**about**' or '**for**' them."[88] PPI has been recognised as a vital activity that may facilitate the acceptability, relevance, timeliness, and quality of research.[89-92] Evidence of the impact of PPI in research is steadily increasing as more reviews [91 93] and realist evaluations are being published.[94 95] A recent scoping review found that overall PPI in PROMs development varied considerably.[96] Of the studies reviewed[96], 58.5% involved patients in item development while 10.9% had PPI input on the determination of outcomes to measure and only 6.7% had patient involvement in all aspects of PROM development.[96] For the renal field, only 10 of the 66 studies included in the systematic review (Chapter 3) reported patient involvement in the process of item generation and/or item selection.[6] These findings show that more work is needed to address the currently suboptimal levels of patient involvement. Evaluation studies are required to determine the best practice and develop clear practical guidelines for conducting PPI in PROMs research.[92]

7.4.2. Implementation of ePROMs intervention

Administration of ePROMs

The qualitative study (Chapter 5) reported no consensus on a number of practical issues such as the frequency and timing of ePROMs administration in routine renal practice. The narrative review (Chapter 1) highlighted the fact that despite the availability of published frameworks and guidelines[97-99] to assist with the implementation of PROM interventions, there is limited evidence from actual implementation studies. More research is needed to evaluate the effectiveness of the various options for each of these issues to facilitate appropriate recommendations.

Interpretation and utilisation of ePROM data

Findings from the qualitative work with clinicians (Chapter 5) confirmed the conclusion by previous studies that the interpretation and utilisation of ePROM data are key challenges facing its adoption.[1 100] Clinicians tend to perceive PROMs data as difficult to interpret and relate to clinical decision-making thus potentially discouraging the adoption of ePROM systems.[101] Efforts are being made to develop training programmes for clinicians on the interpretation and utilisation of ePROM data.[102] Research is needed to develop and evaluate these training programmes and identify the most effective approaches for recommendation.

Clinician engagement

While training clinicians how to interpret and use ePROMs data might foster engagement with such interventions[101] and vice-versa, clinician scepticism might be a barrier to the success of such training programmes. Overcoming scepticism and increasing clinician engagement has been identified as a strong facilitator of the successful implementation of ePROM systems.[1 35] There is therefore an urgent need for research into developing innovative and practical approaches to effectively deal with clinician scepticism and optimise their engagement with ePROMs interventions.

7.4.3. Real life implementation of ePROMs systems

Clinicians interviewed as part of the qualitative study (Chapter 5) were unconvinced that there was a strong evidence base underpinning the use of ePROMs in routine clinical practice. This perception might account to a large extent for the clinician scepticism that has been documented in literature.[1 35] As discussed in previous sections, the evidence to support the use of ePROMs is increasing.[3 5 29] More high quality trials of ePROM systems in routine practice settings are required to generate

further evidence of impact on patient outcomes and other variables such as hospitalization rates. There is also a need for cost-effectiveness analyses to demonstrate economic benefits of ePROMs in clinical practice. It is hoped that the renal ePROM (RePROM) trial being developed by researchers at the Centre for Patient Reported Outcome Research (CPROR), UHB, and Birmingham Clinical Trials Unit (BCTU) may provide this much needed evidence in due course. However, the qualitative study showed that conducting high quality trials and publishing the results in academic journals may not be sufficient in isolation, for example, none of the clinicians interviewed were aware of the studies referenced above by Basch et al. or Schougaard et al. A systematic review, which this author contributed to as a second reviewer, analysed the impact of healthcare research and presented a methodological framework for research impact (See appendix 17).[103] There is a need for research into innovative methods of effectively disseminating the evidence to support the use of ePROMs systems to all stakeholders (including clinicians and patients) following the framework proposed by Cruz Rivera et al.[103] Patients and clinicians as mentioned earlier expressed a keen interest in this PROM research and its findings. A variety of methods are being considered for the dissemination of the findings described in this thesis. Funding from the University of Birmingham's Wellcome Trust Institutional Strategic Support Fund (ISSF) for public engagement has been secured for an interactive engagement event with patients with CKD and clinicians from UHB. Findings from this PhD research and relevant research from CPROR and UHB will be disseminated at this event.

7.5 Conclusions

The studies presented in this thesis revealed a growing evidence base to support the use of PROMs in routine renal practice. Psychometric, statistical and real life evidence was obtained by utilising a mixed-methods approach for the thesis. The systematic review found sufficient evidence to support the use of KDQOL-36, KDQOL-SF and the ESRD-SCLTM for assessing the HRQOL of patients with CKD. Statistical evidence of the predictive capabilities of HRQOL domains was obtained from the analysis of the RIISC dataset which could assist with risk stratification in patients with CKD. Real life evidence was obtained by reviewing the findings of published PROM research.

This PhD research confirmed a strong interest in the use of ePROMs among stakeholders (patients and clinicians). The usability study further demonstrated the high acceptability of a renal ePROM prototype. Previously unreported views were obtained and explored through the qualitative work with stakeholders. The research highlighted the potential applications, benefits and issues to be considered and addressed to ensure the successful development and implementation of an ePROMs intervention in clinical practice.

There is a need for high quality trials in clinical practice to provide further evidence to support the routine use of ePROMs. Patients and clinicians need to be actively involved in ePROMs research and development and efforts should be made to incorporate their views to ensure that the interventions are acceptable, relevant and beneficial.

References

1. Aiyegbusi OL, Kyte D, Cockwell P, et al. A patient-centred approach to measuring quality in kidney care: patient-reported outcome measures and patient-reported experience measures. Curr Opin Nephrol Hypertens 2017;**26**(6):442-49.
2. Schick-Makaroff K, Molzahn A. Brief communication: patient satisfaction with the use of tablet computers: a pilot study in two outpatient home dialysis clinics. Canadian Journal of Kidney Health and Disease 2014;**1**:22.
3. Schougaard LM, Larsen LP, Jessen A, et al. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2016;**25**(3):525-34.
4. Pittman ZCL, John SG, McIntyre CW. Collection of daily patient reported outcomes is feasible and demonstrates differential patient experience in chronic kidney disease. Hemodialysis International 2017;**21**(2):265-73.
5. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. Journal of Clinical Oncology 2016;**34**(6):557-65.
6. Aiyegbusi OL, Kyte D, Cockwell P, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PloS one 2017;**12**(6):e0179733.
7. Snyder CF, Aaronson NK. Use of patient-reported outcomes in clinical practice. Lancet (London, England) 2009;**374**(9687):369-70.
8. Black N. Patient reported outcome measures could help transform healthcare. BMJ (Clinical research ed) 2013;**346**:f167.
9. Greenhalgh J, Long AF, Brettle AJ, et al. Reviewing and selecting outcome measures for use in routine practice. Journal of evaluation in clinical practice 1998;**4**(4):339-50.
10. Antunes B, Harding R, Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: A systematic review of facilitators and barriers. Palliative Medicine 2014;**28**(2):158-75.
11. Basch E, Jia X, Heller G, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. Journal of the National Cancer Institute 2009;**101**(23):1624-32.
12. Porter AC, Lash JP, Xie D, et al. Predictors and Outcomes of Health–Related Quality of Life in Adults with CKD. Clinical journal of the American Society of Nephrology : CJASN 2016;**11**(7):1154-62.
13. Gibbons E, Fitzpatrick , R. A structured review of patient-reported outcome measures for people with chronic kidney disease. Department of Public Health University of Oxford 2010.
14. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2010;**19**(4):539-49.
15. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of clinical epidemiology 2010;**63**(7):737-45.
16. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2013;**22**(8):1889-905.
17. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology 2007;**60**(1):34-42.

18. Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2012;**21**(4):651-7.
19. Moher D, Tsertsvadze A, Tricco AC, et al. When and how to update systematic reviews. The Cochrane database of systematic reviews 2008(1):Mr000023.
20. Bryan S, Davis J, Broesch J, et al. Choosing your partner for the PROM: a review of evidence on patient-reported outcome measures for use in primary and community care. Healthcare Policy 2014;**10**(2):38-51.
21. Snyder CF, Aaronson NK, Choucair AK, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2012;**21**(8):1305-14.
22. Elbers RG, Rietberg MB, van Wegen EE, et al. Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2012;**21**(6):925-44.
23. Green A, Liles C, Rushton A, et al. Measurement properties of patient-reported outcome measures (PROMS) in Patellofemoral Pain Syndrome: a systematic review. Man Ther 2014;**19**(6):517-26.
24. Janssens A, Rogers M, Gumm R, et al. Measurement properties of multidimensional patient-reported outcome measures in neurodisability: a systematic review of evaluation studies. Dev Med Child Neurol 2016;**58**(5):437-51.
25. Raj R, Ahuja K, Frandsen M, et al. Validation of the IPOS-Renal Symptom Survey in advanced kidney disease: a cross-sectional study. Journal of Pain and Symptom Management;**56**(2):281-87.
26. Jesky MD, Dutton M, Dasgupta I, et al. Health-Related Quality of Life Impacts Mortality but Not Progression to End-Stage Renal Disease in Pre-Dialysis Chronic Kidney Disease: A Prospective Observational Study. PloS one 2016;**11**(11):e0165675.
27. Tsai Y-C, Hung C-C, Hwang S-J, et al. Quality of life predicts risks of end-stage renal disease and mortality in patients with chronic kidney disease. Nephrology Dialysis Transplantation 2010;**25**(5):1621-26.
28. Berg SK, Thorup CB, Borregaard B, et al. Patient-reported outcomes are independent predictors of one-year mortality and cardiac events across cardiac diagnoses: Findings from the national DenHeart survey. Eur J Prev Cardiol 2018:2047487318769766.
29. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. Journal of the National Cancer Institute 2011;**103**(24):1808-10.
30. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;**4**(1):50.
31. Foster A, Croot L, Brazier J, et al. The facilitators and barriers to implementing patient reported outcome measures in organisations delivering health related services: a systematic review of reviews. J Patient Rep Outcomes 2018;**2**:46.
32. Snyder CF, Watson ME, Jackson JD, et al. Patient-reported outcome instrument selection: designing a measurement strategy. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2007;**10 Suppl 2**:S76-85.
33. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. Clinical journal of the American Society of Nephrology : CJASN 2014;**9**(10):1813-21.
34. Stevens PE, Levin A. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. Annals of internal medicine 2013;**158**(11):825-30.

35. Rotenstein LS, Agarwal A, O'Neil K, et al. Implementing patient-reported outcome surveys as part of routine care: lessons from an academic radiation oncology department. J Am Med Inform Assoc 2017;**24**(5):964-68.
36. Aiyegbusi OL, Kyte D, Cockwell P, et al. Development and usability testing of an electronic patient-reported outcome measure (ePROM) system for patients with advanced chronic kidney disease. Computers in Biology and Medicine 2018;**101**:120-27.
37. Schrauben SJ, Hsu JY, Rosas SE, et al. CKD Self-management: Phenotypes and Associations With Clinical Outcomes. American Journal of Kidney Diseases 2018;**72**(3):360-70.
38. Martínez A, Everss E, Rojo-Álvarez JL, et al. A systematic review of the literature on home monitoring for patients with heart failure. Journal of Telemedicine and Telecare 2006;**12**(5):234-41.
39. McManus RJ, Mant J, Franssen M, et al. Efficacy of self-monitored blood pressure, with or without telemonitoring, for titration of antihypertensive medication (TASMINH4): an unmasked randomised controlled trial. The Lancet 2018;**391**(10124):949-59.
40. Thaweethamcharoen T, Srimongkol W, Noparatayaporn P, et al. Validity and reliability of KDQOL-36 in thai kidney disease patient. Value in Health Regional Issues 2013;**2**(1):98-102.
41. Wong D, Cao S, Ford H, et al. Exploring the use of tablet computer-based electronic data capture system to assess patient reported measures among patients with chronic kidney disease: a pilot study. BMC nephrology 2017;**18**(1):356.
42. Gatto SL, Tak SH. Computer, Internet, and E-mail Use Among Older Adults: Benefits and Barriers. Educational Gerontology 2008;**34**(9):800-11.
43. Hong YA, Cho J. Has the Digital Health Divide Widened? Trends of Health-Related Internet Use Among Older Adults From 2003 to 2011. The Journals of Gerontology: Series B 2017;**72**(5):856-63.
44. Hargittai E. The digital divide and what to do about it. In: Jones DC, ed. The New Economy Handbook. San Diego, CA: Academic Press, 2003.
45. DiMaggio P, Hargittai E. From the 'Digital Divide' to 'Digital Inequality': Studying Internet Use as Penetration Increases. Woodrow Wilson School of Public and International Affairs, Center for Arts and Cultural Policy Studies, Princeton University 2001.
46. Finkelstein FO, Finkelstein SH. Time to Rethink Our Approach to Patient-Reported Outcome Measures for ESRD. Clinical Journal of the American Society of Nephrology 2017.
47. Finkelstein FO, Finkelstein SH. Assessing Fatigue in the ESRD Patient: A Step Forward. American Journal of Kidney Diseases 2018;**71**(3):306-08.
48. Rotenstein LS, Huckman RS, Wagle NW. Making Patients and Doctors Happier — The Potential of Patient-Reported Outcomes. New England Journal of Medicine 2017;**377**(14):1309-12.
49. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA: a cancer journal for clinicians 2012;**62**(5):337-47.
50. Hjollund NH, Larsen LP, Biering K, et al. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. Interactive journal of medical research 2014;**3**(1):e5.
51. Salmon P, Dowrick CF, Ring A, et al. Voiced but unheard agendas: qualitative analysis of the psychosocial cues that patients with unexplained symptoms present to general practitioners. The British journal of general practice : the journal of the Royal College of General Practitioners 2004;**54**(500):171-76.
52. Wong MG, Ninomiya T, Liyanage T, et al. Physical component Quality of life (QOL) reflects the impact of time and moderate chronic kidney disease, unlike SF-6D utility and mental component SF-36 QOL: an AUSDiab analysis. Nephrology 2018;**0**(ja).
53. van der Veer SN, Aresi G, Gair R. Incorporating patient-reported symptom assessments into routine care for people with chronic kidney disease. Clinical Kidney Journal 2017;**10**(6):783-87.

54. Sommerville I, Baxter G. Socio-technical systems: From design methods to systems engineering. Interacting with Computers 2010;**23**(1):4-17.
55. ISO. Ergonomics of Human–System Interaction – Part 210: Human-centred Design for Interactive Systems. International Standards Organisation, 2010.
56. Das A, Svanæs D. Human-centred methods in the design of an e-health solution for patients undergoing weight loss treatment. International Journal of Medical Informatics 2013;**82**(11):1075-91.
57. Grove BE, Schougaard LM, Hjollund NH, et al. Self-rated health, quality of life and appetite as predictors of initiation of dialysis and mortality in patients with chronic kidney disease stages 4-5: a prospective cohort study. BMC Res Notes 2018;**11**(1):371.
58. Gibbons CJ, Skevington SM, Group W. Adjusting for cross-cultural differences in computer-adaptive tests of quality of life. Quality of Life Research 2018;**27**(4):1027-39.
59. O'Hare AM. Patient-Centered Care in Renal Medicine: Five Strategies to Meet the Challenge. American Journal of Kidney Diseases 2018;**71**(5):732-36.
60. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ (Clinical research ed) 2009;**339**:b2700.
61. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ (Clinical research ed) 2015;**349**:g7647.
62. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. PLoS medicine 2007;**4**(10):e297.
63. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua 2007;**19**(6):349-57.
64. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2009;**12**(4):419-29.
65. Zbrozek A, Hebert J, Gogates G, et al. Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2013;**16**(4):480-9.
66. The Research-Based Web Design & Usability Guidelines, Enlarged/Expanded edition. Washington: U.S. Government Printing Office, 2006.
67. Szajna B. Empirical Evaluation of the Revised Technology Acceptance Model. Management Science 1996;**42**(1):85-92.
68. Holzner B, Giesinger JM, Pinggera J, et al. The Computer-based Health Evaluation Software (CHES): a software for electronic patient-reported outcome monitoring. BMC medical informatics and decision making 2012;**12**:126.
69. Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: challenges and opportunities. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 2009;**18**(1):99-107.
70. Guyatt GH, Deyo RA, Charlson M, et al. Responsiveness and validity in health status measurement: a clarification. Journal of clinical epidemiology 1989;**42**(5):403-8.
71. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury 2011;**42**(3):236-40.
72. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England) 1986;**1**(8476):307-10.

73. de Vet HC, Terwee CB, Ostelo RW, et al. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health and quality of life outcomes 2006;**4**:54-54.
74. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. Journal of clinical epidemiology 2010;**63**(7):804-5; author reply 06.
75. Turner D, Schunemann HJ, Griffith LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. Journal of clinical epidemiology 2010;**63**(1):28-36.
76. Angst F. The new COSMIN guidelines confront traditional concepts of responsiveness. BMC Medical Research Methodology 2011;**11**(1):152.
77. Falissard B. Minimal clinically meaningful difference: measurement error is integrated. Journal of clinical epidemiology 2004;**57**(6):649-50.
78. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. Journal of clinical epidemiology 2003;**56**(5):395-407.
79. Terwee CB, Roorda LD, Knol DL, et al. Linking measurement error to minimal important change of patient-reported outcomes. Journal of clinical epidemiology 2009;**62**(10):1062-7.
80. Baillie AJ. A new approach in the determination of the minimal clinically meaningful difference. Journal of clinical epidemiology 2004;**57**(6):647-48.
81. Gatchel RJ, Mayer TG, Chou R. What does/should the minimum clinically important difference measure? A reconsideration of its clinical value in evaluating efficacy of lumbar fusion surgery. Clin J Pain 2012;**28**(5):387-97.
82. Johnston BC, Ebrahim S, Carrasco-Labra A, et al. Minimally important difference estimates and methods: a protocol. BMJ Open 2015;**5**(10).
83. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of Life Research 2018;**27**(5):1171-79.
84. Gibbons C, Bower P, Lovell K, et al. Electronic Quality of Life Assessment Using Computer-Adaptive Testing. Journal of medical Internet research 2016;**18**(9):e240.
85. Gibbons CJ. Turning the Page on Pen-and-Paper Questionnaires: Combining Ecological Momentary Assessment and Computer Adaptive Testing to Transform Psychological Assessment in the 21st Century. Frontiers in Psychology 2017;**7**(1933).
86. Morris S, Bass M, Lee M, et al. Advancing the efficiency and efficacy of patient reported outcomes with multivariate computer adaptive testing. J Am Med Inform Assoc 2017;**24**(5):897-902.
87. Loe BS, Stillwell D, Gibbons C. Computerized Adaptive Testing Provides Reliable and Efficient Depression Measurement Using the CES-D Scale. Journal of medical Internet research 2017;**19**(9):e302.
88. INVOLVE. What is public involvement in research? Secondary What is public involvement in research? http://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/.
89. Staniszewska S, Ahmed L. Patient involvement in the evaluation of health care: identifying key issues and considering the way forward. Coronary Health Care 2000;**4**(1):39-47.
90. Edwards C, Staniszewska S. Accessing the user's perspective. Health Soc Care Community 2000;**8**(6):417-24.
91. Mockford C, Staniszewska S, Griffiths F, et al. The impact of patient and public involvement on UK NHS health care: a systematic review. International journal for quality in health care : journal of the International Society for Quality in Health Care / ISQua 2012;**24**(1):28-38.
92. Haywood K, Brett J, Salek S, et al. Patient and public engagement in health-related quality of life and patient-reported outcomes research: what is important and why should we care? Findings from the first ISOQOL patient engagement symposium. Quality of Life Research 2015;**24**(5):1069-76.

93. Staley K. ‘Is it worth doing?’ Measuring the impact of patient and public involvement in research. Research Involvement and Engagement 2015;**1**(1):6.
94. Wilson P, Mathie E, Keenan J, et al. ReseArch with Patient and Public invOlvement: a realisT evaluation: the RAPPORT study. Perth Health Services and Delivery Research, NIHR 2015.
95. Thompson J, Bissell P, Cooper CL, et al. Exploring the Impact of Patient and Public Involvement in a Cancer Research Setting. Qual Health Res 2013;**24**(1):46-54.
96. Wiering B, Boer D, Delnoij D. Patient involvement in the development of patient-reported outcome measures: a scoping review. Health Expect 2016;**20**(1):11-23.
97. Osoba D. Translating the science of patient-reported outcomes assessment into clinical practice. Journal of the National Cancer Institute Monographs 2007(37):5-11.
98. Aaronson N, Elliott T, Greenhalgh J, et al. User’s Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice. International Society for Quality of Life Research 2015.
99. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. Journal of Comparative Effectiveness Research 2016;**5**(5):507-19.
100. Bausewein C, Simon ST, Benalia H, et al. Implementing patient reported outcome measures (PROMs) in palliative care - users' cry for help. Health and quality of life outcomes 2011;**9**:27-27.
101. Gibbons E, Black N, Fallowfield L, et al. Patient-reported outcome measures and the evaluation of services. In: Raine R, Fitzpatrick R, Barratt H, eds. Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. Southampton (UK): NIHR Journals Library, 2016.
102. Santana MJ, Haverman L, Absolom K, et al. Training clinicians in how to use patient-reported outcome measures in routine clinical practice. Quality of Life Research 2015;**24**(7):1707-18.
103. Cruz Rivera S, Kyte DG, Aiyegbusi OL, et al. Assessing the impact of healthcare research: A systematic review of methodological frameworks. PLoS medicine 2017;**14**(8):e1002370.

Appendix

Appendix 17 - Systematic review of frameworks for research impact (page 327)

Appendices

Appendix 1: Systematic review search strategy

S1. PROM Search Strategy

MEDLINE (R) (Ovid) 1946 to 21/12/15

Note: Sets 1 to 6 are based on the Oxford PROM Filter.[1] Set 13 is the sensitive search filter for measurement properties developed by Terwee et al. [2]

1. (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab.

2. quality of life.mp.

3. (health index* or health indices or health profile*).ti,ab.

4. health status.mp.

5. ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating or based or assessed or assessment*)).ti,ab.

6. ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab.

7. ((((patient adj reported adj outcome adj measure*) or patient) adj reported adj outcome*) or capability or capabilities).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

8. 1 or 2 or 3 or 4 or 5 or 6 or 7

9. (Renal replacement therapy or APD or Automated Peritoneal Dialysis or CAPD, Continuous Ambulatory Peritoneal Dialysis or CCPD or Continuous cyclic peritoneal dialysis or dialysis or h*emofiltration or h*emodiafiltration or h*emodialysis or kidney transplant* or predialysis or renal replacement or renal transplant*).mp.

10. (CRF or chronic renal failure or CKF or chronic kidney failure or kidney disease* or renal disease or kidney failure or renal failure or CKD or chronic kidney disease or ESKD or end stage kidney disease or ESKF or end stage kidney failure or ESRF or end stage renal failure or ESRD or end stage renal disease or kidney insufficiency).mp.

11. Renal Insufficiency, Chronic/

12. 9 or 10 or 11

13. ((((((((((((((((((((((((((Instrumentation or method* or Validation Studies or Comparative Study).mp. or psychometrics/ or psychometr*.mp. or clinimetr*.mp. or clinometr*.mp. or outcome assessment health care/ or outcome assessment*.ti,ab. or outcome measure*.mp. or observer variation/ or observer variation*.ti,ab. or Health Status Indicators/ or reproducibility of results/ or reproducib*.ti,ab. or discriminant analysis/ or reliab*.ti,ab. or unreliab*.ti,ab. or valid*.ti,ab. or coefficient of variation.ti,ab. or coefficient*.ti,ab. or homogeneity.ti,ab. or homogeneous.ti,ab. or internal consistency.ti,ab. or cronbach*.ti,ab.) and alpha*.ti,ab.) or item*.ti,ab.) and correlation*.ti,ab.) or selection*.ti,ab. or reduction*.ti,ab. or agreement.mp. or precision.mp. or imprecision.mp. or precise value*.mp. or test-retest.ti,ab. or test.ti,ab.) and retest.ti,ab.) or reliab*.ti,ab.) and test.ti,ab.) or retest.ti,ab. or stability.ti,ab. or interrater.ti,ab. or inter-rater.ti,ab. or intrarater.ti,ab. or intra-rater.ti,ab. or intertester.ti,ab. or inter-tester.ti,ab. or intratester.ti,ab. or intra-tester.ti,ab. or interobserver.ti,ab. or inter-observer.ti,ab. or intraobserver.ti,ab. or intra-observer.ti,ab. or intertechnician.ti,ab. or inter-technician.ti,ab. or intratechnician.ti,ab. or intra-technician.ti,ab. or interexaminer.ti,ab. or inter-examiner.ti,ab. or intraexaminer.ti,ab. or intra-examiner.ti,ab. or interassay.ti,ab. or inter-assay.ti,ab. or intraassay.ti,ab. or intra-assay.ti,ab. or interindividual.ti,ab. or inter-individual.ti,ab. or intraindividual.ti,ab. or intra-individual.ti,ab. or interparticipant.ti,ab. or inter-participant.ti,ab. or intraparticipant.ti,ab. or intra-participant.ti,ab. or kappa*.ti,ab. or kappa's.ti,ab. or repeatab*.mp. or replicab*.mp. or repeated.mp.) and measure*.mp.) or finding*.mp. or result*.mp. or test*.mp. or generaliza*.ti,ab. or generalisa*.ti,ab. or concordance.ti,ab. or intraclass.ti,ab.) and correlation*.ti,ab.) or discriminative.ti,ab. or known group.ti,ab. or factor analysis.ti,ab. or factor analyses.ti,ab. or factor structure.ti,ab. or factor structure.ti,ab. or dimension*.ti,ab. or subscale*.ti,ab. or multitrait.ti,ab.) and scaling.ti,ab. and analysis.ti,ab.) or analyses.ti,ab. or item discriminant.ti,ab. or interscale correlation*.ti,ab. or error.ti,ab. or errors.ti,ab. or individual variability.ti,ab. or interval variability.ti,ab. or rate variability.ti,ab. or variability.ti,ab.) and analysis.ti,ab.) or value*.ti,ab. or uncertainty.ti,ab.) and measurement.ti,ab.) or measuring.ti,ab. or standard error of measurement.ti,ab. or sensitiv*.ti,ab. or responsive*.ti,ab. or limit*.ti,ab.) and detection.ti,ab.) or minimal detectable concentration.ti,ab. or interpretab*.ti,ab. or minimal.ti,ab. or minimally.ti,ab. or clinical.ti,ab. or clinically.ti,ab.) and important.ti,ab.) or significant.ti,ab. or detectable.ti,ab.) and change.ti,ab.) or difference.ti,ab. or small*.ti,ab.) and real.ti,ab.) or detectable.ti,ab.) and change.ti,ab.) or difference.ti,ab. or meaningful change.ti,ab. or ceiling effect.ti,ab. or floor effect.ti,ab. or Item response model.ti,ab. or IRT.ti,ab. or Rasch.ti,ab. or Differential item functioning.ti,ab. or DIF.ti,ab. or computer adaptive testing.ti,ab. or item bank.ti,ab. or cross-cultural equivalence.ti,ab.

14. (PRO integration or Clinical PRO application* or telePRO or automated PRO algorithm* or screening purpose* or PRO questionnaire* or Patient-reported outcome

questionnaire* or Patient-reported symptom* or Patient-centred care or Patient self-report* or Self-report health or Self-rated health or Self-reported measure* of health or Health outcome* or Health communication* or Hospital performance evaluation* or Automated telephone survey system* or paper-based survey* or web-based survey* or web-based PRO platform* or web-based system* or PRO collection* or PRO measure* or PRO intervention* or PRO assessment intervention* or PRO data or PRO assessment* or Routine PRO assessment* or Routine PRO collection or Symptom assessment* or Symptom monitoring or Symptom data or Functional status or Electronic PRO assessment* or Electronic PRO system* or ePRO or

ePRO* or ePRO system* or PRO system* or Generic PRO system* or PRO-based clinical alert system*).mp.

15. 13 or 14

16. 8 and 12 and 15

EMBASE (Ovid) 1974 to 21/12/15

Note: Set 4 is the sensitive search filter for measurement properties developed by Terwee et al. and adapted for EMBASE.[2]

1. (Renal replacement therapy or APD or Automated Peritoneal Dialysis or CAPD, Continuous Ambulatory Peritoneal Dialysis or CCPD or Continuous cyclic peritoneal dialysis or dialysis or h*emofiltration or h*emodiafiltration or h*emodialysis or kidney transplant* or predialysis or renal replacement or renal transplant*).mp.

2. (CRF or chronic renal failure or CKF or chronic kidney failure or kidney disease* or renal disease or kidney failure or renal failure or CKD or chronic kidney disease or ESKD or end stage kidney disease or ESKF or end stage kidney failure or ESRF or end stage renal failure or ESRD or end stage renal disease or kidney insufficiency).mp.

3. (PRO integration or Clinical PRO application* or telePRO or automated PRO algorithm* or screening purpose* or PRO questionnaire* or Patient-reported outcome questionnaire* or Patient-reported symptom* or Patient-centred care or Patient self-report* or Self-report health or Self-rated health or Self-reported measure* of health or Health outcome* or Health communication* or Hospital performance evaluation* or Automated telephone survey system* or paper-based survey* or web-based survey* or web-based PRO platform* or web-based system* or PRO collection* or PRO measure* or PRO intervention* or PRO assessment intervention* or PRO data or PRO assessment* or Routine PRO assessment* or Routine PRO collection or Symptom assessment* or Symptom monitoring or Symptom data or Functional status or Electronic PRO assessment* or Electronic PRO system* or ePRO or

ePRO* or ePRO system* or PRO system* or Generic PRO system* or PRO-based clinical alert system*).mp.

4. ((exp questionnaire/ or exp named inventories/ or questionnaires/) and rating scales/) or exp psychometry/ or exp outcome assessment/ or exp pain assessment/ or exp disability/ or exp validity/ or exp reliability/

5. *patient/ or *outcome assessment/ or *questionnaire/ or *"quality of life"/ or *self report/

6. 3 or 4

7. 1 or 2

8. 5 and 6 and 7

PsycINFO (Ovid) 1967 to 21/12/15

Note: Sets 1 to 6 are based on the Oxford PROM Filter.[1] Set 11 is the sensitive search filter for measurement properties developed by Terwee et al.[2]

1. (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab.

2. quality of life.mp.

3. (health index* or health indices or health profile*).ti,ab.

4. health status.mp.

5. ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating or based or assessed or assessment*)).ti,ab.

6. ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab.

7. ((((patient adj reported adj outcome adj measure*) or patient) adj reported adj outcome*) or capability or capabilities).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

8. 1 or 2 or 3 or 4 or 5 or 6 or 7

9. (Renal replacement therapy or APD or Automated Peritoneal Dialysis or CAPD, Continuous Ambulatory Peritoneal Dialysis or CCPD or Continuous cyclic peritoneal dialysis or dialysis or h*emofiltration or h*emodiafiltration or h*emodialysis or kidney transplant* or predialysis or renal replacement or renal transplant*).mp.

10. (CRF or chronic renal failure or CKF or chronic kidney failure or kidney disease* or renal disease or kidney failure or renal failure or CKD or chronic kidney disease or ESKD or end stage kidney disease or ESKF or end stage kidney failure or ESRF or end stage renal failure or ESRD or end stage renal disease or kidney insufficiency).mp.

11. ((((((((((((((((((((((((((Instrumentation or method* or Validation Studies or Comparative Study).mp. or psychometrics/ or psychometr*.mp. or clinimetr*.mp. or clinometr*.mp. or outcome assessment health care/ or outcome assessment*.ti,ab. or outcome measure*.mp. or observer variation/ or observer variation*.ti,ab. or Health Status Indicators/ or reproducibility of results/ or reproducib*.ti,ab. or

discriminant analysis/ or reliab*.ti,ab. or unreliab*.ti,ab. or valid*.ti,ab. or coefficient of variation.ti,ab. or coefficient*.ti,ab. or homogeneity.ti,ab. or homogeneous.ti,ab. or internal consistency.ti,ab. or cronbach*.ti,ab.) and alpha*.ti,ab.) or item*.ti,ab.) and correlation*.ti,ab.) or selection*.ti,ab. or reduction*.ti,ab. or agreement.mp. or precision.mp. or imprecision.mp. or precise value*.mp. or test-retest.ti,ab. or test.ti,ab.) and retest.ti,ab.) or reliab*.ti,ab.) and test.ti,ab.) or retest.ti,ab. or stability.ti,ab. or interrater.ti,ab. or inter-rater.ti,ab. or intrarater.ti,ab. or intra-rater.ti,ab. or intertester.ti,ab. or inter-tester.ti,ab. or intratester.ti,ab. or intra-tester.ti,ab. or interobserver.ti,ab. or inter-observer.ti,ab. or intraobserver.ti,ab. or intra-observer.ti,ab. or intertechnician.ti,ab. or inter-technician.ti,ab. or intratechnician.ti,ab. or intra-technician.ti,ab. or interexaminer.ti,ab. or inter-examiner.ti,ab. or intraexaminer.ti,ab. or intra-examiner.ti,ab. or interassay.ti,ab. or inter-assay.ti,ab. or intraassay.ti,ab. or intra-assay.ti,ab. or interindividual.ti,ab. or inter-individual.ti,ab. or intraindividual.ti,ab. or intra-individual.ti,ab. or interparticipant.ti,ab. or inter-participant.ti,ab. or intraparticipant.ti,ab. or intra-participant.ti,ab. or kappa*.ti,ab. or kappa's.ti,ab. or repeatab*.mp. or replicab*.mp. or repeated.mp.) and measure*.mp.) or finding*.mp. or result*.mp. or test*.mp. or generaliza*.ti,ab. or generalisa*.ti,ab. or concordance.ti,ab. or intraclass.ti,ab.) and correlation*.ti,ab.) or discriminative.ti,ab. or known group.ti,ab. or factor analysis.ti,ab. or factor analyses.ti,ab. or factor structure.ti,ab. or factor structure.ti,ab. or dimension*.ti,ab. or subscale*.ti,ab. or multitrait.ti,ab.) and scaling.ti,ab. and analysis.ti,ab.) or analyses.ti,ab. or item discriminant.ti,ab. or interscale correlation*.ti,ab. or error.ti,ab. or errors.ti,ab. or individual variability.ti,ab. or interval variability.ti,ab. or rate variability.ti,ab. or variability.ti,ab.) and analysis.ti,ab.) or value*.ti,ab. or uncertainty.ti,ab.) and measurement.ti,ab.) or measuring.ti,ab. or standard error of measurement.ti,ab. or sensitiv*.ti,ab. or responsive*.ti,ab. or limit*.ti,ab.) and detection.ti,ab.) or minimal detectable concentration.ti,ab. or interpretab*.ti,ab. or minimal.ti,ab. or minimally.ti,ab. or clinical.ti,ab. or clinically.ti,ab.) and important.ti,ab.) or significant.ti,ab. or detectable.ti,ab.) and change.ti,ab.) or difference.ti,ab. or small*.ti,ab.) and real.ti,ab.) or detectable.ti,ab.) and change.ti,ab.) or difference.ti,ab. or meaningful change.ti,ab. or ceiling effect.ti,ab. or floor effect.ti,ab. or Item response model.ti,ab. or IRT.ti,ab.

or Rasch.ti,ab. or Differential item functioning.ti,ab. or DIF.ti,ab. or computer adaptive testing.ti,ab. or item bank.ti,ab. or cross-cultural equivalence.ti,ab.

12. (PRO integration or Clinical PRO application* or telePRO or automated PRO algorithm* or screening purpose* or PRO questionnaire* or Patient-reported outcome questionnaire* or Patient-reported symptom* or Patient-centred care or Patient self-report* or Self-report health or Self-rated health or Self-reported measure* of health or Health outcome* or Health communication* or Hospital performance evaluation* or Automated telephone survey system* or paper-based survey* or web-based survey* or web-based PRO platform* or web-based system* or PRO collection* or PRO measure* or PRO intervention* or PRO assessment intervention* or PRO data or PRO assessment* or Routine PRO assessment* or Routine PRO collection or Symptom assessment* or Symptom monitoring or Symptom data or Functional status or Electronic PRO assessment* or Electronic PRO system* or ePRO or ePRO* or ePRO system* or PRO system* or Generic PRO system* or PRO-based clinical alert system*).mp.

13. 11 or 12

14. *Organ Transplantation/ or *Hemodialysis/ or *Kidneys/ or *Kidney Diseases/ or Renal Insufficiency, Chronic.mp.

15. 9 or 10 or 14

16. 8 and 13 and 15

CINAHL (via EBSCO host) inception to 21/12/15

Note: S1 is based on the Oxford PROM Filter.[1] S3 is the sensitive search filter for measurement properties developed by Terwee et al. [2]

S1. ((HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab.) OR quality of life.mp. OR ((health index* or health indices or health profile*).ti,ab.) OR health status.mp. OR (((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating or based or assessed or assessment*)).ti,ab.) OR (((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab.) OR (((((patient adj reported adj outcome adj measure*) or patient) adj reported adj outcome*) or capability or capabilities).mp)

S2. (MH "Kidney Failure, Chronic/TH/TM/TD/SU/SS/RF/RH/PF/PR/PC/PP/PA/NU/MO/ME/ET/EI/EP/ED/EC/DT/DI/CO/CL") OR "((Renal replacement therapy or APD or Automated Peritoneal

Dialysis or CAPD, Continuous Ambulatory Peritoneal Dialysis or CCPD or Continuous cyclic peritoneal dialysis or dialysis or h*emofiltration or h*emodiafiltration or h*emodialysis or kidney transplant* or predialysis or renal replacement or renal transplant*).mp.) OR ((CRF or chronic renal failure or CKF or chronic kidney failure or kidney disease* or renal disease or kidney failure or renal failure or CKD or chronic kidney disease or ESKD or end stage kidney disease or ESKF or end stage kidney failure or ESRF or end stage renal failure or ESRD or end stage renal disease or kidney insufficiency).mp.) OR (*Organ Transplantation/ or *Hemodialysis/ or *Kidneys/ or *Kidney Diseases/ or Renal Insufficiency, Chronic.mp.)"

S3. ""(((((((((((((((((((((((((((Instrumentation or method* or Validation Studies or Comparative Study).mp. or psychometrics/ or psychometr*.mp. or clinimetr*.mp. or clinometr*.mp. or outcome assessment health care/ or outcome assessment*.ti,ab. or outcome measure*.mp. or observer variation/ or observer variation*.ti,ab. or Health Status Indicators/ or reproducibility of results/ or reproducib*.ti,ab. or discriminant analysis/ or reliab*.ti,ab. or unreliab*.ti,ab. or valid*.ti,ab. or coefficient of variation.ti,ab. or coefficient*.ti,ab. or homogeneity.ti,ab. or homogeneous.ti,ab. or internal consistency.ti,ab. or cronbach*.ti,ab.) and alpha*.ti,ab.) or item*.ti,ab.) and correlation*.ti,ab.) or selection*.ti,ab. or reduction*.ti,ab. or agreement.mp. or precision.mp. or imprecision.mp. or precise value*.mp. or test-retest.ti,ab. or test.ti,ab.) and retest.ti,ab.) or reliab*.ti,ab.) and test.ti,ab.) or retest.ti,ab. or stability.ti,ab. or interrater.ti,ab. or inter-rater.ti,ab. or intrarater.ti,ab. or intra-rater.ti,ab. or intertester.ti,ab. or inter-tester.ti,ab. or intratester.ti,ab. or intra-tester.ti,ab. or interobserver.ti,ab. or inter-observer.ti,ab. or intraobserver.ti,ab. or intra-

observer.ti,ab. or intertechnician.ti,ab. or inter-technician.ti,ab. or intratechnician.ti,ab. or intra-technician.ti,ab. or interexaminer.ti,ab. or inter-examiner.ti,ab. or intraexaminer.ti,ab. or intra-examiner.ti,ab. or interassay.ti,ab. or inter-assay.ti,ab. or intraassay.ti,ab. or intra-assay.ti,ab. or interindividual.ti,ab. or inter-individual.ti,ab. or intraindividual.ti,ab. or intra-individual.ti,ab. or interparticipant.ti,ab. or inter-participant.ti,ab. or intraparticipant.ti,ab. or intra-participant.ti,ab. or kappa*.ti,ab. or kappa's.ti,ab. or repeatab*.mp. or replicab*.mp. or repeated.mp.) and measure*.mp.) or finding*.mp. or result*.mp. or test*.mp. or generaliza*.ti,ab. or generalisa*.ti,ab. or concordance.ti,ab. or intraclass.ti,ab.) and correlation*.ti,ab.) or discriminative.ti,ab. or known group.ti,ab. or factor analysis.ti,ab. or factor analyses.ti,ab. or factor structure.ti,ab. or factor structure.ti,ab. or dimension*.ti,ab. or subscale*.ti,ab. or multitrait.ti,ab.) and scaling.ti,ab. and analysis.ti,ab.) or analyses.ti,ab. or item discriminant.ti,ab. or interscale correlation*.ti,ab. or error.ti,ab. or errors.ti,ab. or individual variability.ti,ab. or interval variability.ti,ab. or rate variability.ti,ab. or variability.ti,ab.) and analysis.ti,ab.) or value*.ti,ab. or uncertainty.ti,ab.) and measurement.ti,ab.) or measuring.ti,ab. or standard error of measurement.ti,ab. or sensitiv*.ti,ab. or

responsive*.ti,ab. or limit*.ti,ab.) and detection.ti,ab.) or minimal detectable concentration.ti,ab. or interpretab*.ti,ab. or minimal.ti,ab. or minimally.ti,ab. or clinical.ti,ab. or clinically.ti,ab.) and important.ti,ab.) or significant.ti,ab. or detectable.ti,ab.) and change.ti,ab.) or difference.ti,ab. or small*.ti,ab.) and real.ti,ab.) or detectable.ti,ab.) and change.ti,ab.) or difference.ti,ab. or meaningful change.ti,ab. or ceiling effect.ti,ab. or floor effect.ti,ab. or Item response model.ti,ab. or IRT.ti,ab. or Rasch.ti,ab. or Differential item functioning.ti,ab. or DIF.ti,ab. or computer adaptive testing.ti,ab. or item bank.ti,ab. or cross-cultural equivalence.ti,ab.)

S4. ((PRO integration or Clinical PRO application* or telePRO or automated PRO algorithm* or screening purpose* or PRO questionnaire* or Patient-reported outcome questionnaire* or Patient-reported symptom* or Patient-centred care or Patient self-report* or Self-report health or Self-rated health or Self-reported measure* of health or Health outcome* or Health communication* or Hospital performance evaluation* or Automated telephone survey system* or paper-based survey* or web-based survey* or web-based PRO platform* or web-based system* or PRO collection* or PRO measure* or PRO intervention* or PRO assessment intervention* or PRO data or PRO assessment* or Routine PRO assessment* or Routine PRO collection or Symptom assessment* or Symptom monitoring or Symptom data or Functional status or Electronic PRO assessment* or Electronic PRO system* or ePRO or ePRO* or ePRO system* or PRO system* or Generic PRO system* or PRO-based clinical alert system*).mp.)""

S5. S3 or S4

S6. S1 and S2 and S5

References

(1) PROM Group, A.M., Carolina Casañas i Comabella, Monica Hadi, Elizabeth Gibbons, Ray Fitzpatrick, Nia Roberts, PROM GROUP CONSTRUCT & INSTRUMENT TYPE FILTERS. 2010.

(2) Terwee, C.B., et al., Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res, 2009. 18(8): p. 1115-23.

Appendix 2: Summary of study results (systematic review)

S1 Table. Summary of study results

Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Agarwal 2010 (Unnamed) [1]	Cronbach's alpha, 4 factors: > 0.85	ICC: 0.75 - 0.85 (n = 41, 8 weeks)	Regression, NP score vs MHC score(KDQOL) r^2: 0.41 (P<0.001)	CFA (Varimax rotation) KMO: 0.835. 73% of the total variance explained by factors.	Literature review and patient interviews	_
Chinese Dialysis Quality of Life Scale (CDQOL) Suet-Ching 2001 [2]	Cronbach's alpha (overall): 0.85	Pearson r: 0.92, p<0.02 (n = 10, 2 weeks)	CDQOL vs WHO-BREF, Pearson r : +0.79, p = 0.000 Item-total correlations: 0.22 - 0.55	_	CVI: 0.94 Face validity with patients	_
CHOICE Heath Experience Questionnaire (CHEQ) Aiyasanon et al. 2009 [3]	Cronbach's alpha: >0.70 for all domains (except 'social functioning', 0.66 and 'quality of life', 0.58).	_	Significant differences between patients on PD and HD in 10 domains with patients on HD scoring higher. Patients with DM scored significantly lower than those without DM in 5 domains.	_	_	_
Wu et al. 2001 [4]	Cronbach's alpha: >0.70 for all domains (except quality of life 0.68).	_	Item-scale correlations at least two SEs greater than item to other scales. Patients with ESRD and other co-morbidities had significantly lower physical functioning scores (p < 0.001) than those without these conditions.	_	Literature review and patient focus groups	_
CKD Symptom Burden Index (CKD-SBI) Almutary et al. 2015 [5]	Cronbach's alpha: 0.91 - 0.93	_	Pearson's coefficient Inter-correlations between domains,0.82 - 0.97 Pearson's correlations between all symptom domains and each domain. r = -0.33 (prevalence vs burden of KD) to r = -0.78 (distress vs symptom scale).(p<0.01)	_	CVI overall: 0.98 Item range: 0.88 - 1.00	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Dialysis Symptom Index (DSI) Önsoz and Yesilbalkan 2013 [6]	Cronbach's alpha for symptoms: 0.82 - 0.84	Kappa statistics: 0.10 - 0.89 (n = 120, 1 week) 23 items kappa values ≥ 0.60 Diarrhoea and difficulty concentrating had values < 0.20	_	_	Expert rating of items, CVI: 0.82	_
Weisbord et al. 2004 [7]	_	Kappa statistics: 0.06 - 0.90 (n = 20, 4 -7 days) 10 items demonstrated kappa values ≥0.60.	_	_	Expert review and focus groups with patients and providers.	_
Ferrans and Powers Quality of Life Index 3.0 (Dialysis) Dehesh et al. 2014 [8]	Cronbach's alpha: 0.64 - 0.88	_	Spearman's correlation between items and hypothesised domains: 0.75 - 0.94 (scaling success 100%)	Factor analysis (Quartimax rotation): Variance explained by the four factors was 89%.	_	_
Ferrans and Powers 1985 [9]	Cronbach's alpha: 0.90	Test-retest: 0.81 (n = 20, 4 weeks)	Correlation between scores from the QLI and the life satisfaction question for dialysis patients = 0.65.	_	Literature review and patient consultation	_
Halabi 2006 [10]	Cronbach's alpha: 0.79 - 0.86	_	_	_	_	_
Korkut 2007 [11]	Cronbach's alpha: 0.92	Test-retest: 0.65 (n = 31, 4 weeks)	Correlations between QLI and 6 SF-36 domains (r = 0.36 - 0.47, p < 0.05) Correlations between QLI score and the Beck's score (r = -0.566, p<0.01)	_	_	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Kidney Disease Quality of Life(KDQOL Dialysis version) Hays 1994 [12]	Cronbach's alpha: all >0.75 (except quality of social interaction 0.68)	_	Number of medications currently taken correlated significantly with 9 domains with physical functioning being the highest (r = -0.23, p <0.01)	CFA (oblique Promax rotation) Guttman's weakest lower bound was consistent for 3 to 4 factors. 4 factors were indicated by parallel analysis and the scree test.	Literature review and patient focus groups	_
KDQOL (Modified) Rao et al 2000 [13]	Cronbach's alpha: 0.66 - 0.92	_	Correlation of pain domain with corresponding SF36 pain domain, r = 0.70, p<0.001 Sleep and SF36 energy/fatigue domains, r = 0.57, p<0.001	_	_	_
KDQOL-36 questionnaire (KDQOL-36) Chao et al. 2016 [14]	Cronbach's alpha (excluding item 28): 0.76 - 0.92	_	Correlation between item to total and item to domain: 0.21 - 0.67 and 0.43 - 0.92 respectively (p< .001). Pearson's Coefficient correlation of the KDQOL-36 (excluding Item 28) to the question "How satisfied are you with your health?" of the WHOQOL-BREF): 0.28 to 0.42 (p < 0.001)	After improving the model fit (deleting item 28 (dialysis) and item 35 (sex life), the modified model 2 was presented as follows: $\chi 2 = 1390.903$, $\chi 2/df = 2.711$, p < 0.001, GFI = 0.833, adjusted GFI = 0.806, RMSEA = 0.063.	The equivalence, clarity, and readability of the translated KDQOL-36™ CVI: 91%, 93%, and 95%, respectively.	_
Chow and Tam 2014 [15]	Cronbach's alpha (kidney disease domains): 0.65 - 0.83. Cronbach's PCS: 0.32 and MCS: 0.53.	Test-retest (ICC) all subscales: >0.98. (n = 20, 10 - 14 days)	Spearman Rho correlations depression (HADS) and all domains, r: −0.25 (p < 0.001) to −0.50 (p < 0.001). Independent t-test, to compare QoL between the patient groups, p <0.001 for all domains.	_	_	_
Mateti 2015 [16]	Cronbach's alpha: 0.72 - 0.77	Test-retest (ICC): 0.83 - 0.97 (n = 45, 7 days)	EQ-VAS in kidney disease targeted domains correlated significantly with symptoms (0.393), effects of kidney disease (0.320) and burden of kidney disease (0.264).	_	_	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Ricardo et al. 2013 [17]	Cronbach's alpha: 0.80 - 0.87	_	MCS correlated with BDI for all groups: -0.56 to -0.61, p<0.001. The PSF score correlated significantly with the symptoms domain, r= −0.70 to −0.77.	_	_	_
Tao et al. 2014 [18]	Cronbach's alpha: 0.69 - 0.78	Test-retest (ICC): 0.70 - 0.86(n = 28, 10 - 14 days)	Spearman's correlations between all KDQOL-36 scores and the BDI score, rho: -0.395 to -0.654, p <0.01)	_	Item-level content validity index (I-CVI) and scale-level CVI (S-CVI): 1.0	_
Thaweethamcharoen et al. 2013 [19]	Cronbach's alpha, kidney disease domains: HD = 0.79 - 0.83 PD = 0.71 - 0.78	Test-retest (ICCs): for all scales > 0.95 (n = 30)	For HD, Spearman's correlations of disease specific scores with PCS, MCS and utility scores (EQ5D): all +ve and significant. For PD, Spearman's correlations of symptoms domain: all +ve and significant (p<0.001)	_	_	_
Yang 2013 [20]	Cronbach's alpha, kidney disease domains:0.82- 0.91 Cronbach's alpha:0.74 (PCS) and 0.70 (MCS)	_	Pearson's Coefficient, item-to-scale correlation: 0.76 - 0.90	CFA* (kidney disease items), a 3-factor model reflected the grouping of items to form the 3 kidney scales. Good fit to data. (CFI 0.934, RMSEA 0.085).	_	_
KDQOL-Short Form questionnaire (KDQOL-SF) Abd ElHafeez et al. 2012 [21]	Cronbach's alpha, kidney-specific domains: >0.70 (except 'work status' 'cognitive function', 'quality of social interaction' 0.28, 0.60. 0.23, respectively) Cronbach's alpha, generic domains: >0.70	Test-retest (ICC): 0.79 - 0.95 (n = 50, 1 week)	All the disease specific domains correlated significantly with PCS, r: 0.24 - 0.68 (p<0.05) All the disease specific domains, except 'sexual function', correlated significantly with MCS, r:0.24 - 0.55 (p<0.05)	EFA (Varimax rotation) KMO: 0.73. Bartlett's test of sphericity: X2 = 1.83, p < 0.001. PCA was done. Initial eigenvalues showed that 39 items explained 70.9% of the variance in 10 components. Factor loading ranged from 0.50 to 0.94.	_	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Barotfi et al. 2006 [22]	Cronbach's alpha, HD, all generic domains: > 0.70 Disease-specific: 0.54 - 0.83 Cronbach's alpha, Tx, all generic domains: > 0.70 Disease-specific: 0.59 - 0.95	Test–retest (Pearson's coefficients) Disease-specific domains: 0.40 – 0.88 Generic domains: 0.56 – 0.81 (n = 63)	Differences in mean scores between Tx and HD for all generic domains and for 7 of the 11 disease-specific domains. >10 points for most domains that show a difference (p<0.001). QoL scores significantly different between Tx and HD for generic domains.	–	–	–
Bataclan and Dial 2009 [23]	Cronbach's alpha, Disease-specific domains: 0.60 - 0.78 Generic domains: 0.61 - 0.80	Test-retest (ICC) : Disease-specific domains: 0.58 - 0.97 Generic: 0.65 - 0.98 (n = 30, 10 - 14 days)	All the items in generic domains had significant correlation with overall health rating: 0.61 - 0.80, (p<0.05) except for role – emotional (p = 0.19). Disease-specific domains correlated with the overall health rating: 0.18 – 0.39, p <0.05	–	–	–
Boini et al. 2007 [24]	Cronbach's alpha, Disease-specific domains: >0.7 for all except 'work status' (0.38) and 'quality of social interactions' (0.48). Generic domains: all >0.75 except 'Social functioning' (0.64	Test-retest correlation coefficient, specific domains: >0.70 for all except 'social interactions' (0.51), 'social support' (0.42) Generic domains: >0.70 for all except 'physical functioning' (0.56) and 'role physical' (0.67) (7 days)	Item-dimension correlations, for all the specific domains: 0.03 - 0.83 Correlation between burden of kidney disease (specific) and general health (generic), r = 0.72	–	–	–

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Bouidida et al. 2014 [25]	Cronbach's alpha, Kidney-specific domains: 0.38 - 0.90. Note that sexual function only had 40% response rate. Generic domains: 0.70 - 0.86	Test-retest (ICC), Kidney specific domains: 0.78 - 0.90 Generic domains: 0.67 - 0.90 (n = 20, 10 - 14 days)	Correlations of items within domains were higher than that of items outside domains in 87% of cases. Men scored significantly higher than women on 5 subscales: on symptom/problem, cognitive function, physical functioning, social functioning, and pain.	_	_	_
Cheung et al. 2012 [26]	Cronbach's alpha for specific ranged from 0.32 (work status) to 1.00 (dialysis staff encouragement and sexual function) Cronbach's for RAND 36 domains ranged: 0.78 - 0.98	_	Spearman's rank coefficient (rho), KPS scores most strongly correlated with 'effects of kidney disease' (rho = 0.56, p<0.001), 'physical functioning' (rho = 0.87, p<0.001), and 'energy/fatigue' (rho = 0.62, p<0.001). eGFR correlated with 'burden of kidney disease' (rho = 0.32, p<0.01), 'effects of kidney disease' (rho = 0.29, p<0.01), 'symptoms/problems' (rho = 0.22, p<0.01).	_	_	Starting RRT was associated with improvement in the following domains: symptoms/problems (p <0.001), cognitive function (p =0.047), and sleep (p =0.044).
Duarte et al. 2005 [27]	_	Intra-observer reliability (Cronbach's alpha): all >0.70 except 'patient satisfaction' (0.65). (ICC): 0.49 - 0.93, p<0.001 Inter-observer reliability (Cronbach's alpha): all >0.70 except 'patient satisfaction' (0.50). (ICC): 0.33 - 0.99, p<0.001	Pearson's correlation coefficients, physical functioning and physical capacity (NHP): 0.72, p<0.05. Energy/fatigue and Fatigue (KDQ): 0.68, p<0.05.	_	_	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Fardinmehr et al. 2012 [28]	Cronbach's alpha: 0.85	_	Pearson's coefficients correlations, Overall health score correlated significantly with Kidney specific domains: symptoms/problems (0.47), effect of KD (0.67), burden of KD (0.43).	_	_	_
Green et al. 2001 [29]	Cronbach's alpha, Specific (before item deletion): 0.35 - 0.92 (after deletion): 0.64 - 0.81 Generic domains: 0.73 - 0.92	Test and retest (ICC), Specific domains: 0.42 - 0.95 Generic domains: 0.60 - 0.83 (n = 60, 10 - 14 days)	11 out of 42 items correlated with their intended scale within 2 standard errors of their correlation with at least one other scale.	_	_	_
Hays et al. 1995 [30]	Cronbach's alpha: all were >0.80 except quality of social interaction (0.61) and cognitive function (0.68)	_	_	_	Literature review and focus groups with patients	_
Joshi et al., 2010 [31]	Cronbach's alpha: Overall: >0.7 All domains: >0.7 except social function (0.66)	_	Significant correlations between overall health score and kidney-disease targeted domains: 0.21 - 0.37	EFA (Varimax rotation) Total variance explained by all 8 subscales was 68.35%. Physical function explained the highest percentage of variance at 31.39%	_	_
Klersy et al. 2007 [32]	Cronbach's alpha: Kidney specific domains: Only 3 domains were <0.70. Generic domains: Only 2 domains were <0.70	Test-retest, Pearson's correlation coefficient and ICC ranged from 80% to 100% for all domains except staff encouragement and work status (both 60%). (n = 10, within 4 weeks)	Item internal consistency was 74.5% for kidney specific items and 87.5% for generic items.	Factor analysis (PCA, oblique rotation, scree plot) Eight scales related to the burden of the disease, were loaded on the first axis while scales social support, staff encouragement and patient satisfaction, were loaded on the second axis.	_	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Kontodimopoulos & Niakas 2005 [33]	Cronbach's alpha >0.7 for all domains except social interaction, (0.59) and social support, (0.68).		Item-scale correlations (Item-internal consistency) for both generic and specific domains: 0.31 - 0.94. Scaling success (>0.40) in all except one generic item and 3 specific items. T-test and ANOVA showed statistically significant differences due to age, gender, education, co-morbidities and hospitalizations ($p<0.05$).			
Kontodimopoulos & Niakas 2007 [34]	Cronbach's alpha: >0.75 for all generic domains. Specific domains: >0.75 for all except social interaction (0.59), Social support (0.67)	Item-internal consistency between items and their hypothesised scales for both generic and specific sections: 0.32 - 0.93.	Pearson's correlations between generic and specific domains: 0.27 - 0.67 ($p<0.01$). T-test and ANOVA showed statistically significant differences due to age, gender, education, co-morbidities and hospitalizations.			
Korevaar et al. 2002 [35]	Cronbach's alpha, specific domains: >0.7 for all except 'work status' and 'quality of social interactions' (both 0.39) and 'social support' (0.67).	–	EuroQoL overall health rating (VAS-score) was significantly correlated with all the generic and specific domains of KDQOL: 0.31 - 0.76, except 'patient satisfaction'. Residual GFR correlated significantly with 5 out the 8 specific domains, 0.15 - 0.21	–	–	Significant correlations between changes in HRQOL scores and residual GFR in 10 domains within a 9 month period.
Malindretos 2010 [36]	Cronbach's alpha: >0.90 for all domains.	Test-retest (ICC): 0.88 - 0.99 (n = 50, within 4 weeks)	All items showed significant correlation with their hypothesized scale and this correlation was stronger than the correlation with other domains.	–	–	–
Molsted et al. 2005 [37]	Cronbach's alpha: > 0.70 for 8 disease-specific domains. Cronbach's alpha: Generic domains: 0.77 - 0.93	–	–	–	–	–

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Moreira et al. 2009 [38]	Cronbach's alpha, Specific domains: 0.1 -1.0 Generic domains: 0.7 - 0.9	–	Ordinal logistic regression, There were significant correlations between men and 'pain'. Men had higher scores in this dimension compared to women.	–	–	–
Pakpour et al. 2011a [39]	Cronbach's alpha, Specific domains: 0.71 - 0.92 Generic domains: 0.73 - 0.93	Test–retest (ICC), Specific domains: 0.79 - 0.92. Generic domains: 0.77 - 0.91 (n = 20, within 10 days)	Pearson's Correlations, Overall health rating correlated significantly with symptoms, effects, burden of kidney disease, work status, quality of social interaction, sexual function and sleep (0.11 - 0.439, P < 0.01).	PCA on the 43 kidney targeted items. KMO: 0.71, Bartlett's test of sphericity: x2-test (946) = 4.570, P < 0.001. EFA on the 43 items: initial eigenvalues showed all items explained 79.81% of the variance in 11 components.	–	–
Park et al. 2007 [40]	Cronbach's alpha, Disease specific domains: 0.45 - 0.93 Generic domains : 0.65 - 0.92	Test-retest, (ICC): >0.75 for all KDQOL-SF domains, except social functioning (0.69) (n = 30, 1 week)	Pearson's correlation coefficients, overall health rating correlated significantly with symptoms (0.35), effects of KD (0.35), burden of KD (0.29), cognitive function (0.22), and quality of social interaction (0.25).	–	–	–
Perneger 2003 [41]	Cronbach's alpha, Generic domains: 0.74 - 0.90 Disease-specific: 8 domains were >0.70.	–	Six kidney disease specific scales correlated significantly with PCS: 0.29 (work status) to 0.56 (symptoms/problems) All the disease specific scales correlated significantly with MCS: 0.27 (work status) to 0.70 (cognitive function).	–	–	–
Vasilieva 2007 [42]	Cronbach's alpha: 0.72 to 0.86 (for 7 of the kidney specific domains)	–	–	–	–	–

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Yildirim et al. 2007 [43]	Cronbach's alpha: 0.84	Test-retest (Cronbach's alpha): 0.91.	Pearson's coefficient, Significant correlations between the scores of similar domains of 15D and KDQOL-SF. Physical functioning and mobility (r = -0.81), emotional well-being and mental function (r = -0.78), sexual function and sexual activity (r = -0.78), patient satisfaction and distress (r = -0.80).	_	_	_
Kidney Disease Questionnaire (KDQ) Alvarez-Ude et al. 1997 [44]	Cronbach's alpha: All ≥0.7 except relationship with others (0.68) and frustration (0.67) Overall: 0.93	Test-retest (ICC): All ≥0.7 except frustration (0.62) (n = 34, 2 weeks)	Spearman's rank correlation coefficient, KDQ correlated with the domains and overall score of the SIP, r: -0.45 to -0.83. Correlation with the EuroQoL tariff values: 0.39 - 0.65.	_	_	_
Laupacis et al. 1992 [45]	_	Test-retest (ICC) placebo-treated patients: All domains >0.80 (8 weeks)	The physical and fatigue domains correlated significantly with all the other measures (Physical SIP, Psychosocial SIP, Time trade-off, Stress test)	Factor analysis was done to assist with item selection but details not reported.	Literature review, clinician and patient interviews	Significant improvement over 6 months with erythropoietin treatment in 4 of 5 dimensions: physical, fatigue, relationship with others and depression.
ESRD Symptom Checklist–Transplantation Module (ESRD-SCL) Franke et al. 1999 [46]	Cronbach's alpha: 0.76 - 0.85	Reproducibility (ICC) between the first and second measurement for 88 stable patients after 12 months were moderate in 5 and modest in 1 of the 6 scales.	All the domains of the ESRD-SCL correlated negatively and significantly with all the domains of SF-36 except the 'increased growth of gum' and hair' dimension.	Factor analysis (Varimax rotation) was used to create the final questionnaire. Six factors were defined which contained 43 items.	Literature review, clinician and patient interviews	_

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Ortega et al. 2007 [47]	Cronbach's alpha: 0.92 (overall) Range: 0.62 - 0.86	Test-retest (ICC): all domains >0.7. Range: 0.73 - 0.81	Correlations between the ESRD-SCL and MCS: moderate for psycho-neurological problems, Limited Physical Capacity, and Limited Cognitive Capacity and weak for the rest.	Factor analysis (Varimax rotation) Scree test showed that the 6 factor model explained 56.07% variance.	–	The ESRD-SCL scores decreased throughout follow up while the scores for SF-36 and EQ5D increased, indicating better HRQoL.
Stavem and Ganss 2006 [48]	Cronbach's alpha: 0.72 - 0.81 (overall: 0.94)	Test-retest (ICC): 0.87 - 0.95 (n = 48, over 2 weeks)	Pearson's coefficients, Overall ESRD-SCL correlated negatively and significantly with all SF-36 domains. 'Limited physical capacity' correlated negatively with all SF-36 domains (-0.52 to -0.74, p<0.01).	–	–	–
Gastrointestinal Quality of Life Index (GIQLI) Kleinman et al. 2006 [49]	Cronbach's alpha, GIQLI for domains: 0.71 - 0.90 Overall: 0.93	–	Pearson's correlations, GIQLI total score and EQ-5D: r = 0.63. GSRS scores were significantly different by clinical severity ratings from 'none' to 'severe and very severe'.	–	–	–
Gastrointestinal Symptom Rating Scale (GSRS) Kleinman et al. 2006 [49]	Cronbach's alpha, GSRS: 0.71 - 0.91 (except 'abdominal pain' 0.58)	–	Pearson's correlations, 'abdominal pain' and EQ-5D: r = -0.54 GSRS scores were significantly different by clinical severity ratings from 'none' to 'severe and very severe'.	–	–	–
Kidney Transplant Questionnaire (KTQ) Chisholm-Burns et al. 2011 [50]	Cronbach's alpha: 0.62 - 0.90.	–	Pearson's correlations, PCS (SF-12) with uncertainty/fear (0.2), physical, fatigue (0.4). MCS with uncertainty/fear (0.3), emotional, fatigue (0.4), Physical (0.26), appearance (0.28).	–	–	–

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Laupacis et al. 1993 [51]	Cronbach's alpha: 0.61 - 0.94	ICC: 0.70 - 0.87 (n = 26 stable Tx patients, assessed 6 and 12 months post-transplantation)	The physical domains of the SIP correlated most closely with the physical symptoms and fatigue domains of the KTQ ($p < 0.001$).	A combination of factor analysis and clinical judgement was used to create the final questionnaire. Details were not reported.	Literature review, clinician and patient interviews	Significant improvement in all KTQ domains before and 6 months post-transplantation, except appearance.
Niu et al. 2015 [52]	Cronbach's alpha: 0.70 - 0.90	Test-retest (Pearson's coefficient): 0.70 - 0.91 (2 weeks)	Significant positive correlations with KTQ scores were obtained for PCS and MCS. There were higher correlation co-efficients between the MCS and the KTQ domains (0.34-0.62).	CFA* analysis, (GFI 0.903), RMSEA (0.001), CFI (1.00)	Expert consultation. CVI: >0.80	_
Rebollo et al. 2003 [53]	Cronbach's Alpha: > 0.80 for all except 'Appearance' (0.69).	ICC: > 0.80 for all domains except, 'Physical symptoms' (0.63) and 'Appearance' (0.67). ICC was calculated in 'stable' patients between months 6 and 12 after Tx.	Correlations among the dimensions of the KTQ: 0.32 - 0.72. Correlation coefficient between MCS and fatigue was 0.60. MCS and Energy was 0.78.	_	_	For the 'physical symptoms' domain of the KTQ the effect size at 12th month was 0.48
Modified Edmonton symptom assessment system (ESAS) Davison et al. 2006a [54]	_	Test-retest (ICC), Individual symptoms: 0.53 - 0.71. Total symptom distress score: 0.70 (n = 165, 1 week)	Pearson's Coefficient, all ESAS items significantly correlated with HRQOL scores of KDQOL-SF. Total symptom distress score with PHC ($r = -0.54$, $p<0.01$) with MHC ($r = -0.62$, $p<0.001$)	_	Expert review and patient participation	_
Davison et al. 2006b [55]	_	_	_	_	_	Change in overall symptom distress score correlated with changes in disease-specific (KDQOL-SF) domains, $r = -0.46$ to -0.73, $p<0.01$.

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Modified Transplant Symptom Occurrence and Symptom Distress Scale (MTSOSD) Moons et al. 2001 [56]	–	–	Pearson's correlation coefficient between the ridits of symptom occurrence and symptom distress: 0.54, p=0.0001. Positive correlations between depression and the ridits of symptom occurrence (r=0.32, p=0.001), and symptom distress (r=0.54, p<0.0001) indicating a higher level of symptom experience in depressed people.	–	Literature review Multidisciplinary expert review	–
ReTransQoL (RTQ) Version 1 Beauger et al. 2013 [57]	–	–	–	PCA (Varimax rotation) showed an unsatisfactory structure of the items with the dimensions with a KMO of 0.841. The 5 dimensions were confirmed with a variance of about 50%.	–	–
Gentile et al. 2008 [58]	Cronbach's alpha: 0.70 - 0.86	–	Physical health correlated with all domains of SF36. Significant correlation with PHC and MHC: >0.60. Mental health correlated with all related domains of the SF36 (>0.60). Significant correlation with MHC: >0.76.	Factor analysis (Varimax rotation). Five factors identified which accounted for 46.3% of the total variance. Nine out of 45 items had for their specific dimension a factor loading under 0.40.	Tx patients confirmed the pertinence of the five dimensions, and the relevance of the items.	Significant differences were neither found for the five domains of RTQ nor domains of SF36 among any of these groups of patients.
ReTransQoL (RTQ) Version 2 Beauger et al. 2013 [57]	Cronbach's alpha: 0.7 - 0.9.	–	The physical health domain showed high correlations with domains of SF36 namely: Vitality, General Health, Physical Functioning and Bodily Pain (r >0.65). Women had lower HRQOL scores in comparison with men in every domain (p<0.05).	PCA (Varimax rotation): The structure of the 5 factors was fixed and supported by PCA and the 5 factors accounted for 53.1% of total variance. The structure of the RTQ V2.0 was established with a CFA* providing satisfactory indicators: RMSEA = 0.05, SRMR = 0.049, CFI = 0.97 and GFI = 0.91.	–	–

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
EuroQol EQ-5D Cleemput et al. 2004 [59]	–	–	Chi-square test show older patients have more mobility problems than younger patients ($p<0.01$) and have lower EQ-5D index scores (Kruskal-Wallis test, $p = 0.01$). Spearman rank correlation coefficients between EQ-5D index scores and the SF-36 domain: 0.41 - 0.68 ($P<0.01$).	–	–	–
Medical Outcomes Study Short Form-12 Questionnaire Pakpour et al. 2011b [60]	Cronbach's alpha, PCS (SF-12): 0.89 MCS (SF-12): 0.90	Teat-retest: 0.67 (bodily pain) - 0.90 (physical functioning). All >0.7 except bodily pain. PCS: 0.76 MCS: 0.74 (n = 70, weeks)	Patients that were female, older, less educated, dialysed for longer, and had lower levels of albumin and Kt/V had significantly ($p< 0.05$) poorer quality of life compared to other patient subgroups. All items showed higher correlation when compared with their own scale than when compared with other scales.	CFA was done on the 8 scales loaded with the two latent factors (PCS and MCS). The model provided a good fit for the data, with a fit index of chi2 = 23.30, degree of freedom = 13, P = 0.04, GFI = 0.96, adjusted GFI = 0.90, normed fit index = 0.99, CFI = 0.99, standardized root mean square residual = 0.036, and RMSEA = 0.079.	–	–
Medical Outcomes Study 36-item Short Form Health Survery (SF- 36) v 2.0 Feurer et al. 2004 [61]	Cronbach's alpha, PCS: 0.76 MCS: 0.80	–	Correlation between functional performance (KPS) and PCS: 0.38, $p<0.001$.	–	–	–
Mingardi et al. 1999 [62]	Cronbach's alpha: 0.70 - 0.94	–	Correlation between items and hypothesized scale, overlap corrected): 0.32 - 0.82. Scaling success: 100% for all except general health (80%). Significant but low associations between SF-36 domains and albumin values	–	–	–

S1 Table. Summary of study results						
Measure (Study)	**Measurement properties**					
	Internal consistency	**Reliability** (Sample size, time interval)	**Hypothesis testing**	**Structural validity**	**Content validity**	**Responsiveness**
Nottingham Health Profile (NHP) Badia et al. 1994 [63]	Cronbach's alpha: 0.58 - 0.85 Overall: 0.91	Test-retest, Spearman's coefficient: 0.69 - 0.85. Agreement percentage (AP): 0.48 - 0.65 Overall AP: 0.56	_	_	_	_
Zengin et al. 2014 [64]	Cronbach's alpha NHP: 0.73 - 0.97	_	Spearman's correlations for similar domains of NHP and SF-36: 0.33 - 0.76	_	_	_
TIME TRADE-OFF (modified TTO) Churchill et al. 1987 [65]	_	Test-retest, (ICC): Overall = 0.81 (n = 171, 4 - 8 weeks)	TTO mean scores and Patient VAS scores, Tx patients scored the highest on both scales 0.84 and 0.86 respectively.	_	_	_
WHOQOL-BREF (Dialysis Module) Yang et al. 2006 [66]	Cronbach's alpha: 0.72 - 0.79	Test-retest: 0.61 - 0.79 (n = 20, 4 - 8 weeks)	Pearson's correlation coefficients between the 4 domains and the WHOQOL-BREF global items were less than 0.50 ($p < 0.005$). Correlations between items and their corresponding domains were significant ($p < 0.001$) and greater than those between the items and other domains.	EFA identified 4 factors: general non-social, socioeconomic, physical-environment and physico-psychological factors. CFI (0.93), NNFI (0.92) All factor loadings for the items were significant	_	_

[a] The quality of translations carried out by studies was evaluated as cross-cultural validity was not assessed.

ICC - Internal correlation coefficients, NP - Neuropsychiatric, MHC - Mental health composite, R - Regression, EFA - Exploratory factor analysis, CFA - Common factor analysis, CFA* - Confirmatory factor analysis, SE - Standard error, CVI - Content validity index, PCS - Physical composite summary score, MCS - mental composite summary score, KDCS - Kidney disease composite summary, PCA – Principal component analysis, KMO - Kaiser-Meyer-Olkin, GFI - Goodness-of-fit index, RMSEA - Root mean square error of approximation, CFI - Comparative fit index, BMI - Body mass index, GFR - Glomerular filtration rate

CAD - coronary artery disease, DM – Diabetes, HD – Haemodialysis, PD – Peritoneal dialysis, Tx – Renal transplant, RRT - Renal replacement therapy, QoL – Quality of life

BDI - Beck depression inventory, KPS - Karnofsky performance scale, SIP - Sickness impact profile, PSF - Patient Symptom Form

Appendix 3: Characteristics of included studies (systematic review)

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Abd ElHafeez et al., 2012 [21]	Adaption, reliability and validation study of pre-dialysis CKD patients	Randomly selected pre-dialysis CKD patients (stages 1 – 4) from an out-patient nephrology clinic.	Egypt	100 (46.0)	Median: 54 (42-60)	Not reported
Agarwal 2010 [1]	Development and validation study in non-dialysed patients with CKD	Purpose was to measure the presence and severity of symptoms. Patients with pre-dialysis CKD attending renal clinics. CKD stages were not specified.	USA	92 (5.0)	67.5 (±11.0)	Not reported
Aiyasanon et al., 2009 [3]	Validation and reliability study	Patients attended the outpatient nephrology clinic, chronic PD clinic and chronic HD unit.	Thailand	Main study 110 (49.1)	55.6	HD or PD > 6
Almutary et al., 2015 [5]	Developmental study based on the DSI instrument	The measure was developed to assess symptoms in different stages of CKD. A convenience sample of CKD patients from three hospitals. The majority of the patients used for the study were on dialysis while <25% were pre-dialysis (CKD stages 4 & 5).	Saudi Arabia	433 (47.1)	48.2 (± 14.9)	Not reported
Alvarez-Ude et al., 1997 [44]	Cultural adaptation and validation study in HD patients.	Patients receiving long-term HD (> 3 months) from a General Hospital were enrolled for the study.	Spain	35 (43.0)	67.7 (± 10.9)	CRF: 49.6 (± 61.8) years
Badia et al., 1994 [63]	Validation study in ESRD patients	Patients from a Hospital HD programme.	Spain	At follow up 121 (42.4)	53.0 (±13.0)	HD: 48.0 (± 35.0)

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Barotfi et al., 2006 [22]	Translation and validation study Hungarian version of KDQOL-SF	Patients on maintenance HD at 9 dialysis units and patients from an outpatient transplant centre.	Hungary	HD 418 (44.0) Tx 418 (41.0)	HD 53.0 (±14.0) Tx 49.0 (±12.0)	Maintenance dialysis >3
Bataclan and Dial 2009 [23]	Cultural adaptation and validation study in dialysis patients	Patients with CRD, undergoing renal replacement therapy at a dialysis centre	Philippines	80 (56.0)	53 (±2.0)	Not reported
Beauger et al., 2013 [57]	Phase 1 - Re-evaluation and adaptation study Phase 2 & 3 – Validation study in Tx patients	Patients were Tx recipients. Phase 1 and phase 2 samples were obtained from a cross-sectional multicentre study. Phase 3 sample came from a cross-sectional multicentre study (QUAVIREIN: French translation of Renal Quality of Life study).	France	Phase 1 & 2 1059 (38.1) Phase 3 1591 (39.5)	55.2 (±12.4) (Phase 1 &2) 55.3 (±14.2) (Phase 3, QUAVIREIN)	Not reported
Boini et al., 2007 [24]	Translation and validation study in dialysis patients	Dialysis patients from 5 centres in France	France	68 (45.6)	66.0 (±15.2)	Renal failure: 3.9 (± 3.7) years
Bouidida et al., 2014 [25]	Cultural adaptation and validation study	Hospital patients on HD and PD.	Morocco	80 (70.0)	43.9 (± 14.2)	Dialysis 104.8 (± 120)
Chao et al., 2016 [14]	Validation study in patients with CKD 1 - 5	Participants were patients with CKD stages 1 - 5 from 3 CKD clinics.	Taiwan	428 (41.0)	62.0 (±12.3)	Not reported

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Cheung et al., 2012 [26]	Translational and validation study in ESRD patients	This analysis is part of a larger study on the QOL and role of palliative care in ESRD (CKD stage 5). Participants were from renal wards and outpatient clinics of a tertiary hospital. A proportion later commences dialysis.	Singapore	78 (44.9)	Median 76 (52 - 98)	Not reported
Chisholm-Burns et al., 2011 [50]	Validation study of KTQ	Renal transplant recipients enrolled in an RCT, on immunosuppressant therapy.	USA	100 (48.0)	51.0 (±13.2)	Post-Tx 39.7 (± 29.8)
Chow and Tam 2014 [15]	Validation and reliability study of Cantonese Chinese version of KDQOL-36	Patients attending the renal dialysis unit of a regional hospital and its satellite dialysis centre.	China	D 110 (31.8) Tx 122 (44.3)	D 58.2 (±15.2) Tx 51.8 (±10.3)	Dialysis ≥ 3 Tx≥ 1 year
Churchill et al., 1987 [65]	Cross-sectional study of modified Time Trade-Off (TTO)	ESRD patients from regional hospital nephrology programme (dialysis and transplant patients)	Canada	171 (?)	?	≥3months
Cleemput et al., 2004 [59]	Validation study	The secondary data analyses of a study of Tx patients following an immunosuppressive regimen. Tx patients from 3 centres.	Belgium, Netherlands	361 (40.3)	51.7 (± 13.0)	Median post-Tx: 16.7 (Q1 = 7.9; Q3 = 38.6)
Davison et al., 2006a [54]	Validation study of modified ESAS in dialysis patients	Peritoneal dialysis patients and in-centre and HD patients from satellite units.	Canada	507 (47.3)	63.5 (±16.0)	Dialysis: 3.4 (±2.8) years
Davison et al., 2006b [55]	Validation study of a modified ESAS in HD patients	Longitudinal study of HD patients, in-centre and from satellite units..	Canada	At follow up 261 (43.7)	At follow up 64.1 (±15.9)	Dialysis: 3.8 (± 3.2) years

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Dehesh et al., 2014 [8]	Validation study in HD patients	HD patients from a dialysis centre and two hospitals.	Shiraz, Iran	150 (41.0)	50.5 (±15.1)	HD ≥ 6
Duarte et al., 2005 [27]	Cultural adaptation and validation study in ESRD patients	Random sample of ESRD patients on chronic dialysis from a dialysis unit.	Sao Paulo, Brazil	Reliability test 93 (45.0) Validity test sample 74	49.0 (±13.0)	Median time ESRD 2 years
Fardinmehr et al., 2012 [28]	Reliability and validation study	ESRD patients undergoing regular HD at hospitals.	Iran	50 (34.0)	52.7 (± 4.2)	Dialysis: 37.8 (± 36.0)
Ferrans and Powers 1985 [9]	Developmental study of instrument for use in healthy and ill individuals	The QLI was developed to measure QOL of healthy individuals, as well as those on dialysis.	USA	Patients 37 (28)	50 (±14.18)	Not reported
Feurer et al., 2004 [61]	Re-validation study in Tx patients	Clinics for solid organ transplant candidates and recipients. Data for this study was collected during an earlier study exploring HRQOL after solid organ transplantation.	USA	112 (39.0)	44.0 (±1.3)	Not reported
Franke et al., 1999 [46]	Developmental study	The ESRD SCL-TM was developed to assess the QOL of Tx recipients with a focus on side effects of immunosuppression therapy. The study was conducted during regular outpatient visits.	Germany	458 (45.0)	48 (± 13.0)	Tx: 77.6 (± 51.6)
Gentile et al., 2008 [58]	Development and validation study in Tx patients	Tx recipients were randomly selected from the registry of a transplant centre.	France	Validation phase 130 (36.5)	Validation phase 49.4 (± 12.8)	Validation phase 5.4 (± 3.1) years

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Green et al., 2001 [29]	Cultural adaptation and validation study in HD patients	Patients from dialysis centres	Japan	712 (793?) unknown (39.1?)	55	Dialysis > 3
Halabi 2006 [10]	Translation, cultural adaptation and validation study	Dialysis patients		270 (?)	?	Not reported
Hays et al., 1994 [12]	Developmental study	Patients on dialysis at 9 outpatient dialysis centres.	USA	165 (52.0)	53 (22 - 87)	78% on dialysis at the same centre ≥1 year 1 month.
Hays et al., 1995 [30]	Developmental study	The KDQOL Short Form was developed as a short version of the original KDQOL instrument.	USA	165 (?)	Not reported	Not reported
Joshi et al., 2010 [31]	Validation study	HD patients at 22 dialysis centres.	Singapore	'Evaluable' sample 980 (43.9)	'Evaluable' sample 56 (± 21)	HD > 3
Kleinman et al., 2006 [49]	Validation study	Post-Tx patients on immunosuppression therapy attending 5 clinics in 4 countries.	Australia, Canada, Germany, and Switzerland.	96 (43.7)	47.4 (± 12.3)	Tx: 3.6 (± 3.6) years
Klersy et al., 2007 [32]	Cultural adaptation and validation study	A convenience sample of pre-dialysis (CKD stage not specified) and dialysis patients was enrolled at 2 referral centres. Some of the participants were hospitalized patients.	Italy	188 (37.0)	60.0 (± 16.0)	Dialysis (IQR): 1.7 (0.8-5.4) years

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Kontodimopoulos & Niakas 2005 [33]	Reliability and validation study	ESRD patients in 20 dialysis units	Greece	483 (38.8)	59.9 (±14.6)	HD: 6.9 (± 5.7) years
Kontodimopoulos & Niakas 2007 [34]	Reliability and Validation study	ESRD patients were recruited from 25 dialysis units.	Greece	642 (38.7)	58.1 (±14.9)	Dialysis: 6.2 (± 5.7) years
Korevaar et al., 2002 [35]	Validation study of disease-specific scales only	Multi-centre prospective study in 32 centres, the Netherlands Cooperative Study on the Adequacy of Dialysis (NECOSAD).	Netherlands	375 (39.0)	60.0 (±16.0)	Hospitalization in last 3 months: 8.8 (+ 15.0) days
Korkut 2007 [11]	Translational, Reliability and validation study	HD patients at a dialysis centre	Turkey	31 (48.4)	45.7 (±13.5)	Dialysis: 5.1 (1-12) years
Laupacis et al., 1992 [45]	Developmental study	The questionnaire was developed for use in clinical trials in patients on chronic HD. Stable HD patients from 3 hospitals which provide in-hospital HD were recruited.	Canada	50 (40.0)	55.0 (± 15.0)	Dialysis: 41 (± 2.0)
Laupacis et al., 1993 [51]	Developmental study	The KTQ was assessed in an on-going prospective cohort study of patients before and after renal transplantation at two university centres.	Canada	Development sample: 50 (28.0) Psychometric study sample: 73 (?)	Development sample: 43 (19 - 65)	Development sample (Post-Tx): 1.3 (0.2 - 3.0) years Psychometric sample (Post-Tx): ≤ 6
Malindretos 2010 [36]	Cultural adaptation and validation study.	Randomly selected patients undergoing HD in 6 Renal Units.	Greece	200 (45.0)	62.9 (± 14.7)	HD: 47.3 (± 57.9)

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Mateti 2015 [16]	Cross-cultural adaptation. reliability and validation study	Patients were selected randomly from out-patient HD units from three centres.	Indian	82 (28.1)	50.7(± 12.3)	HD: 38.1(± 25.3)
Mingardi et al., 1999 [62]	Validation study	This study was part of a large, multicentre trial (DIA-QOL). Patients were treated at 4 dialysis units.	Italy	246 (48.2)	60.8 (±13.1)	Not reported
Molsted et al., 2005 [37]	Reliability study	HD and PD patients from 2 hospitals were recruited.	Denmark	130 (33.0)	59.0	HD: 39 (± 59) PD: 27 (± 26)
Moons et al., 2001 [56]	Translation and Validation study	Post Tx out-patients of a university hospital on a maintenance immunosuppression.	Belgium	108 (39)	47.0	Median post-Tx 5.5 years
Moreira et al., 2009 [38]	Reliability and validation study	CRF patients on HD treatment at a clinic.	Brazil	147 (48.3)	51.8 (± 14.2)	HD:28.6 (± 30.5)
Niu et al., 2015 [52]	Cultural adaptation, Reliability and validation study	Tx recipients from a follow-up clinic.	China	136 (27.9)	43.9(± 11.4)	Tx: 40.4(± 32.9)
Önsoz and Yesilbalkan 2013 [6]	Cultural adaptation and reliability study	Chronic HD patients from two dialysis units.	Turkey	120 (44.2)	54.5 (± 13.8)	HD: 48.1 (41.2)
Ortega et al., 2007 [47]	Validation and reliability study	To validate for routine use in clinical practice. Patients were recruited from Tx waiting list at 16 hospitals.	Spain	307 (40.8)	51.6 (±12)	Not reported
Pakpour et al., 2011a [39]	Cultural adaptation and validation study	Randomly sampled patients who had been referred to dialysis centres.	Iran	212 (43.8)	57.5 (+ 14.7)	Dialysis: 47.9 (± 30.2)

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Pakpour et al., 2011b [60]	Reliability and validation study	HD patients from dialysis centres.	Iran	144 (42.4)	54.6(± 11.4)	Dialysis: 30.9 (± 22.2)
Park et al., 2007 [40]	Reliability and validation study	HD and CAPD patients at a university dialysis centre.	Korea	164 (41.5)	Men: 54.0 (± 12.7) Women: 54.3 (± 13.6)	Dialysis: 56.7 (± 61.3)
Perneger 2003 [41]	Cross-sectional study of ESRD patients	ESRD patients treated by in-centre HD identified from patient lists provided by 4 dialysis units.	Geneva canton, Switzerland	83 (37.0)	60.0 (±15.5)	Not reported
Rao et al., 2000 [13]	Developmental and validation study of subscales of KDQOL instrument.	Secondary analysis of data collected for another study of HD patients.	USA	165 (52.0)	20 - 65+	Not reported
Rebollo et al., 2003 [53]	Cross-cultural adaptation and validation study.	ESRD patients undergoing chronic HD or PD on the Tx waiting list were recruited at the moment of the pre-transplant examination.	Spain	54 (32.0)	Median age: 51.0 Interc. Range: (38.0 – 57.0)	Not reported
Ricardo et al., 2013 [17]	Reliability and validation study	Individuals with mild-to-moderate CKD (CKD stages not clearly defined) enrolled in the Chronic Renal Insufficiency Cohort (CRIC) Study and the Hispanic CRIC (HCRIC) Study.	USA	829 (40.0)	57.0 (±11.6)	Not reported
Stavem and Ganss 2006 [48]	Reliability and validation study	Tx recipients of a hospital nephrology outpatient clinic.	Norway	53 (42.0)	57.5 (±13.2)	Tx median (range): 4.4 (0.4 - 29.1) years

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Suet-Ching 2001 [2]	Development and validation study	Patients from dialysis centres.	Hong Kong, China	164 (51.0)	41 - 70 (70% total sample)	Not reported
Tao et al., 2014 [18]	Reliability and validation study	Patients with mild-to-severe CKD (CKD stages 1 – 4) were recruited from the renal wards and outpatient dialysis clinics of a tertiary hospital.	Mainland China	103 (44.8)	47.6 (14.2)	Dialysis: 45.9 (±41.4)
Thaweethamcharoen et al., 2013 [19]	Reliability and validation study	HD and PD patients were recruited from hospital outpatient clinics.	Thailand	HD 167 (52.1) PD 62 (49.1)	HD: 57.5 (±15.1) PD: 67.6 (±11.5)	HD: 7.4 (±5.4) years PD: 1.9 (±1.2) years
Vasilieva 2007 [42]	Cross-cultural adaptation and validation study	Chronic HD patients	Russia	91 (42.0)	50.2 (± 12.8)	Not reported
Weisbord et al., 2004 [7]	Developmental study	This study is part of a larger study of symptoms in ESRD. HD patients were selected from a dialysis unit.	USA	20 (5) (Test re-test)	Not reported	Not reported
Wu et al., 2001 [4]	Developmental study	This instrument was developed as part of CHOICE study, for patients with ESRD to complement the SF-36. A random sample of HD and PD patients attending 4 dialysis centres.	USA	928 (46.0) HD: 694 PD: 234	58	Dialysis< 1 year
Yang 2013 [20]	Validation study	Secondary data analysis study of data from a cross-sectional survey of HD patients treated at a dialysis centre.	Singapore	394 (44.2)	52.4 (±11.7)	Dialysis: 69.0 (±50.6)

S2 Table. Characteristics of included studies						
Study	**Type of study**	**Study population and setting**	**Country**	**Sample size** (% of females)	**Mean age in years** (SD or range)	**Duration of condition/treatment** Months, Mean (SD) unless otherwise stated
Yang et al., 2006 [66]	Development and validation study	Patients undergoing regular HD at the dialysis centres of 10 regional hospitals or outpatient clinics.	Taiwan	249 (49.0)	62% < 50 years old; 20% < 40 years old.	Not reported
Yildirim et al., 2007 [43]	Cultural adaptation and validation study	Patients with ESRD randomly selected from dialysis units of 3 educational and research hospitals.	Turkey	82 (65.0)	51.0 (± 12.0)	Dialysis: median 3 years.
Zengin et al., 2014 [64]	Cross-sectional study	Patients receiving treatment in 3 different HD centres.	Turkey	172 (40.7)	49.7 (± 14.7)	HD: 7.5 years

D - Dialysis; HD - Hemodialysis; Tx - Renal transplant; ESRD – End stage Renal Disease; CKD – Chronic Kidney Disease; CRD – Chronic Renal Disease; CRF – Chronic Renal Failure; SD - Standard deviation; CAPD - Continuous ambulatory peritoneal dialysis; QOL – Quality of life.

References

1. Agarwal R. Developing a self-administered CKD symptom assessment instrument. Nephrology Dialysis Transplantation 2010;**25**(1):160-66.
2. Suet-Ching WL. The psychometric properties of the Chinese Dialysis Quality of Life Scale for Hong Kong dialysis patients. Journal of advanced nursing 2001;**36**(3):441-49.
3. Aiyasanon N, Premasathian N, Nimmannit A, et al. Validity and reliability of CHOICE health experience questionnaire: Thai version. Journal of the Medical Association of Thailand 2009;**92**(9):1159-66.
4. Wu AW, Fink NE, Cagney KA, 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 : the official journal of the National Kidney Foundation 2001;**37**(1):11-21.

5. Almutary H, Bonner A, Douglas C. Arabic translation, adaptation and modification of the dialysis symptom index for chronic kidney disease stages four and five. BMC nephrology 2015;**16**(1):1-8.
6. Onsoz H. UYO. Reliability and validity of the Turkish version of the Dialysis Symptom Index in chronic hemodialysis patients. Turkish Nephrology, Dialysis and Tranplantation Journal 2013;**22**(1):60-67.
7. Weisbord SD, Fried LF, Arnold RM, 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):226-40.
8. Dehesh T, Zare N, Jafari P, et al. Psychometric assessment of the Persian version of the Ferrans and Powers 3.0 index in hemodialysis patients. International Urology and Nephrology 2014;**46**(6):1183-89.
9. Ferrans CE, Powers MJ. Quality of life index: development and psychometric properties. Ans 1985;**Advances in nursing science. 8**(1):15-24.
10. Halabi JO. Psychometric properties of the Arabic version of Quality of Life Index. Journal of Advanced Nursing 2006;**55**(5):604-10.
11. Korkut Y. The reliability and validity study of the Turkish version of Ferrans and Powers' Quality of Life Index for dialysis patients. Archives of Neuropsychiatry 2007;**44**(14).
12. Hays RD, Kallich JD, Mapes DL, et al. Development of the kidney disease quality of life (KDQOL) instrument. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 1994;**3**(5):329-38.
13. Rao S, Carter WB, Mapes DL, 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):1099-111.
14. Chao S, Yen M, Lin TC, et al. Psychometric Properties of the Kidney Disease Quality of Life-36 Questionnaire (KDQOL-36). West J Nurs Res 2016;**38**(8):1067-82.
15. Chow SK, Tam BM. Is the kidney disease quality of life-36 (KDQOL-36) a valid instrument for Chinese dialysis patients? BMC nephrology 2014;**15**:199.
16. Mateti UV, Nagappa AN, Attur RP, et al. Cross-cultural adaptation, validation and reliability of the South Indian (Kannada) version of the Kidney Disease and Quality of Life (KDQOL-36) instrument. Saudi J Kidney Dis Transpl 2015;**26**(6):1246-52.
17. Ricardo AC, Hacker E, Lora CM, et al. Validation of the Kidney Disease Quality of Life Short Form 36 (KDQOL-36TM) US Spanish and English versions in a cohort of Hispanics with chronic kidney disease. Ethnicity and Disease 2013;**23**(2):202-09.
18. Tao X, Chow SK, Wong FK. Determining the validity and reliability of the Chinese version of the Kidney Disease Quality of Life Questionnaire (KDQOL-36). BMC nephrology 2014;**15**:115.
19. Thaweethamcharoen T, Srimongkol W, Noparatayaporn P, et al. Validity and reliability of KDQOL-36 in thai kidney disease patient. Value in Health Regional Issues 2013;**2**(1):98-102.
20. Yang F, Wang VW, Joshi VD, et al. Validation of the English version of the Kidney Disease Quality of Life questionnaire (KDQOL-36) in haemodialysis patients in Singapore. Patient 2013;**6**(2):135-41.
21. Abd Elhafeez S, Sallam SA, Gad ZM, et al. Cultural adaptation and validation of the "kidney Disease and Quality of Life - Short Form (KDQOL-SFTM) version 1.3" questionnaire in Egypt. BMC nephrology 2012;**13 (1) (no pagination)**(170).
22. Barotfi S, Molnar MZ, Almasi C, et al. Validation of the Kidney Disease Quality of Life-Short Form questionnaire in kidney transplant patients. Journal of Psychosomatic Research 2006;**60**(5):495-504.

23. Bataclan RP, Dial MAD. Cultural adaptation and validation of the Filipino version of Kidney Disease Quality of Life - Short Form (KDQOL-SF version 1.3). Nephrology 2009;**14**(7):663-68.
24. Boini S, Leplege A, Loos Ayav C, et al. Measuring quality of life in end-stage renal disease. Transcultural adaptation and validation of the specific Kidney Disease Quality of Life questionnaire. [French]. Nephrologie et Therapeutique 2007;**3**(6):372-83.
25. Rhou H, Ezaitouni F, Ouzeddoun N, et al. Translation, cultural adaptation and validation of the kidney disease quality of life-short form 1.3 in an African country. Transplantation Proceedings 2014;**46**(5):1295-301.
26. Cheung YB, Seow YY, Qu LM, et al. Measurement properties of the Chinese version of the Kidney Disease Quality of Life-Short Form (KDQOL-SFTM) in end-stage renal disease patients with poor prognosis in singapore. Journal of Pain and Symptom Management 2012;**44**(6):923-32.
27. Duarte PS, Ciconelli RM, Sesso R. Cultural adaptation and validation of the "Kidney Disease and Quality of Life - Short Form (KDQOL-SFTM 1.3)" in Brazil. Brazilian Journal of Medical and Biological Research 2005;**38**(2):261-70.
28. Fardinmehr O, Farajzadegan Z, Naini AE, et al. The validity and reliability of the persian version of kidney disease quality of life questionnaire-short form (KDQOL-SF) in Iranian patients. [Persian]. Journal of Isfahan Medical School 2012;**29**(165).
29. Green J, Fukuhara S, Shinzato T, et al. Translation, cultural adaptation, and initial reliability and multitrait testing of the Kidney Disease Quality of Life Instrument for use in Japan. Quality of Life Research 2001;**10**(1):93-100.
30. Hays R. D. KJD, Mapes D. L., Coons S. J., Amin N., Carter W. B., Kamberg C. Kidney Disease Quality of Life Short Form (KDQOL-SF) Version 1.3: A manual for use and scoring. RAND 1995.
31. Joshi VD, Mooppil N, Lim JFY. Validation of the kidney disease quality of life-short form: A cross-sectional study of a dialysis-targeted health measure in singapore. BMC nephrology 2010;**11 (1) (no pagination)**(36).
32. Klersy C, Callegari A, Giorgi I, et al. Italian translation, cultural adaptation and validation of KDQOL-SF, version 1.3, in patients with severe renal failure. J Nephrol 2007;**20**(1):43-51.
33. Kontodimopoulos N, Niakas D. Determining the basic psychometric properties of the Greek KDQOL-SFTM. Quality of Life Research 2005;**14**(8):1967-75.
34. Kontodimopoulos N, Niakas D. Measuring health-related quality of life of Greek dialysis patients with the KDQOL-SFTM. [Greek]. Archives of Hellenic Medicine 2007;**24**(6):590-99.
35. Korevaar JC, Merkus MP, Jansen MAM, et al. Validation of the KDQOL-SFTM: A dialysis-targeted health measure. Quality of Life Research 2002;**11**(5):437-47.
36. Malindretos P, Sarafidis P, Spaia S, et al. Adaptation and validation of the Kidney Disease Quality of Life-Short Form questionnaire in the Greek language. Am J Nephrol 2010;**31**(1):9-14.
37. Molsted S, Heaf J, Prescott L, et al. Reliability testing of the Danish version of the Kidney Disease Quality of Life Short Form. Scand J Urol Nephrol 2005;**39**(6):498-502.
38. Moreira CA, Garletti Junior W, Lima LF, et al. Assessment of the basic psychometric properties for the portuguese version of the KDQOL-SFTM. [Portuguese]. Rev Assoc Med Bras 2009;**55**(1):22-28.
39. Pakpour AH, Yekaninejad M, Molsted S, et al. Translation, cultural adaptation assessment, and both validity and reliability testing of the Kidney Disease Quality of Life - Short Form version 1.3 for use with Iranian patients. Nephrology 2011;**16**(1):106-12.

40. Park HJ, Kim S, Yong JS, et al. Reliability and validity of the Korean version of kidney disease quality of life instrument (KDQOL-SFTM). Tohoku Journal of Experimental Medicine 2007;**211**(4):321-29.
41. Perneger TV, Leski M, Chopard-Stoermann C, et al. Assessment of health status in chronic hemodialysis patients. Journal of Nephrology 2003;**16**(2):252-9.
42. Vasilieva IA. Russian version of the questionnaire The Kidney Disease and Quality of Life Short Form (KDQOL-SF) a valuable diagnostic instrument for assessing quality of life of dialysis patients. Nefrologiya 2007;**11**(1):64-70.
43. Yildirim A, Ogutmen B, Bektas G, et al. Translation, cultural adaptation, initial reliability, and validation of the Kidney Disease and Quality of Life-Short Form (KDQOL-SF 1.3) in Turkey. Transplant Proc 2007;**39**(1):51-4.
44. Alvarez-Ude F, Galan P, Vicente E, et al. Adaptation and validation of the Spanish version of 'Kidney Disease Questionnaire'. [Spanish]. Nefrologia 1997;**17**(6):486-96.
45. Laupacis A, Muirhead N, Keown P, et al. A disease-specific questionnaire for assessing quality of life in patients on hemodialysis. Nephron 1992;**60**(3):302-6.
46. Franke GH, Reimer J, Kohnle M, et al. Quality of life in end-stage renal disease patients after successful kidney transplantation: development of the ESRD symptom checklist - transplantation module. Nephron 1999;**83**(1):31-9.
47. Ortega T, Valdes C, Rebollo P, et al. Evaluation of reliability and validity of Spanish version of the end-stage renal disease symptom checklist-transplantation module. Transplantation 2007;**84**(11):1428-35.
48. Stavem K, Ganss R. Reliability and validity of the ESRD Symptom Checklist--Transplantation Module in Norwegian kidney transplant recipients. BMC nephrology 2006;**7**:17.
49. Kleinman L, Kilburg A, Machnicki G, 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):1223-32.
50. Chisholm-Burns MA, Erickson SR, Spivey CA, et al. Concurrent validity of kidney transplant questionnaire in US renal transplant recipients. Patient Preference and Adherence 2011;**5**:517-22.
51. Laupacis A, Pus N, Muirhead N, et al. Disease-specific questionnaire for patients with a renal transplant. Nephron 1993;**64**(2):226-31.
52. Niu Y, Zhang W, Mao S, et al. Pilot feasibility research of Chinese version of kidney transplant questionnaire in recipients of living donor kidney transplantation. Int J Clin Exp Med 2015;**8**(12):22570-6.
53. Rebollo P, Ortega F, Ortega T, et al. Spanish validation of the "kidney transplant questionnaire": a useful instrument for assessing health related quality of life in kidney transplant patients. Health and quality of life outcomes 2003;**1**:56.
54. 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 2006;**69**(9):1621-5.
55. Davison SN, Jhangri GS, Johnson JA. Longitudinal validation of a modified Edmonton symptom assessment system (ESAS) in haemodialysis patients. Nephrology Dialysis Transplantation 2006;**21**(11):3189-95.
56. Moons P, De Geest S, Versteven K, et al. Psychometric properties of the "Modified Transplant Symptom Occurrence and Symptom Distress Scale". Journal of nursing measurement 2001;**9**(2):115-34.
57. Beauger D, Gentile S, Jouve E, et al. Analysis, evaluation and adaptation of the ReTransQoL: A specific quality of life questionnaire for renal transplant recipients. Health and quality of life outcomes 2013;**11 (1) (no pagination)**(148).

58. Gentile S, Jouve E, Dussol B, et al. Development and validation of a French patient-based health-related quality of life instrument in kidney transplant: The ReTransQoL. Health and quality of life outcomes 2008;**6 (no pagination)**(78).
59. Cleemput I, Kesteloot K, Moons P, et al. The construct and concurrent validity of the EQ-5D in a renal transplant population. Value in Health 2004;**7**(4):499-509.
60. Pakpour AH, Nourozi S, Molsted S, et al. Validity and reliability of short form-12 questionnaire in Iranian hemodialysis patients. Iran J Kidney Dis 2011;**5**(3):175-81.
61. Feurer ID, Moore DE, Speroff T, et al. Refining a health-related quality of life assessment strategy for solid organ transplant patients. Clinical Transplantation, Supplement 2004;**18**(12):39-45.
62. Mingardi G, Cornalba L, Cortinovis E, et al. Health-related quality of life in dialysis patients. A report from an Italian study using the SF-36 Health Survey. Nephrology Dialysis Transplantation 1999;**14**(6):1503-10.
63. Badia X, Alonso J, Brosa M, et al. Reliability of the Spanish version of the Nottingham health profile in patients with stable end-stage renal disease. Social Science and Medicine 1994;**38**(1):153-58.
64. Zengin N, Oren B, Gul A, et al. Assessment of quality of life in haemodialysis patients: A comparison of the Nottingham Health Profile and the Short Form 36. International Journal of Nursing Practice 2014;**20**(2):115-25.
65. Churchill DN, Torrance GW, Taylor DW, et al. Measurement of quality of life in end-stage renal disease: the time trade-off approach. Clin Invest Med 1987;**10**(1):14-20.
66. Yang SC, Kuo PW, Wang JD, et al. Development and pyschometric properties of the dialysis module of the WHOQOL-BREF Taiwan version. Journal of the Formosan Medical Association 2006;**105**(4):299-309.

Appendix 4: Characteristics of included PROMs (systematic review)

S3 Table. Characteristics of included PROMs

Instrument name	Study ID	Study language(s)	Summary of instrument	Scoring method	Mode of administration	Recall period
Agarwal (unnamed)	Agarwal	English	37-item instrument Mental and physical domains	0 - 100 scale (worst to best)	Self-administered	4 weeks
Chinese Dialysis Quality of Life Scale (CDQOL)	Suet-Ching	Cantonese	29-item instrument	5-point Likert scale	Self-administered	Not reported
CHOICE Health Experience Questionnaire (CHEQ)	Wu Aiyasanon	English Thai	83-item instrument. 8 domains from SF36 (36 items) with 7 supplementary items for the CHEQ. 13 specific domains (38 items) 2 person-specific quality of life items.	Responses are transformed by simple linear regression to a scale of 0 - 100.	Self-administered	4 weeks
CKD Symptom Burden Index (CKD-SBI)	Almutary	Arabic	32-item instrument. CKD-SBI is a modification of the DSI for use in CKD stages IV and V. The CKD-SBI has 4 symptom dimensions.	The prevalence scale uses a Yes/No method. The range of scores is 0-32. A higher overall score indicates a higher symptom occurrence. Distress, severity and frequency are rated using a 0 to 10 scale. A maximum total score for each of these scales is 320 (range 0-320). Higher scores indicate the higher symptom burden.	Self-reported	Extended from 1 week in the original DSI to 4 weeks to reflect the chronicity of CKD symptoms

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
Dialysis Symptom Index (DSI)	Weisbord Onsoz	English Turkish	30-item instrument	A 5-point Likert scale. Two scores are obtained. An overall symptom burden score and a total symptom severity score	Self or interviewer administered	Past week
End stage Renal Disease Symptom Checklist - Transplant Module (ESRD-SCL)	Franke Ortega Stavem	German Spanish Norwegian	43-item instrument. 6 dimensions.	5-point Likert scale. A high score indicates a worse HRQOL. Overall scores range from 0 - 100.	Interviewer and self-administered	Not reported
EuroQol EQ-5D	Cleemput	Dutch, French	The EQ-5D has a self-classifier and a visual analogue scale (VAS) which patients can use to value their current health state. The self-classifier includes 5 dimensions and each dimension has 3 levels of severity (no problems, some problems, and severe problems).	Using the EQ-5D it is possible to describe 243 different health states. Each of these health states corresponds with a single index measure between 0 (dead) and 1 (perfect health), obtained from the preferences of the general public.	Self-report	Not reported
Gastrointestinal Symptom Rating Scale (GSRS)	Kleinman	German, English	15-item instrument. Covers important GI symptoms that may occur with various immunosuppressive therapies. There are 5 subscales	Each subscale gives an average score ranging from 1 (no discomfort) to 7 (very severe discomfort). Higher scores indicate worse impact	Self-administered	Not reported
Gastrointestinal Quality of Life Index (GIQLI)	Kleinman	German, English	36-item instrument. Focuses on the impact GI complaints on an individual's HRQOL. There are 5 subscales:	Subscale scores range from 0 to 4. A total score ranging from 0–144) can be obtained. Higher scores indicate better HRQOL.	Self-administered	Not reported

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
Kidney Disease Questionnaire (KDQ)	Alvarez-Ude Laupacis	Spanish English	26-item instrument. 5 dimensions Note that the physical symptom dimension is patient-specific, thus the symptoms most important to individual patients are identified and used.	7-point Likert scale.	Interviewer	2 weeks
Kidney Disease Quality of Life-36 questionnaire (KDQOL-36)	Chao Chow Mateti Ricardo Tao T'charoen Yang 2013	Taiwanese Cantonese Kannada Spanish English Mandarin Thai English	36-item instrument. 3 specific scales from the KDQOL-SF™v1.3. Summary scales from generic SF-12.	The raw scores are transformed linearly to a range of 0 to 100. Higher scores indicate better HRQOL The scores for the PCS and MCS are converted to T-scores with a mean of 50 and a standard deviation of 10.	Self, proxy and interviewer administered	Not reported

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
Kidney Disease Quality of Life-Short Form questionnaire (KDQOL-SF™)	Abd ElHafeez	Arabic	80-item instrument. Derived from KDQOL-SF. Version 1.3 differs from the 1.2 version only by the addition of a screening question about sexual activity. 8 generic (SF-36) dimensions, 8 disease-specific dimensions and 3 additional dimensions	Scores are calculated so that each domain has a potential range from 0 to 100. Higher scores indicate better HRQOL	Self or interviewer administered	4 weeks
	Barotfi	Hungarian				
	Bataclan	Filipino				
	Boini	French				
	Bouidida	Moroccan				
	Cheung	Chinese				
	Duarte	Portuguese				
	Fardinmehr	Persian				
	Green	Japanese				
	Hays	English				
	Joshi	English				
	Klersy	Italian				
	Kontodimopoulos	Greek				
	Kontodimopoulos	Greek				
	Korevaar	Dutch				
	Malindretos	Greek				
	Molsted	Danish				
	Moreira	Portuguese				
	Pakpour	Farsi				
	Park	Korean				
	Perneger	French				
	Vasilieva	Russian				
	Yildirim	Turkish				

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
Kidney Disease Quality of Life - KDQOL (Dialysis version)	Hays 1994	English	134-item instrument. KDQOL consists of SF36 generic core with 8 domains, 11 kidney disease targeted scales and an item that assesses change in health over a year (overall health rating).	Scores transformed linearly into 0 - 100 point scale. Higher scores indicate better HRQOL.	Self or interviewer administered	Last 30 days
Kidney Disease Quality of Life - KDQOL (Modified)	Rao	English	55 items from S/P and effects of KD. Using affinity mapping, 11 subscales were identified while 4 items were ungrouped.	0 to 100 scale.	Not reported	Not reported
Kidney Transplant Questionnaire (KTQ)	Chisholm-Burns Laupacis Niu Rebollo	English English Chinese Spanish	25-item disease-specific measure, 5 domains. The physical symptom domain is patient-specific, thus the symptoms most important to individual patients are identified and used.	For each domain, an average score from 1 to 7 is calculated.	Self or interviewer administered	2 weeks
Modified Edmonton Symptom Assessment System (ESAS)	Davison a & b	English	10 symptom specific items (9 original items)	10 visual analogue scales with superimposed 0-10 scale. Total score range from 0 to 100.	Self-completed	Not reported
Modified Transplant Symptom Occurrence and Symptom Distress Scale (MTSOSD)	Moons	Dutch	29-item instrument. 20 items relate to the symptom occurrence domain. 9 items relate to symptom distress domain.	Vertical scaling for symptom occurrence and a horizontal technique for symptom distress. 5-point scale (0-4)	Self-reported	Not reported

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
Nottingham Health Profile (NHP)	Badia Zengin	Spanish Turkish	38 yes/no questions. Part 1 has 6 dimensions while Part 2 is optional.	Each dimension is scored by multiplying the total positive answers with 100 and dividing the result by the number of items. The NHP scores range between 0 (good health status) and 100 (poor health status).	Self and interviewer administered	Not reported
Quality of Life Index 3.0	Dehesh Ferrans Halabi Korkut	Persian English Arabic Turkish	68-items in total. Instrument is divided into 2 sections: One section measures satisfaction with various domains of life, while the second measures the importance of the domain to the individual. Three additional items relative to dialysis treatment were added to each section.	Ratings are made on a scale of 1 to 6 for both parts. Scores are calculated by weighing each satisfaction response with its paired importance response. The total QOL score and the four subscale scores range between 0 and 30.	Self or interviewer administered	Not reported
ReTransQoL (RTQ) v1	Beauger Gentile	French French	45-item instrument, 5 dimensions	Questions are equally weighted. The score for each individual is obtained by computing each item's mean score within every dimension. All dimensions are linearly transformed to a 0–100 scale. A score of 100 indicates the highest QOL.	Self-administered	Not reported

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
ReTransQoL (RTQ) v2	Beauger	French	32-item instrument, 5 dimensions.	Questions are equally weighted. The score for each individual is obtained by computing each item's mean score within every dimension. All dimensions are linearly transformed to a 0–100 scale. A score of 100 indicates the highest QOL.	Self-administered	Not reported
Short Form-12 (SF-12)	Pakpour (b)	Persian	12 items (questions) and 8 scales. These 8 scales can also be computed into 2 distinct clusters of PCS-12 and MCS-12	Not reported	Interviewer	Not reported
Short Form-36 (SF-36)	Feurer Mingardi	English Italian	36-item instrument, 8 scales	After standardizing the 8 scales to z-scores based on means and standard deviations for the general US population, aggregate physical and mental component summary scales are calculated by adding the eight differentially weighted scales. Finally, each component score is transformed to a normalized T-score with a mean of 50 and a standard deviation of 10.	Self-reporting Administration by a trained interviewer in person or over the telephone	Standard (4-week) version or Acute (1-week) version
WHOQOL-BREF (Dialysis module)	Yang 2006	Taiwanese	32-item instrument. The core WHOQOL-BREF (TW) has 4 domains. 4 HD specific items incorporated into the existing domains. 2 global items.	5-point Likert scale. The score for each domain ranges from 4 to 20. This is calculated by multiplying the average scores for all items in the domain by 4.	Self or interviewer administered	Recent month

S3 Table. Characteristics of included PROMs						
Instrument name	**Study ID**	**Study language(s)**	**Summary of instrument**	**Scoring method**	**Mode of administration**	**Recall period**
Modified Time Trade-Off (TTO)	Churchill 1987	English	Utility measure	The lower value of 0 represents an extremely low QOL in those patients for whom death would be as preferable as their present state of health while the upper value of 1 represents a very high QOL in those patients whose present health state would be equivalent to their perception of perfect health.	Interviewer	2 – 3 weeks

Appendix 5: Patient saturation table

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Comments on selected questionnaires												
Comments applicable to both questionnaires												
Acceptable	X		X	X			X					
Appropriate lengths	X	X	X	X			X					X
Easy to understand	X	X	X			X	X		X			
Questionnaire preference		X							X			
Response options												
Enough response options						X						
Limited response options				X						X		
Scope	X	X					X		X			
IPOS Renal												
Comprehension												

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Interpretation of question 6		X					X					
Language is appropriate					X							
Similarity of Q2c and Q2i						X						
Straightforward questions		X										
Relevance of questions												
Question 6							X					
Question 8							X					
Question 9							X					
Specificity of questions												
Question on pain not specific					X	X	X					X
KDQOL-36												
Acceptability of question 35									X			
Comprehension of questions 9, 11, 29												

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Comparison of question 9 and 11		X										
Interpretation of question 11	X	X		X	X		X					
External influences on mood					X							
Interpretation of question 9	X	X										
Rephrasing of questions 9 and 11		X					X					
Response options				X								
Better recall period		X										
Potential barriers and facilitators of the adoption of ePROM system												
Potential barrier												
Patient related												
Burden of CKD					X							
Dislike of technology	X											

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Inconsistent responses					X							
Patient adherence	X				X							
Stable condition	X			X								
Questionnaire administration												
Frequent administration					X							
Questionnaire related												
Opinion it requires too much time										X		
Repetition of questions					X							X
Too many questions					X			X				
Potential facilitators												
ePROM design												
Anonymity					X							
Easy to complete						X						

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Electronically rather than paper			X									
Initial acknowledgement of completion				X								
Option to indicate no change	X											
PDF downloads	X											
Feedback												
Feedback on results				X						X		
Simple explanation of results				X								
Need for non-electronic options	X											
Prompt response				X					X			
Raise awareness of benefits										X		
Potential benefit of an ePROM system												
Assist with care												
Help doctors	X					X						

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Help raise alerts and flag issues								X				
Determinants of utility												
If actually patient completed	X											
Patient honesty	X											
Results need to be acted upon				X								
Improve efficiency												
Cut appointments	X	X										
Save paper	X										X	
Improve patient involvement												
Capture patients' feelings and opinions	X											
Self monitoring	X				X			X				
Indirect patient benefit												
Increase patient awareness of CKD	X											

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Quickly allay patient concerns										X		
Practical considerations for implementation												
Administration of questionnaire												
Frequency of administration												
Decision about frequency	X		X		X	X		X	X		X	
Preferred frequency of administration	X	X	X		X	X	X	X	X	X	X	
Mode of administration												
Awareness of myHealth			X	X				X				
Opinion of paper questionnaires	X				X							
Preferred mode of administration												
Electronic	X	X	X		X	X		X		X	X	
Paper	X						X		X			

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Reasons for preference	X	X	X		X				X	X		
Use of electronic devices and the internet												
Able to use the internet				X	X	X	X			X	X	X
Access to internet	X	X	X	X	X	X		X	X	X		
Electronic devices owned		X	X	X	X	X		X	X	X	X	X
Opinion of technology				X	X					X		
Personnel to receive results								X		X		
Feedback from clinicians												
Feedback preference	X	X	X	X			X	X		X		
Preferred mode for feedback	X	X	X	X								
Potential issues												
Health issues	X											
Management after completing an ePROMs		X		X								

Patient saturation table	Interview participants											
	1	2	3	4	5	6	7	8	9	10	11	12
Patient confidentiality		X		X	X							X
Reliability and validity of results												
Different results	X											
Influence of mood					X							
Influence of perception of CKD severity				X								
Keeping the doctor happy					X							
Partner's role					X							
Security issue		X			X			X				

Red crosses represent the initial occurrence of a code
Black crosses represent subsequent occurrences of the code

Appendix 6: Clinician saturation table

Clinician saturation table	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Concerns, potential barriers and facilitators															
Concerns															
Clinician-related															
Effect on appraisals			X						X						
Indirect influence on responses			X						X						
More work for clinicians		X				X					X	X	X		
Counter argument							X					X			
Responding to results	X		X	X	X		X	X					X	X	
Effect on clinical practice															
Change	X			X											

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Increase consultation time		X					X						X		
Legal and ethical risks	X						X								
Lots of referrals	X														
Medicalisation of normality	X							X			X				
Patient-related															
Feelings of obligation	X														
Hawthorn effect			X												
Influence of cultural beliefs						X									
Influence of family and friends		X				X									
May not mean what they say			X												
Patient adherence	X				X								X		
Patient burden	X	X		X		X	X								

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Patient expectations	X		X												
Patients' feeling about referrals									X						
Some patients may not be captured															
Elderly patients				X								X		X	
Language issues				X					X					X	
Literacy issues				X					X					X	
Patients with learning difficulties									X						
Questionnaire-related															
Patient input in development			X				X								
Questions might become stilted over time				X		X									
Threshold issues	X									X			X	X	

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Potential barriers															
Clinician-related															
Attitudes															
Alert fatigue	X		X												
Reason for many alerts			X												
Discomfort with non-physical complaints	X			X		X	X		X			X			
Nature of medical training			X	X			X		X			X			
Reluctance to refer outside the MDT															
Counter argument				X	X			X				X			
Reasons for reluctance															

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Not needed											X				
Clinicians cherry picking	X														
Difficulties with follow up after referral								X							
Difficulty knowing who to refer patients to								X							
Limited resources				X							X				
Patient access to services											X				
Limited benefit											X				
Protocol for junior clinicians									X						
Referrals to			X					X							

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
GPs might not be acted on															
Considered outside area of practice									X		X				
Scepticism and cynicism			X	X		X	X				X				
Perceptions															
Perceived inability to address non-medical issues			X	X	X				X		X				
Counter argument							X					X	X	X	X
Perception as time consuming				X		X	X		X		X			X	
Perception of PROMs as not individualised							X								
Conflicting agendas and expectations															

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Different expectations										X					
Doctor's agenda	X	X	X	X		X	X	X	X						
Patient's agenda	X							X		X		X			
ePROM-related															
Complex presentation	X		X		X			X	X				X	X	X
Information overload	X		X		X	X			X				X	X	X
Long questionnaires with many domains	X					X									
Numerous graphs	X														
Opening another window	X													X	
Numerical display		X								X				X	X
Pdf, Word files								X							X
Institutional															

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Infrastructure	X			X		X			X		X				
Lots of other surveys	X			X											
Separation of mental health services from physical health								X							
Incompatible with survival targets			X					X							
Time constraints															
Alternative view of time constraints							X								
Brief consultation	X			X		X	X		X						
Busy clinics	X	X	X	X			X	X			X				
Potential facilitators															
Clinician-related															
Buy-in	X														

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Consultant involvement in implementation	X														
Guidelines on interpretation of results										X					
Staff training	X			X				X				X			
ePROM-related															
Initial anecdotal evidence of value			X	X		X	X	X	X		X			X	
Subsequent scientific evidence															
Balance of cost and benefit											X				
Establish a link between QOL and adverse outcomes									X						
Improve patient outcomes			X			X		X	X		X		X		

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Predictive properties							X								
Reduced clinic visits			X		X										
Responsiveness			X				X	X						X	
Validity			X									X	X		
System design-related															
Alert in clinical portal	X							X		X					
Easy access to results	X			X					X	X		X		X	
ePROM needs to identify problem areas			X												
Mandatory acknowledgement of results								X							
Overall score with interpretations										X					
Reminders	X				X										

Clinician saturation table	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Suggestions for appropriate response			X												
Suggestions for probing highlighted issues			X												
Trigger points	X									X					
Very good signposting										X					
General opinions of PROMs															
Cautious optimism															
Fine balance	X					X								X	
Limited role							X		X	X	X		X		
Comparison with history taking				X			X		X					X	
Criticism of PROMs															
Does not take experience into account														X	

Clinician saturation table	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
PROMs do not capture the concept of QOL							X							X	
PROMs gained prominence as a result of issues with current clinical practice							X								
PROMs not really patient reported							X								
Interest in PROMs research	X		X		X		X		X				X		
Important role for PROMs in healthcare				X	X	X	X		X	X	X	X	X	X	X
Potential benefits and applications of an ePROM system															
Determinants of benefits															
Consultation style				X											
Patient expectations								X							
Purpose for use				X				X	X			X		X	

Clinician saturation table

	Interview participants													
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Type of patient		X	X						X	X			X		
Potential applications															
Evaluate interventions												X	X	X	
Provides trends	X	X						X	X						
Remote monitoring		X	X		X				X	X		X	X	X	
Indicator for drug compliance	X														
Use in clinic	X	X		X	X			X	X					X	
Use in MDT setting	X	X		X	X	X	X	X		X			X		X
Potential benefits															
Assist with care															
Assist with identifying psychological distress		X		X		X		X		X				X	
Early identification of signs and symptoms		X		X	X										

Clinician saturation table

	Interview participants													
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Identification of symptoms	X									X				X	
Improve coverage of health related issues	X	X						X				X		X	
Prioritize aspects of care	X					X				X					X
Provides a baseline	X														
Improve efficiency															
Assist with triaging	X				X					X					
Facilitate referrals						X	X	X							X
Focus clinical consultations		X	X					X	X						
Improve patient involvement and engagement															
Avenue for patient expression	X							X			X				
Helps patients reflect on their care	X					X		X							X

Clinician saturation table

	Interview participants													
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Involve patients in their care	X													X	X
Improve patient-clinician communication															
Bridge the gap between agendas								X		X					
Communicate clinician interest to patients	X						X							X	
Help clinicians understand what matters to patients										X					X
Promote discussions	X	X	X			X	X	X						X	
Understand expectations of patient cohorts										X					
Practical considerations for the implementation of ePROMs															
Administration of ePROMs															
Frequency of administration	X	X		X	X	X	X		X	X		X	X	X	X

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Mode of administration															
Patient ability to complete ePROMs	X	X				X						X		X	
Patient access and use of computers	X							X		X			X		
Preference for electronic over paper format										X				X	
Need for privacy		X							X						
Patients should be free to opt out								X							
Personnel to receive results		X	X	X		X	X			X			X	X	X
Timing of administration	X	X			X			X		X	X				
Feedback										X					
Interpretation of ePROM data															
Clinical meaning	X					X	X			X	X				

Clinician saturation table

	Interview participants														
	FG	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Weighing of symptoms	X														
Presentation of ePROM data															
Format															
Graphical display	X			X				X				X			X
Numerical values			X	X			X			X					
Pictorial format	X	X												X	
Summary with highlights			X			X		X	X				X		X
Tabular format							X								
Traffic light system	X	X	X	X				X		X			X	X	
Platform															
Clinical Portal and PICS				X	X	X	X					X		X	
PRO tab	X					X		X		X					
Email					X										X

FG - Focus group

Red crosses represent the initial occurrence of a code
Black crosses represent subsequent occurrences of the code

Appendix 7: Supplementary table of quotations

S3. Supplementary Quotations by Theme and sub-themes	
Sub-themes	**Illustrative Quotations**
	General opinion of PROMs
The role of PROMs in healthcare	"At the moment we see patients on a three monthly basis in clinic and simply look at the thing and say yes it looks fine based on our memory of what it looked like last time. So actually a reported patient outcome measure there would be better than what we're currently doing, would be an improvement on what we're doing." Doctor 4 "You know I think we are used to dealing with lots of numerical data, also very binary data, survival versus no survival. But you simply can't, you know the sort of data that reflects patient wellbeing and patient quality of life is not numerical in that way." Doctor 11 "But I think the data would also suggest that many aspects of kidney disease affect quality of life in a way that we can't measure on a blood test in hospital." Doctor 11 "If the clinician is worried that the questionnaire will open a big bag of worms that they're then going to have to deal with, I think that's something that we should be doing because actually." Doctor 6
Interest in PROMs research (clinician generated)	".....start piloting questionnaires or start using questionnaires or asking me to look at them and see what I.....I'll be like the altruistic patient, I'll help you knowing that it might not help the patient in front of me but it might help generations of patients and generations of doctors in the future." Doctor 2
Cautious optimism	"I think they're useful to get a greater understanding of what patients' expectations and feelings around situations are...but I think you've got to be very careful because I think they're very useful for broad sense and context and they're good for analysing population and bits and pieces." Doctor 9 "So you need to have that balance between what you can achieve and what you can't and communicate that to the patient..." Doctor 5 "Do you mean benefit to us as clinicians in their management? Yes. Does it help us manage them? Yeah. It might do, in certain circumstances, yes." Doctor 10 "Even within a CKD Clinic, them knowing that you recognise that life is a little difficult may open that debate. It would nice

	to think that they offered something more. I'm not sure that they will." Doctor 6
Comparison with history taking (clinician generated)	"As a clinical tool to detect people who have symptoms related to worsening kidney function, well one could say it's just a structured way of asking the same questions that you would ask in a consultation." Doctor 3 "I suppose that's why there are those of us, myself included but definitely others, whose focus has been to highlight what patients are concerned about; so not necessarily to administer pre-defined questionnaires but to highlight what troubles patients and more particularly, to highlight the burden of certain symptoms or problems that traditionally, we've not dealt with. You might argue that both paths serve the same function but I think that if you can uncover and make people aware that there are issues with patients that they may not necessarily admit to at first sight, or that you didn't realise there would be, but... 'Look, it's present in 20% of patients', that's kind of a way of educating ourselves." Doctor 6
	Potential benefits and applications of ePROMs
Potential benefits	"I'd do it on line preferably because it saves bloody loads of paper [laughter] but I don't know what everybody else's attitude is." Patient 1 "Yes, I think if you put these in front of patients, such as myself, it reminds me of the breadth of the discussion with regard to this particular disease." Patient 1 "Sometimes with kidney problems myself sometimes you start noticing little things and you start panicking or start thinking what's happening here? But if you have someone on board that can put your mind at ease it's much more easier than waiting maybe a month or two months for the next appointment to come through as well." Patient 10 "I think that's really good and it gives us a bit of a base line…" FG "My experience is that they do open lines of communication and that is what Clinical Care is all about. Because particularly Nephrology is so numbers based, the old joke in Nephrology is that you get the patient in, you tell them what their numbers are and you tell them to get out…." Doctor 6 "One can argue that actually we're probably quite bad at picking up psychological distress and we as doctors tend to ignore the psychological complaints of people's health and just worry about their physical symptoms…" Doctor 3 "Clinicians think and definitely sometimes patients think that all we're interested in is the kidney transplant and we're not interested in anything else. Even by using these questionnaires, it hints to people that we are interested in other stuff." Doctor 6 "And I can see what I think you're getting at and actually if every three months or every six weeks for example, they filled in a questionnaire and I got a flash on my screen saying so and so is doing really well then we can delay their coming in. The

	logistics will have to be thought through but I can see a point in the low clearance clinic. There are some very specific things that we're looking for. I'm going to use the word 'educated', I know that sounds patronising, patriarchal but informed of what to look out for and if they start getting these things, I want to know about it immediately." Doctor 2 "Okay, so would this help you then to perhaps steer people towards different Consultants or Doctors in the clinic if you'd seen this yourselves before the clinic…if you'd seen somebody was flagging up an emotional problem…?" FG "You can then at least have a look down what they're looking to get out of a consultation and say, 'Actually, I can't do very much about that. That looks like a deep-seated problem that you have that isn't related to this' and divert that through either back to their GP, or back to Psychiatric Care and that side of things, or a psychologist." Doctor 7 "Like a traffic light and then if clinic is running behind or I know the patient is likely to have said there is a problem with most things I can just focus on the red and ignore the amber or green whereas if I know the clinic is running slow for example I know this patient is not somebody that's very forthcoming with symptoms I might be able to go on the red and amber but yes it has got be a) time b) obvious." Doctor 2 "It's a reasonably good tool for a patient to also consider their own care and think about their care in a different way because I don't think they necessarily always think about why they're at the hospital." Doctor 7 "Often, patients have an agenda when they go into a clinic room and a consultant has an agenda when they go into a clinic room. Usually, you're somewhere in the middle and I think being able to utilise something where the patient has managed to verbalise what they want to talk about and how they're feeling, they may not necessarily talk to you about that, as a consultant, for a variety of different reasons. If it's used as a tool to be able to improve communication between the healthcare professional and the patient, and for both of them to get out of a consultation what their agenda is, then I think that's helpful." Doctor 7 "So I think having an understanding about what is important to them as part of their healthcare journey but also what is the most problematic thing and by the very nature of what we do they're quite complex patients you see, reality is in most clinic visits you're going to be able to address one or maybe two issues at most and I think these surveys and questionnaires can help to try and target how you do that." Doctor 9 "He's been telling the nurses, been telling the carers, I can tell that there is something wrong but he or she won't divulge anything to me in which case a questionnaire could prompt or focus the conversation rather than me repeatedly asking are you sure everything is okay, sure everything is okay and I keep getting fine, fine so I can see that really useful in that extreme." Doctor 2 "Maybe it could help to involve the patient in their care and make them feel that they're being given the time to express their experience and their feelings and concerns in a way that they may not feel they have time to do in a brief outpatient's consultation." FG

Determinants of benefits	"Well it would partly depend on the circumstances in which this was going to be used." Doctor 8 "I think it probably will but it depends on what the patient wants to get out of consultation and what's needed. If the patient's expectations are that they're going to go into a clinic and you're going to sort out all of their problems, maybe even mental health problems as well, then in some respects, using a PROM to be able to say, 'Actually, that's not really my area. That needs to be dealt with in this other place', then that's probably an appropriate thing too." Doctor 7 "It can be a double edged sword you see you're asking me vague general questions and I can try and answer for 50, 60% of my patients for example. But trouble is in my head I keep getting one extreme coming in or the other extreme coming in. And a face flashes in front of me that says you do not really want to ask that patient what their symptoms are. And on the other hand I get another flash saying I wish I could get into that person's head to find out what is going on that I can help with." Doctor 2 "I think so, especially if they were going to do this on their own and not in the presence of the doctor, so if you've either got the time, you know they can spend as much time as they need doing it and they don't feel any sort of pressure in only answering the questions in the way that, you know, only letting the doctor hear what the doctor wants to hear." Doctor 8
Potential applications	"Yeah but then also might be really interesting is the time around starting renal replacement therapy and I don't know whether these would be able to compare from before and after starting…" Doctor 1 "I think the questionnaires might be useful to identify people who need assessment but I think it needs to be backed up by a second layer before you just drag them to a clinic because actually the risk is you just delay their therapy." Doctor 9 "I think these questions will come into effect with patient compliance, if the patient just feels like this, then there's a lot of work to put in to convince them to keep to their medication, their appointments and, because you get the odd ones that…so a lot of input, this patient was yes, yes, to every question here, so I think it's, I think it's relevant." FG "So it could be enhanced surveillance you could use it for, you could use it for overall surveillance, I'm not sure that's quite as useful. And even after doing a procedure we could actually have it as a surveillance thing so you could use this for something like their first month of dialysis and whilst it's not necessarily patient reported if the dialysis nurse you know is proxy reported it's still PROM, proxy reported outcome measure." Doctor 4 "For instance, if someone was anaemic, they might flag us as being short of breath or having occasional chest pain, lack of libido, lack of energy, so those things flagging up could trigger an alert to the anaemia team to assess certain bloods and things like that in clinic." FG "It would help with some sort of collaborative report wouldn't it, because some of the; I was thinking about the questions on the back, you might not, you might say that that's a real problem but you don't particularly want to talk about it with the

	Consultant. It would perhaps form part of the Forward Look MDT wouldn't it which is already quite a long process…" FG
	Practical considerations
Administration	"If they were looking for a timescale and measuring how long the progression of a disease was, I can appreciate that would be useful to them but if it is to be useful to them, the patient (or the person who's got cancer) would like to be prompted as to how often. They may do it every other day which would be embarrassing to you and every ten years would be useless to you. But you're the guys who know what the likely progression of the disease is." Patient 1 "How often? I suppose it's all about how often they want the information. If it's like once a fortnight or once a month or whatever you don't mind just doing it for ten minutes do you, it's easy, it's quite easy isn't it?" Patient 6 "I think monthly would be okay. I mean I won't want to be doing this every week. Because with people working and still managing their CKD, you do not have time to do this too often." Patient 3 "I wouldn't have a problem with filling it in twice a year, as long as there was an option of 'No Change' which would save filling it in." Patient 1 "I'd fill it up on the internet, as long as you gave me the url to the website. I'd fill it in there. If paper arrived on my doorstep, I'd probably fill that in. If there was a choice that I received by email, I'd most likely download it and fill it in (for my reference) and print a page out. It really wouldn't make any difference." Patient 1 "Smartphone. I think it will do because I think these questionnaires are very similar to the ones that you can do on smartphones anyway at the moment so, it's easier to do and I think it's more convenient for people to sit there and just maybe do it instead of being in front of a big screen at times as well." Patient 10 "Well I've always got me iPad with me so I can make the, you know, the page bigger or smaller or whatever it takes but I always, I do all my work more or less off my iPad." Patient 2 "…if I request my wife to do it with me you'll have two people doing it. And sometimes what happens it can be, something may crop up which I don't want to tell her because it will make her upset you see so filling this form up and using the smartphone." Patient 5 "I would be quite happy for them to complete this before every clinic visit actually, you know, if it does prove to really provide a focus for what is important to the patient and the particular symptoms they are experiencing I would be quite happy if they completed this before every clinic visit and like I said hopefully that would make the clinic visits more efficient actually and actually just make the clinic visits more satisfactory for the patient so hopefully they get more out of it than they would have if they hadn't completed this questionnaire." Doctor 8 "Yeah I mean I guess it might depend on their trajectory of their disease. If they've got very stable CKD they're maybe

	unlikely to change but I would have thought, I mean in CKD clinic for example we would see people on a two to three monthly basis in local clinic and that seems like a reasonable length of time between questionnaires." Doctor 11 "Again those who haven't got the capacity to answer the questions so they will be relying on other people and again that might be tricky so the people might not be able to answer these for themselves it would be some other people answering for them which may not be accurate so it's going to be tricky." Doctor 1 "Hadn't thought of that but that's actually important as well because it feeds into this conversation of what they tell the nurse compared to what they tell me. Yeah that's a good point because just thinking out loud here, part of the problem with that patients' care is me. If the patient doesn't like me, thinks I'm too dogmatic and thinks I don't pay enough concern, they're never going to fill in a form to say that if they know the form is coming to me. If we anonymised the form it makes it a bit of a waste of time because I can't help…" Doctor 2 "Arguably everybody involved. So the patients to see a summary of their responses and I think it might be important to the patients to see whether their answers had changed over time so that might help some people. And then doctors, other people in the MDT who were looking after them in that context, clinic system or whatever should see it." Doctor 3 "I think that if you've got questions like that, then there does need to be a defined portal whereby someone gets to know about it. Again, it doesn't have to be a doctor or a nurse but it does have to get to someone. My understanding of questionnaires is that you're not going to immediately act on any of the results, so it can just come back in a standardised way." Doctor 6 "So if you feel that you're less well of your disease is more active, maybe you should think about filling in this questionnaire because it could be a good way to try and flag up exactly why you're feeling non-specifically unwell because often people with kidney disease do feel non-specifically unwell." Doctor 9 "Because they are quite busy and actually I think, I mean, my concern about all these things is the practicalities of them you know. I mean I've got a clinic this afternoon. I've got 25 patients coming to it. I might have a registrar, well hopefully I'll have a registrar who will see about 10 of those so I'll end up seeing the other 15. If I were to have a load of patient self-reported forms to look at before or during the clinic or if patients were trying to do it during the clinic it would be quite impractical actually." Doctor 10 "Certainly, the cultural beliefs also do have an effect and most of the ethnic population do have very strong family ties and all of that will influence how many responses you get for your questionnaires. When you're talking with families in the clinic and patients, when they are using translators, you do know that there are things that can be translated a little bit more loosely or not, depending on their person's own beliefs." Doctor 5
Clinical interpretation	"For instance, if someone was anaemic, they might flag us as being short of breath or having occasional chest pain, lack of

(Clinician)	libido, lack of energy, so those things flagging up could trigger an alert to the anaemia team to assess certain bloods." FG "I mean obviously for specific symptoms like, you know, if patients say well I don't have any energy, I feel tired all the time, a very common complaint, so you look for the physiological explanations for that like anaemia for example so that you can identify if they're anaemic then you can treat that. But very often they are not actually and you know that's a complex psychological symptom really, fatigue, and once you've ruled out basic physiological explanations for it you are left with something for which there is no easy medical intervention and patients have to just learn to cope with that on their own." Doctor 10 "So I mean I think presumably it's going to generate an overall score with some sort of interpretation at the bottom." Doctor 9
Presentation (Clinician)	"On the computer screen in graphical form." Doctor 11 "How do you then flag those up through the portal onto the clinicians and I think it probably needs almost like traffic light system, red, this needs to be reviewed, amber, there is a bit of cause for concern or green, yeah, and you could either do that through a tab which lights up at the top." Doctor 9 "In principle I think that could have value, the obstacles that I would say exist at the moment are that doctors use Clinical Portal but they don't use My Health so you would have to populate it into Portal." Doctor 3
Feedback	"It's useful to know about it. From a patient point of view, it's interesting and useful to know about it because you want to get on with your life and get as much done as possible before what you've got actually affects your ability to do it. That's useful to a patient as well as to you guys." Patient 1 "Yes. I would like to have some feedback. It will be nice to know what the doctor thinks. It will also be good to see, I know data protection and all, but to see what type of results other patients are getting..." Patient 3
	Concerns, potential barriers and facilitators
Concerns	"The only thing that I think that is possible could be an issue is security, patient confidentiality." Patient 2 "I think the only thing would have to be that if you're going to use something like this, you have to have confidence that there's going to be a response to it." Patient 4 "I think with all these questionnaires, you're limited. Sometimes, the answer isn't exactly what you want to give…." (Patient 4) "So every month somebody looks at my haemoglobins to make sure they are high enough. They look at my blood sugar

	control to make sure it's good. They look at my calcium phosphate to say they are more than the targets set by learned societies. Are my patients dialysing for long enough? How many of them are dialysing through a fistula rather than a line? So a patient might say to me "I want to dialyse for two hours a day. I don't want to use a fistula and I'll miss the occasional day if I want to." All of which would improve their quality of life in the short term no doubt. In the long term all the evidence tells me they're going to die a lot sooner and their quality of life will become impaired in the future. If I agree to what they're saying then the next time they come they tick the 'I'm very happy' in the satisfaction questionnaire, I'll get a phone call from my boss saying your unit has got the worst results…" Doctor 2 "I suppose that's a really good point and not one I'd thought about, and it would certainly make you anxious, because it's true, if you've asked a question and had a reply and then not done anything about it, that's much worse than just having not really being aware of it in the first place…" FG "I think you are at risk of medicalising normality actually. Because everyone has those issues, haven't they actually? You know if someone's got health problems, I don't know, let's say they've got renal failure and they're upset as a consequence or have some psychological issues as a consequence, I mean that's a normal response isn't it to that? I mean I'm not sure it's a pathology as such that requires you know assessment by another specialist and you know assessment by a psychologist." Doctor 10 "…because if the patient thinks that they've filled in all in and everybody's going to, oh we're all going to be looking at every single detail of it but that's not what's happening then…" FG "There are a lot of things that you can ask patients and the KDQOL-36 questionnaire - typical examples, it's all very well asking the question but then this could come back to what can you actually do about it?" Doctor 2 "…so it could change our practice in a way that has an unintended consequence." FG "I keep trying to tell my students about the Hawthorne Principle, that just the mere act of looking at something changes it...as soon as you observe something it automatically changes nature so by the fact of asking the question you have already changed what a patient is going to say. Because a) they might never have thought of that as a problem and b) they may never have thought that that potentially was a solution to that problem." Doctor 2 "…but I just feel like it's another thing for them to be doing. I mean a lot would appreciate doing something like that and I think they can put their feelings onto something like that, but I don't know, I just think they're being bombarded with things every day." FG
Potential barriers	"Yeah it will, if there's too many questions and too many similar types of questions. That will discourage me because I will say ah I've just told them this thing and they've repeated the question on the next thing and it's a similar question which they've asked me so what should I do." Patient 5

	"Yeah. You'd only do it if something had changed. If you felt as though you didn't have as much energy, it could be other things; not just CKD, although, it probably would be as a result of CKD but you couldn't actually be positive about it." Patient 1 "Oh there is very real alert fatigue where people ignore alerts most time because most alerts are a complete waste of time and second because there are just far too many of them and far too many about people that I am barely peripherally involved with." Doctor 2 "I think you might have to do some training with staff though on how to respond to that, because there's, from the psychology side, I could, there's a few Consultants that we don't get a single referral from, so I can only assume that they are either cherry picking or they just go down that emotional route at all." FG "So we have the Renal Psychologist which is a fairly straightforward thing and I send quite a lot of my patients to them. Brilliant. I've got a referral form. Great. Psychiatry isn't in the QE and so knowing where to refer that person to and how is difficult. That's probably my biggest barrier because people come from all sorts of different places and they may be in a different catchment area for a different service. I would usually then refer them back to their GP to make that referral because I don't know where to refer them to." Doctor 7 "I think the perception that it will increase time required for consultation would be a significant risk and then just scepticism about the benefit of using something like this if you could just say.....well my perception of just asking somebody two or three questions that I believe are the things that I most need to know is going to be quicker than your system. So why should I change? I think that's the first fundamental thinking that you are up against a bit." Doctor 3 "I think first of all, obviously, in the absence of any of that evidence that I've just talked about, if that evidence wasn't there people might feel that it might be a time-consuming thing to do with little benefit gained from it." Doctor 8 "Often, patients have an agenda when they go into a clinic room and a consultant has an agenda when they go into a clinic room. The time constraint within a ten minute consultation of having lots and lots of patients going through a clinic; if you know that you've only got ten or 15 minutes, you're not necessarily letting somebody tell you all their woes. You kind of want to tick your boxes of knowing the specific things that you want." Doctor 7 "One of the major pressures in our clinical practice is time and these are patients that we see in clinic whereby a clinic appointment is 15 minutes but actually those 15 minutes, part of that is taken up in other things such as dictating letters, looking at results, etc, the actual time spent with a patient is probably ten minutes or less so really you know you need to make sure that this is presented in a way that is useful, bearing in mind the pressures with regards to time that we are under. If it wasn't then I think that would be a major resistance to have been taken up in clinical practice." Doctor 8 "It does but in a questionnaire form, you get bogged down by the details of the questionnaire. The lengthier it is, the more people disengage. If it's too lengthy. It comes back to that. It really comes back to that. It is important. If it's too lengthy

	[laughter].” Doctor 5 “When you’re in clinic, the first thing you open is portal and then from portal you go into pics and those are the only things that you look at in clinic, anything that involves opening another window or going online you just, people aren’t you know, they just won’t do it.” FG “…..having a non-discriminatory form. So by that I mean if every patient clicks everything is terrible then I see little point in that or if every patient ticks everything is fine I will see little point in that.” Doctor 2
Potential facilitators	“So why can’t we have a form like this, without no name, nothing at all, no doctor’s name, no my name and just fill it up and send it to you? And that way you’ve got confidentiality to start off, I don’t know which doctor I’m talking to start off, it can be A, B, C and the doctor doesn’t know who the patient is and your research goes on.” Patient 5 “Yeah, why would I bother if I don’t get anything from it? It needs to be a two-way thing. I appreciate why it might be helping you but, at the same time, I think I have to feel that I’m getting something from it, to take the time to do it.” Patient 4 “I mean you can’t get away from blood tests and data, that is what we do and that is how we treat disease. But I think increasingly the patient is becoming a client rather than just being a patient and much more attention is being paid to more holistic aspects of their care. And teaching students that from an early stage can’t be a bad thing.” Doctor 11 “The time before a doctor interrupts a patient is astonishingly limited actually and not in a good way. Doctors do interrupt really quickly so it would be quite a lot of discipline actually. You could maybe train people to do it maybe again you could train other non-medical healthcare professionals to do it more effectively…” Doctor 3 “I think trying to set some sort of absolute feedback for the results, the clinician so they know where they are is probably going to be useful.” Doctor 9 “So if they can answer this questionnaire on their own without anybody around, so they can answer it truthfully, that is another important factor about completing a questionnaire I think.” Doctor 1 “I mean the main consideration is the balance of the benefit of it and the cost and of course the cost of that is the clinician’s time and the patient’s time.” Doctor 10 “First of all having some evidence that they are useful. One would be that evidence that these in some way are a link to adverse outcomes for example…acting on this improves outcomes so there are two different forms of evidence but I think if there was good quality evidence that that was the case then I think that would be the strongest factor in physicians then taking this up and actually using them in their clinical practice.” Doctor 8 “I think one would need some sort of validation of the method that’s being used as well.” Doctor 11 “There will always be a period of 'Oh, another paperwork!' and there’s that side that comes in and another set of things to

	do. If one gets passed that, it would be alright and I suspect it is more of information and evidence that you need to present that gets people going towards it." Doctor 5 "Yeah, and you have to acknowledge you've seen it, so that they're not just filling them in and nobody's looking at them and it's a waste of time." Doctor 7 "And it may be the patient needs a reminder, you know. So if they're, if we know a dialysis patient and they're looking up their results it flags up or there's an email or a message that comes through, please send your return in today." Doctor 4 "In time I would like to see some proper scientific evidence actually this has made a difference to either patient's outcomes, patient's wellbeing, maybe even reduce the number of clinic visits but yes at the beginning the first ones might be burdensome but if actually somebody is cured or made better they might not need to come back as often." Doctor 2 "It's got to be sensitive as well so that I don't falsely reassure myself that the patient has ticked everything is clear whereas in fact they've got an axe in the back of their head because there isn't a "I've got a headache" tick box." Doctor 2

FG – Focus group participant

Appendix 8: Comments on selected questionnaires

Patients were asked for their opinions on the Kidney Disease Quality of Life-36 (KDQOL-36) and the Integrated Patient Outcome Scale-Renal (IPOS-Renal) questionnaires which were selected as potential questionnaires to adapt for the ePROM system based on an earlier systematic review.[1]

Comments applicable to both questionnaires

Patients were of the opinion that both questionnaires were clear and easy to understand, and of reasonable lengths despite being comprehensive. Some patients however pointed out that questionnaires by design have predetermined response options which may not always be applicable to the individual. They also felt that the amount of information patients can provide is restricted. They suggested the inclusion of additional comment boxes. A clinician approved of the fact that the questionnaires had questions assessing patient's psychological/emotional wellbeing which the individual felt was an uncomfortable aspect of health to discuss with patients.

Specific comments on the KDQOL-36

Although patients and clinicians perceived the KDQOL-36 to be more comprehensive than the IPOS-Renal, some clinicians felt that it was too long to use in routine clinical settings. Patients and clinicians considered the relevance and meaning of the individual questions. Clinicians felt that some questions may not be culturally relevant to their patients. For example, the question about the degree to which a patient's health limits their ability to bowl. The clinician suggested that the questionnaire was probably referring to tenpin bowling (which is more popular in America) and not crown green bowling which patients from Birmingham, UK might

be more familiar with. Some patients also questioned the meaning and relevance of question 11 which asks if patients have felt ‘downhearted and blue’ in the past four weeks. Again this was seen as an American expression which some British patients might not be familiar with. A patient felt that question 35 which inquires about the patient’s sex life might be embarrassing for older men and deemed too personal in some cultures. The individual suggested changing the question to ‘Do you have any problems in your personal life?’

Specific comments on the IPOS-Renal

Participants perceived the measure to be succinct and on the whole, easy to understand. However, some patients felt some of its questions were not specific enough. For example, the question on pain did not specify the location of the pain. They also questioned the meaning and relevance of question 6 ‘Have you felt at peace over the last week?’

Preference

Participants weighed the pros and cons of both questionnaires and a few decided that they preferred the KDQOL-36 despite its longer length because it covers more domains, therefore potentially providing more information. As it assesses patient’s health over a longer time period of 4 weeks, a patient believed that it would provide a better overall picture of patients’ health status.

Reference

1. Aiyegbusi OL, Kyte D, Cockwell P, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PloS one 2017;**12**(6):e0179733.

Appendix 9: Illustrative quotes for the selected questionnaires

	Illustrative quotes
Comments applicable to both questionnaires	"I think questions about mood are also important.....Just a cultural thing. I think it is the sort of doctor/patient relationship for a lot of us it is probably easier to ask about physical health than to ask about physiological wellbeing but that's not to say we shouldn't be doing it. I think it is useful when it's in there." Doctor 3 "No. I mean the questions were not really personal." Patient 3 "They're a good length, not too long…" Patient 7 "I felt they were clear and easy to understand. It didn't even appear as are there any trick questions where a lot of these questionnaires do have or, and even you can see it. You know, you think hang on a minute, 'I've already answered that in a different way', you know." Patient 9 "On the length? Well, it's not burdensome to read." Patient 1 "Yeah, no problem at all in understanding the questions and even getting some insight as to what you're aiming at." Patient 1 "If it's like once a fortnight or once a month or whatever you don't mind just doing it for ten minutes do you, it's easy, it's quite easy isn't it?" Patient 6 "The only issues I would say is probably sometimes when you give an answer you may cross your wires so you might be trying to explain something but electric questionnaire might not be able to give you the options so sometimes you might need to have an additional comment box." Patient 10 "Yes, I do. I think it's comprehensive. I can't think of anything else that I would think was pertinent to the disease I apparently have that you haven't addressed." Patient 1 "Yes I think it does, most of the things on there are things that you do, that do affect you when you have CKD." Patient 9
Specific comments	"Bowling or playing golf, this sounds like an American form doesn't it? Because I'm sure they are referring to ten pin

on KDQOL-36	bowling, not crown green bowling, which is probably not popular in Birmingham…climbing several flights of stairs… have you felt calm and peaceful? So, ok, so that's all very useful, yeah." Doctor 10 "Yeah it's definitely more symptoms especially are covered by it and I think these are important aspects to ask about which the IPOS questionnaire doesn't include in there, a lot of the renal related are included and a few more even in here as well which are quite important I think." Doctor 1 "36 questions is a lot so again if you wanted to ask all this repeatedly I think they would fatigue from doing so. I suspect that completion rate for a 36 item questionnaire would be quite low in the first place. A lot of people would give up before they get to the end. Repeating this every time someone came to clinic that would be too much. I don't think you can do 36 questions at a clinic visit. I think it's the sort of thing some people you put them in clinical trials and they will do this sort of thing but to generalise to the patient or cohort you see in the clinic then it's clearly not the same sort of phenomenon." Doctor 3 "I feel it should be, 'Do you have any problems in your personal life?'. I think men in particular would have a difficulty in speaking, maybe to another man they might. I don't know as I'm not a man but I don't know what the age, I think it depends on the age. Again, I think embarrassment. I think it's still a no-go subject with a lot of people and a lot of cultures, you know, that would not be willing to talk about it or answer questions on it (Question 35 about sex)." Patient 9 "Question 11 sort of mirror imaging number nine." Patient 2 "Slightly bothered or not bothered. I don't really actually know how I feel about it. I think that's what I'm trying to say. I think 'bothered' is not the right word. I don't know why I feel that 'bothered' is not the right word. 'I'm not at all bothered' – I don't know why. It's a personal thing *[laughter]*." Patient 4 "I think there may be some that are fairly similar and some ask some things that people just won't relate to. Playing golf and no they will do actually, this is a very wide spectrum of activities and you've got some males and your females doing all sorts of things at home and I suppose that's quite a broad question but the kind of activities that those are, they've kind of got a spectrum of how you physically will be able to accomplish those things so going out golfing compared to doing vacuuming at home I think there are quite different levels of exercise so…" CC1 "So I think it's nice that it is quite well laid out. I like the fact that it has got a bit of a front sheet telling you a little bit about it. It's nice and straightforward and quite concise. I think that's helpful for patients. A lot of times patient information is far too detailed. The sort of stuff people will never read. But that is quite short. I quite like the look of

	that, it's got nice subheadings, I think that would be reassuring if I were a patient. The survey itself I think again I quite like the way it's laid out, five or six pages long, the text is not too dense…" Doctor 3 "Trying and doing the questionnaire on your own so you can answer truthfully because there are a few personal questions in there. I don't think they are actually unacceptable, I think there are things that should be asked and may not be asked, things like they're maybe thinking and they don't feel comfortable asking, they are being asked in this question. So for a patient to be able to let them know but not actually to say out loud especially if their family are with them." Doctor 1 "I think they're quite long. I think they probably need to be, by the nature of them." Doctor 9 "If this was something that was going to be done in the patient's own time I think that would be perfectly reasonable, it doesn't seem excessively long, having said that, if this was something that I would need to do in clinic with somebody then whether or not I would have the time to do all of this is a different issue." Doctor 8 "So I mean those patients who are at college, university, work a lot of these questions... Does the disease impact their work? and I think that's an important question to ask and if they worry about the disease because again some people are scared to get a relapse and actually worry about it a lot. And some people don't worry about it because they've been so far so long time spent being well that they don't worry about it. It is an important question and that would be a nice thing to ask those people because if they're spending a long time worrying and actually they're doing really, really well they don't need to worry and you can spend that time reassuring them." Doctor 1 "Accomplishing less than you would like or limiting the kind of work or other activities I think those are fairly similar questions but just asked in a different way and I think trying to mm…" Doctor 1 "My age group, sex was never mentioned in or out of the home. It just wasn't a subject, you know, when you were young. I don't think people would answer it. I think they, I wouldn't say they'd be embarrassed by it but I think they will skip it, you know, or lie." Patient 9 "What does that mean, downhearted and blue?" Patient 7 "It's fairly obvious to me and I imagine it probably would be to most people. Everybody knows what downhearted means. Blue? You've qualified it with 'downhearted' [laughter] (Question 11)." Patient 1

Specific comments on IPOS-Renal	"I think the IPOS questionnaire is quite easy to understand, is very nicely worded in terms of especially the first side, very simplistic, yeah." Doctor 1 "This questionnaire seems very straightforward to me, I would have thought most patients assuming they can speak English will be able to answer those questions." Doctor 3 "Oh very brief, very succinct yes so the first one IPOS renal patient is obviously predominantly symptom based and then some questions about mood really which I think are all perfectly reasonable questions to ask people with chronic kidney disease and I would regard many of these symptoms as quite advanced so things like vomiting, if someone had chronic kidney disease and we were pre-dialysis one would have hoped they would have initiated dialysis before symptoms like that had become manifest as it were…" Doctor 3 "The wording's okay, the wording's fine yeah pain, shortness of breath, weakness or lack of energy, obviously vomiting, poor appetite, constipation, sore or dry mouth, drowsiness, poor mobility, itching, language no problem, it's okay." Patient 5 "Not all of these symptoms are obviously per se going to be related to renal dysfunction. Things like poor mobility and constipation I wouldn't regard particularly as symptoms of advanced kidney disease or uraemia. They obviously probably driven by co morbidity, osteoarthritis or drug induced constipation or lots of other things couldn't they?" Doctor 3 "They should specify which, they can't just put pain." Patient 7 "Have you felt at peace? How it is connected with kidney? I don't understand meaning of that (Question 6)." Patient 7 "Doesn't explain very well what it means by 'felt at peace' (Question 6)." Patient 2 "Yeah, I think it was lack of energy and feeling drowsy. Yeah, that's near enough the same question really." Patient 6 "Pain related to what? That's what I was trying to ask." Patient 6
Preference	"At the end of the day I think you want to find out in detail how they are coping so I'd probably go for the longer questionnaire because it does cover the detail that you want and if you are doing this less frequently this detail would be quite important so you have more areas that you can explore and see if you can help." Doctor 1

	“I think the plus side obviously from the week one is you’re more likely to remember what you’ve done over that last week or what’s happened over that last week but to give an overall picture it needs, it would need to be longer. That’s why I would probably go with the four week one.” Patient 2

Appendix 10: HRA approval letter

NHS
Health Research Authority

Dr Olalekan Lee Aiyegbusi
Centre for Patient Reported Outcomes Research,
Institute of Applied Health Research
College of Medical and Dental Sciences
University of Birmingham, Edgbaston. Birmingham
B15 2TT

Email: hra.approval@nhs.net

24 February 2017

Dear Dr Aiyegbusi

Letter of HRA Approval

Study title:	**Using Patient-Reported Outcome measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney Disease. (PRO-trACK Project)**
IRAS project ID:	**200657**
REC reference:	**17/WM/0010**
Sponsor	**University of Birmingham**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.

It is critical that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details and further information about working with the research management function for each organisation can be accessed from www.hra.nhs.uk/hra-approval.

Appendices

The HRA Approval letter contains the following appendices:

- A – List of documents reviewed during HRA assessment
- B – Summary of HRA assessment

After HRA Approval

The document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

In addition to the guidance in the above, please note the following:

- HRA Approval applies for the duration of your REC favourable opinion, unless otherwise notified in writing by the HRA.
- Substantial amendments should be submitted directly to the Research Ethics Committee, as detailed in the *After Ethical Review* document. Non-substantial amendments should be submitted for review by the HRA using the form provided on the HRA website, and emailed to hra.amendments@nhs.net.
- The HRA will categorise amendments (substantial and non-substantial) and issue confirmation of continued HRA Approval. Further details can be found on the HRA website.

Scope

HRA Approval provides an approval for research involving patients or staff in NHS organisations in England.

If your study involves NHS organisations in other countries in the UK, please contact the relevant national coordinating functions for support and advice. Further information can be found at http://www.hra.nhs.uk/resources/applying-for-reviews/nhs-hsc-rd-review/.

If there are participating non-NHS organisations, local agreement should be obtained in accordance with the procedures of the local participating non-NHS organisation.

IRAS project ID	200657

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please email the HRA at hra.approval@nhs.net. Additionally, one of our staff would be happy to call and discuss your experience of HRA Approval.

HRA Training

We are pleased to welcome researchers and research management staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

Your IRAS project ID is **200657**. Please quote this on all correspondence.

Yours sincerely

Beverley Mashegede
Assessor

Email: hra.approval@nhs.net

Copy to: Dr Sean Jennings, Sponsor Contact

Dr Chris Counsell, Lead NHS R&D Contact

NIHR CRN Portfolio Applications Team

Appendix A - List of Documents

The final document set assessed and approved by HRA Approval is listed below.

Document	*Version*	*Date*
Contract/Study Agreement [Statement of Activities]		24 February 2017
Covering letter on headed paper [cover letter]		15 December 2016
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Confirmation of insurance]		13 December 2016
Interview schedules or topic guides for participants [Topic guide - Patient interviews]	1.0	27 November 2016
Interview schedules or topic guides for participants [Topic guide - focus group]	1.0	27 November 2016
IRAS Application Form [IRAS_Form_15122016]		15 December 2016
IRAS Application Form XML file [IRAS_Form_03012017]		03 January 2017
IRAS Checklist XML [Checklist_01022017]		01 February 2017
Letter from funder [PhD award]		20 July 2015
Letter from sponsor [Confirmation letter of sponsorship]		13 December 2016
Letters of invitation to participant [Patient invitation letter]	2.0	29 January 2017
Other [GCP Certificate]	1.0	
Other [Transcription service agreement 1]		
Other [Transcription service agreement 2]		
Other [Transcription service agreement 3]		
Participant consent form [Consent Form - interview]	2.0	29 January 2017
Participant consent form [Consent form - Usability testing]	2.0	29 January 2017
Participant consent form [Consent form - Validation study]	2.0	29 January 2017
Participant consent form [Consent form - Focus group]	2.0	29 January 2017
Participant information sheet (PIS) [Interviews]	3.0	29 January 2017
Participant information sheet (PIS) [Usability testing]	3.0	29 January 2017
Participant information sheet (PIS) [Validation study]	3.0	29 January 2017
Participant information sheet (PIS) [Clinician and other staff – Focus Groups]	3.0	29 January 2017
Research protocol or project proposal [Pro-trACK Protocol]	1.0	25 November 2016
Response to Request for Further Information [Response to Validation Queries]		03 January 2017
Summary CV for Chief Investigator (CI) [Curriculum Vitae Lee]		01 November 2016
Summary CV for student [Curriculum Vitae Lee]		01 November 2016
Summary CV for supervisor (student research) [Prof Calvert CV]		
Summary CV for supervisor (student research) [Derek Kyte CV]		
Summary CV for supervisor (student research) [Tom Marshall]		
Summary CV for supervisor (student research) [Paul Cockwell CV]		
Summary, synopsis or diagram (flowchart) of protocol in non-technical language [PRO-trACK lay summary]	1.0	25 November 2016
Validated questionnaire [EQ-5D-5L]		
Validated questionnaire [IPOS-Renal]		
Validated questionnaire [KDQOL-36]		
Validated questionnaire [KDQOL-SF]		

Appendix B - Summary of HRA Assessment

This appendix provides assurance to you, the sponsor and the NHS in England that the study, as reviewed for HRA Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England to assist in assessing and arranging capacity and capability.

For information on how the sponsor should be working with participating NHS organisations in England, please refer to the, *participating NHS organisations, capacity and capability* and *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* sections in this appendix.

The following person is the sponsor contact for the purpose of addressing participating organisation questions relating to the study:

Name: Sean Jennings

Tel: 01214158011

Email: researchgovernance@contacts.bham.ac.uk

HRA assessment criteria

Section	HRA Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	No comments
2.1	Participant information/consent documents and consent process	Yes	All participant information sheets have been updated to bring them in line with HRA Approval amendments via a minor amendment.
3.1	Protocol assessment	Yes	No comments
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	The sponsor intends to use a Statement of Activities as the form of agreement with the participating NHS organisation. The transcription service will be provided by an external company and the confidentiality agreements that will be used have been submitted.
4.2	Insurance/indemnity	Yes	Where applicable, independent contractors (e.g. General Practitioners)

Section	HRA Assessment Criteria	Compliant with Standards	Comments
	arrangements assessed		should ensure that the professional indemnity provided by their medical defence organisation covers the activities expected of them for this research study.
4.3	Financial arrangements assessed	Yes	No funds will be provided to the participating organisation to support this study.
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	No comments
5.3	Compliance with any applicable laws or regulations	Yes	No comments
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	Provisional Opinion issued 26 January 2017. Further Information Favourable Opinion issued 13 February 2017.
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	No comments
6.3	Devices – MHRA notice of no objection received	Not Applicable	No comments
6.4	Other regulatory approvals and authorisations received	Not Applicable	No comments

Participating NHS Organisations in England

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.
This is a non-commercial student (PhD) study and there is one site type. The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the r esearch management function at the participating organisation. For NIHR CRN Portfolio studies, the Local LCRN contact should also be copied into this correspondence. For further guidance on working with participating NHS organisations please see the HRA website. If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England which are not provided in IRAS or on the HRA website, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net. The HRA will work with these organisations to achieve a consistent approach to information provision.

Confirmation of Capacity and Capability

This describes whether formal confirmation of capacity and capability is expected from participating NHS organisations in England.
Participating NHS organisations in England **will be expected to formally confirm their capacity and capability to host this research**. • Following issue of this letter, participating NHS organisations in England may now confirm to the sponsor their capacity and capability to host this research, when ready to do so. How capacity and capacity will be confirmed is detailed in the *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* section of this appendix. • The Assessing, Arranging, and Confirming document on the HRA website provides further information for the sponsor and NHS organisations on assessing, arranging and confirming capacity and capability.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and the minimum expectations for education, training and experience that PIs should meet (where applicable).
A Local Collaborator is expected at the participating organisation. GCP training is not a generic training expectation, in line with the HRA statement on training expectations.

HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken
Where arrangements are not already in place, network staff (or similar) undertaking any research activities that may impact on the quality of care of the participant, would be expected to obtain an honorary research contract from one NHS organisation (if university employed), followed by Letters of Access for subsequent organisations. This would be on the basis of a Research Passport (if university employed) or an NHS to NHS confirmation of pre-engagement checks letter (if NHS employed). These should confirm enhanced DBS checks, including appropriate barred list checks, and occupational health clearance. For research team members undertaking activities that do not impact on the quality of care of the participant (for example, administering questionnaires), a Letter of Access based on standard DBS checks and occupational health clearance would be appropriate.

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in England to aid study set-up.
The applicant has indicated that they intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix 11: Patient invitation letter

Queen Elizabeth Hospital Birmingham NHS
Part of University Hospitals Birmingham NHS Foundation Trust

Patient name
Address

[redacted], *Clinical Research Offices*
Old Nuclear Medicine, 1st Floor Pharmacy Building,
Heritage Building, Research & Development
University Hospitals Birmingham NHS Foundation Trust
Birmingham
B15 2TT

Mary Dutton
Email: Mary.Dutton@uhb.nhs.uk
Phone number: [*to be inserted*]

Date:

Dear [*Name to be inserted*],

Title of Project: Using Patient-Reported Outcome measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney Disease. (PRO-trACK Project)

We would like to invite you to take part in the above study being conducted by researchers from the University of Birmingham and the Queen Elizabeth Hospital, Birmingham (QEHB).

You have been invited because we believe that as a patient with chronic kidney disease, you are in the best position to provide the information we are seeking.

Before you decide, please take some time to read the information sheet which contains more details about the study.

If you are interested in taking part in the study or would like more information, please contact us using the email address or phone number above.

Yours sincerely

Mary Dutton *RN, MSc*

Senior Nurse Renal Research, QEHB

Appendix 12: Patient information sheet (qualitative study)



Title of project: Using Patient-Reported Outcome measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney Disease. (PRO-trACK Project)

Information about study 2 - Interviews

PART 1

We would like to invite you to take part in the research being conducted by researchers from the University of Birmingham and the Queen Elizabeth Hospital, Birmingham. This study will be conducted as part of an educational programme (PhD) at the University of Birmingham.

- We will ask patients with chronic kidney disease (CKD) to use a computer or smartphone to provide regular information about their quality of life and symptoms, in between their hospital appointments.
- This information can be collected using questionnaires known as 'electronic Patient-Reported Outcome Measures' or ePROMs.
- Doctors believe they could use ePROMs to find out if the patient needs urgent care, so they can take action straight away, rather than waiting for the patient's next clinic appointment.
- The purpose of this stage of our study is to find out the opinions of patients with chronic kidney disease on the use of ePROMs.

Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Feel free to discuss the study with other people if you wish.

Please ask us if there is anything that is not clear or if you would like more information. Take your time to decide whether you wish to take part or not.

Why have I been invited?

You have been invited because as a patient with chronic kidney disease (CKD), you are in the best position to provide the information we are seeking. We would like to know your thoughts on the use of an electronic questionnaire in the management of CKD. Please note that the ePROM system is not designed to be used as a replacement for direct consultation with your clinicians. Therefore, your health care needs will continue to be met through your general practice and renal service. **Please contact these services if you feel you require assistance.**

What is the purpose of the study?

The purpose of this study is to find out the opinions of patients with chronic kidney disease on the use of an electronic patient reported outcome measure (ePROM) questionnaire hosted on 'myhealth@QEHB' to remotely self-report their symptoms and quality of life. It is hoped that the ePROM system will assist clinicians in the management of patients' condition. These interviews may allow us to know the possible strengths and weaknesses of using the ePROM system in managing patients. The collection of ePROMs could make it possible to remotely monitor symptoms in patients with CKD and generate 'real time' data which may assist clinicians with the tailoring of treatment to the needs of each patient in future. In this manner, the use of PROMs may significantly improve the quality of life of patients with CKD.

What are the possible benefits of taking part?

Although there will be no direct benefit to you for participating in this study, your input will inform the development of an effective ePROM system which may enhance the care you and other patients receive in future. In addition, by taking part you will have the chance to provide your opinions/experiences which you may find rewarding and worthwhile.

Do I have to take part?

You do not have to take part. It is up to you to decide. Taking part in this research will be based around your commitments. Times for the interview can be arranged for your convenience.

If you decide to take part you will be given this information sheet to keep and be asked to sign a consent form. You are free to withdraw within 5 working days following your interview, without giving a reason. A decision not to take part or to withdraw at any time, will not affect the standard of care you receive from the NHS.

What will happen to me if I decide to take part?

You will be asked to take part in one face-to-face interview. The interview will take place either at the Institute for Translational Medicine, Queen Elizabeth Hospital or in one of the meeting rooms at the Murray Learning Centre, University of Birmingham or at another location convenient for you. The interview will be tape recorded. We expect each interview to last up to an hour.

In the unlikely event that you become distressed during your interview, you will be given the option to either delay or discontinue the interview. Contact details of the senior clinician (a consultant nephrologist) on the research team will be given to you if you wish to seek support.

Expenses

If you decide to take part in this study, we will reimburse you with a £20 gift voucher for your time and also cover any reasonable travel costs. Please retain travel receipts and tickets.

What if there is a problem?

Any complaint about the way you have been dealt with during the study, or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2

What if there is a problem?

If you have concerns about any aspect of this study, you should ask to speak to the researcher:

Dr Lee Aiyegbusi, Centre for Patient Reported Outcomes Research, University of Birmingham, Edgbaston, Birmingham B15 2TT

Tel: *[to be inserted]*

Email: oxa238@bham.ac.uk

Or the supervisor, Professor Paul Cockwell, Department of Nephrology
Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston Birmingham, B15 2GW

Tel: +44 (0) 121 371 5839

Email: Paul.Cockwell@uhb.nhs.uk

If you remain unhappy you are advised to contact the UHB Patient Advisory Liaison service (PALs) on 0121 371 3280.

Will my taking part in the study be kept confidential?
Yes.

We will follow ethical procedures and all information about you will be handled in confidence. All information provided throughout this research is protected under the *UK Data Protection Act (1998)* and will only be used by those on the research team. The study will do its utmost to preserve your anonymity and keep comments

unidentifiable, therefore your names will not be mentioned in any report of this research.

What will happen to any data I give?

Your data will be handled according to the Data Protection Act (1998), the Information Security Policy of the University of Birmingham and the NHS England Data Protection Policy 2014. The data you provide will only be accessed by the research team at the Hospital and the University of Birmingham, or by authorised people to check that the study is being carried out correctly.

The audio-recording of the interview will be used to produce a typed record of the discussion, known as a transcript, which will be anonymised. A professional transcription company will be contracted to carry out the transcription. The company will sign a confidentiality agreement before receiving such recordings which will be anonymised before we send to the company. The transcript will be analysed as part of our research. The audio recording will be securely stored at the University of Birmingham until the end of the study when it will be destroyed.

All electronic data will be saved in an encrypted file and stored on a secure server, which is backed up daily, at the University of Birmingham. Paper copies of documents such as consent forms will be stored in a locked cabinet, in a secure locked office within a secure building housed in the University of Birmingham.

Personal data such as contact details will be stored for up to 24 months after which it will be destroyed. Study data will be stored using the secure Research Data Archive (RDA), developed by the University of Birmingham IT Services, for 10 years before it is destroyed.

What will happen to the results of the research study?

The results of the study will be used by the research team to finalise development of the ePROM system. We plan to publish the results in a leading scientific journal and present at international scientific conferences. You will not be identifiable in any published report or presentation. You will be sent a summary of the preliminary and

final results (unless you tell us that you do not want one). The ePROM system will be tested further in future research conducted by the study team.

Who is organising and funding this research?

This project is funded as part of the Health Foundation's Improvement Science programme. The Health Foundation is an independent charity working to improve the quality of healthcare in the UK.

Who has reviewed the study?

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has received approval from West Midlands – Edgbaston Research Ethics Committee (REC Ref: 17/WM/0010).

Contact details for further information

For further information please contact Dr Lee Aiyegbusi. Tel: *[to be inserted].*

Appendix 13: Clinician information sheet (Qualitative study)

Title of project: Using Patient-Reported Outcome measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney Disease. (PRO-trACK Project)

Information about study 2 – Focus group

PART 1

We would like to invite you to take part in the research being conducted by researchers from the University of Birmingham and the Queen Elizabeth Hospital, Birmingham. This study will be conducted as part of an educational programme (PhD) at the University of Birmingham.

- We will ask patients with chronic kidney disease (CKD) to use a computer or smartphone to provide regular information about their quality of life and symptoms, in between their hospital appointments.
- This information can be collected using questionnaires known as 'electronic Patient-Reported Outcome Measures' or ePROMs.
- The purpose of this stage of our study is to find out the opinions of clinicians who manage patients with CKD at the Renal Unit, QEHB, members of the myHealth team and hospital management staff on the use of ePROMs.

Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Feel free to discuss the study with other people if you wish.

Please ask us if there is anything that is not clear or if you would like more information. Take your time to decide whether you wish to take part or not.

Why have I been invited?

You have been invited because as a clinician, member of the myHealth team or hospital management staff at QEHB, you are in the best position to provide the information we are seeking. We would like to know your thoughts on the use of an electronic questionnaire in the management of patients with CKD. Please note that the ePROM system is not designed to be used as a replacement for direct consultation with clinicians.

What is the purpose of the study?

The purpose of this study is to find out the opinions of clinicians; members of the myHealth team and hospital management staff, on the use of an electronic patient reported outcome measure (ePROM) questionnaire hosted on 'myhealth@QEHB' to remotely monitor the symptoms and quality of life of patients. The focus group discussions may allow us to know the possible strengths and weaknesses of using the ePROM system in managing patients. We may contact you later for interviews if necessary. The collection of ePROMs could make it possible to remotely monitor symptoms in patients with CKD and generate ‘real time’ data

which may assist clinicians with the tailoring of treatment to the needs of each patient in future. In this manner, the use of PROMs may significantly improve the quality of life of patients with CKD.

What are the possible benefits of taking part?

Although there will be no direct benefit to you for participating in this study, your input will inform the development of an effective ePROM system which may enhance the

care patients receive in future. In addition, by taking part you will have the chance to provide your opinions/experiences which you may find rewarding and worthwhile.

Do I have to take part?

You do not have to take part. It is up to you to decide. Taking part in this research will be based around your commitments. If you decide to take part you will be given this information sheet to keep and be asked to sign a consent form.

You are free to withdraw for any reason before the focus group session without giving a reason. A decision not to participate or to withdraw will not affect your employment.

During and after participating in a focus group discussion, it will be impossible to withdraw your data due to the fact that the opinions of participants will be audio recorded as a group and the discussion will evolves as individuals express their views and respond to the views of others.

If you agree to be interviewed later, you are free to withdraw within 5 working days following your interview, without giving a reason. A decision not to take part or to withdraw will not affect the standard of care you receive from the NHS.

What will happen to me if I decide to take part?

You will be asked to take part in a focus group discussion. The focus group discussion will take place either at the Institute for Translational Medicine, Queen Elizabeth Hospital or in one of the meeting rooms at the Murray Learning Centre, University of Birmingham. The focus group discussion will be tape recorded and it is expected to last between 1 ½ and 2 hours. Light refreshments will be provided during the discussion. If we contact you later for an interview this will take no longer than 1 hour.

What if there is a problem?

Any complaint about the way you have been dealt with during the study, or any possible harm you might suffer will be addressed. The detailed information on this is given in Part 2.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2

What if there is a problem?

If you have concerns about any aspect of this study, you should ask to speak to the researcher:

Dr Lee Aiyegbusi, Centre for Patient Reported Outcomes Research, University of Birmingham, Edgbaston, Birmingham B15 2TT

Tel: *[to be inserted]*

Email: oxa238@bham.ac.uk

Or the supervisor, Professor Paul Cockwell, Department of Nephrology
Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Mindelsohn Way, Edgbaston Birmingham, B15 2GW

Tel: +44 (0) 121 371 5839

Email: Paul.Cockwell@uhb.nhs.uk

Will my taking part in the study be kept confidential?
Yes.

We will follow ethical procedures and all information about you will be handled in confidence. All information provided throughout this research is protected under the *UK Data Protection Act (1998)* and will only be used by those on the research team. The study will do its utmost to preserve your anonymity and keep comments unidentifiable, therefore your names will not be mentioned in any report of this research.

What will happen to any data I give?

Your data will be handled according to the Data Protection Act (1998), the Information Security Policy of the University of Birmingham and the NHS England Data Protection Policy 2014. The data you provide will only be accessed by the research team at the Hospital and the University of Birmingham, or by authorised people to check that the study is being carried out correctly.

The audio-recording of the focus groups and likely interviews will be used to produce a typed record of the discussion, known as a transcript, which will be anonymised. A professional transcription company will be contracted to carry out the transcription. The company will sign a confidentiality agreement before receiving such recordings which will be anonymised before we send to the company. The transcript will be analysed as part of our research. The audio recording will be securely stored at the University of Birmingham until the end of the study when it will be destroyed.

All electronic data will be saved in an encrypted file and stored on a secure server, which is backed up daily, at the University of Birmingham. Paper copies of documents such as consent forms will be stored in a locked cabinet, in a secure locked office within a secure building housed in the University of Birmingham.

Personal data such as contact details will be stored for up to 24 months after which it will be destroyed. Study data will be stored using the secure Research Data Archive (RDA), developed by the University of Birmingham IT Services, for 10 years before it is destroyed.

What will happen to the results of the research study?

The results of the study will be used by the research team to finalise development of the ePROM system. We plan to publish the results in a leading scientific journal and present at international scientific conferences. You will not be identifiable in any published report or presentation. You will be sent a summary of the preliminary and final results (unless you tell us that you do not want one). The ePROM system will be tested further in future research conducted by the study team.

Who is organising and funding this research?

This project is funded as part of the Health Foundation's Improvement Science programme. The Health Foundation is an independent charity working to improve the quality of healthcare in the UK.

Who has reviewed the study?

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee to protect your safety, rights, wellbeing and dignity. This study has received approval from West Midlands – Edgbaston Research Ethics Committee (REC Ref: 17/WM/0010).

Contact details for further information

For further information please contact Dr Lee Aiyegbusi. Tel: *[to be inserted].*

Appendix 14: Patient consent form (Qualitative study)

UNIVERSITY OF BIRMINGHAM | CPROR
CENTRE FOR PATIENT REPORTED OUTCOMES RESEARCH

CONSENT FORM – INTERVIEW

Title of Project: Using Patient-Reported Outcome measures (PROMs) to promote quality of care and safety in the management of patients with Advanced Chronic Kidney Disease. (PRO-trACK Project)

Chief Investigator: Dr Lee Aiyegbusi

Identification Number:

Please initial box

1. I confirm that I have read and understand the information sheet dated..... (version.....) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. ☐

3. I understand that the interview will be tape recorded but that I can refuse to answer a question if I wish and stop the interview at any time without giving an explanation. ☐

4. I understand that the interview will be typed up and that the conversations in that interview may be used when the research team write about the study. However, any quotes will be anonymised and that no information will be used in any presentations or reports that could lead to my identification. ☐

5. I understand that relevant sections of my medical notes and data collected during the study may be looked at by individuals from the University of Birmingham, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records ☐

6. I agree to take part in the above study. ☐

7. **OPTIONAL**: I agree for the research team to contact me regarding participation in future research studies. YES ☐

NO ☐

Name of Patient	Date	Signature

Name of researcher	Date	Signature

Appendix 15: Topic guide for patient interviews

S1. Topic guide for patient interviews

- Was the purpose for the study clear from reading the patient information sheets?
- What do you think of this study?

  Prompts:

  Do you think the study will benefit patients? If so why and how?

- Views on the content of the questionnaire(s). What do you think of the questionnaires?

  Prompts:

  9. What do you think of the lengths?
  10. What do you think of the questions?
      Prompts:

      Do you understand them? Are the questions (and their response options)

      worded in a way that would be easily understood by patients? How easy do

      you think it is to complete the questionnaires?

  11. Do individual questionnaires cover all aspects that might be important to patients with advanced CKD, or are any elements missing? If so which?
  12. Are there questions that might be seen as unacceptable? If so why?

- How often would you be prepared to complete the questionnaire(s) as part of your care?
- How would you prefer to complete the questionnaires?

  Prompt: by PC, Smartphone, tablet, telephone voice recognition or paper completion?

- What might discourage you from completing the questionnaires?
- What might encourage you to complete the questionnaires?
- Do you think the benefits you mentioned earlier are sufficient to motivate other patients to fill the questionnaire on a regular basis?
- If not mentioned, ask if feedback is desired. What kind of feedback would you prefer?

- Are there any issues about completing these questionnaires that worry you?
- Any final comments?

Appendix 16: Topic guide for clinician focus group/interviews

S2. Topic Guide for clinician focus group/interviews

- What are your thoughts on the use of ePROMs in the management of CKD patients?

  Prompt: Do you think the use of ePROMs will benefit patients? If so why and how? If not why?

- How often would you like patients to complete questionnaires like these?
- How would you like the data to be displayed?
- Are there any issues about using these questionnaires that you would like to mention?
- What factors may improve the use of ePROM data by clinicians?
- What factors may discourage the use of ePROM data by clinicians?
- Views on KDQOL-36 and IPOS-Renal

  13. What are your thoughts on the questionnaires?

      Prompts:

      What are your opinions of the wording of the questions and their response options?

      What about the scope of the questionnaires? Do individual questionnaires cover all aspects that might be important to patients with advanced CKD, or are any elements missing?

      Are there questions that might be seen as unacceptable to patients? If so which questions and why?

- Is there anything else anyone would like to say on the use of PROMs for the management of patients with CKD?

Appendix 17: Review of frameworks for research impact (Co-authored publication)

PLOS | MEDICINE

RESEARCH ARTICLE

Assessing the impact of healthcare research: A systematic review of methodological frameworks

Samantha Cruz Rivera, Derek G. Kyte*, Olalekan Lee Aiyegbusi, Thomas J. Keeley, Melanie J. Calvert

Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom

* d.g.kyte@bham.ac.uk

Check for updates

OPEN ACCESS

Citation: Cruz Rivera S, Kyte DG, Aiyegbusi OL, Keeley TJ, Calvert MJ (2017) Assessing the impact of healthcare research: A systematic review of methodological frameworks. PLoS Med 14(8): e1002370. https://doi.org/10.1371/journal.pmed.1002370

Academic Editor: Mike Clarke, Queens University Belfast, UNITED KINGDOM

Received: February 28, 2017

Accepted: July 7, 2017

Published: August 9, 2017

Copyright: © 2017 Cruz Rivera et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and supporting files.

Funding: Funding was received from Consejo Nacional de Ciencia y Tecnología (CONACYT). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript (http://www.conacyt.mx/).

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: MJC has received consultancy fees from Astellas and Ferring pharma

Abstract

Background

Increasingly, researchers need to demonstrate the impact of their research to their sponsors, funders, and fellow academics. However, the most appropriate way of measuring the impact of healthcare research is subject to debate. We aimed to identify the existing methodological frameworks used to measure healthcare research impact and to summarise the common themes and metrics in an impact matrix.

Methods and findings

Two independent investigators systematically searched the Medical Literature Analysis and Retrieval System Online (MEDLINE), the Excerpta Medica Database (EMBASE), the Cumulative Index to Nursing and Allied Health Literature (CINAHL+), the Health Management Information Consortium, and the Journal of Research Evaluation from inception until May 2017 for publications that presented a methodological framework for research impact. We then summarised the common concepts and themes across methodological frameworks and identified the metrics used to evaluate differing forms of impact. Twenty-four unique methodological frameworks were identified, addressing 5 broad categories of impact: (1) 'primary research-related impact', (2) 'influence on policy making', (3) 'health and health systems impact', (4) 'health-related and societal impact', and (5) 'broader economic impact'. These categories were subdivided into 16 common impact subgroups. Authors of the included publications proposed 80 different metrics aimed at measuring impact in these areas. The main limitation of the study was the potential exclusion of relevant articles, as a consequence of the poor indexing of the databases searched.

Conclusions

The measurement of research impact is an essential exercise to help direct the allocation of limited research resources, to maximise research benefit, and to help minimise research waste. This review provides a collective summary of existing methodological frameworks for

and travel fees from the European Society of Cardiology outside the submitted work. TJK is in full-time paid employment for PAREXEL International.

Abbreviations: AIHS, Alberta Innovates—Health Solutions; CAHS, Canadian Academy of Health Sciences; CIHR, Canadian Institutes of Health Research; CINAHL+, Cumulative Index to Nursing and Allied Health Literature; EMBASE, Excerpta Medica Database; ERA, Excellence in Research for Australia; HEFCE, Higher Education Funding Council for England; HMIC, Health Management Information Consortium; HTA, Health Technology Assessment; IOM, Impact Oriented Monitoring; MDG, Millennium Development Goal; NHS, National Health Service; MEDLINE, Medical Literature Analysis and Retrieval System Online; PHC RIS, Primary Health Care Research & Information Service; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PROM, patient-reported outcome measures; QALY, quality-adjusted life year; R&D, research and development; RAE, Research Assessment Exercise; REF, Research Excellence Framework; RIF, Research Impact Framework; RQF, Research Quality Framework; SDG, Sustainable Development Goal; SIAMPI, Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society.

research impact, which funders may use to inform the measurement of research impact and researchers may use to inform study design decisions aimed at maximising the short-, medium-, and long-term impact of their research.

Author summary

Why was this study done?

- There is a growing interest in demonstrating the impact of research in order to minimise research waste, allocate resources efficiently, and maximise the benefit of research. However, there is no consensus on which is the most appropriate tool to measure the impact of research.
- To our knowledge, this review is the first to synthesise existing methodological frameworks for healthcare research impact, and the associated impact metrics by which various authors have proposed impact should be measured, into a unified matrix.

What did the researchers do and find?

- We conducted a systematic review identifying 24 existing methodological research impact frameworks.
- We scrutinised the sample, identifying and summarising 5 proposed impact categories, 16 impact subcategories, and over 80 metrics into an impact matrix and methodological framework.

What do these findings mean?

- This simplified consolidated methodological framework will help researchers to understand how a research study may give rise to differing forms of impact, as well as in what ways and at which time points these potential impacts might be measured.
- Incorporating these insights into the design of a study could enhance impact, optimizing the use of research resources.

Introduction

In 2010, approximately US$240 billion was invested in healthcare research worldwide [1]. Such research is utilised by policy makers, healthcare providers, and clinicians to make important evidence-based decisions aimed at maximising patient benefit, whilst ensuring that limited healthcare resources are used as efficiently as possible to facilitate effective and sustainable service delivery. It is therefore essential that this research is of high quality and that it is impactful—i.e., it delivers demonstrable benefits to society and the wider economy whilst minimising research waste [1,2]. Research impact can be defined as 'any identifiable 'benefit to, or

positive influence on the economy, society, public policy or services, health, the environment, quality of life or academia' (p. 26) [3].

There are many purported benefits associated with the measurement of research impact, including the ability to (1) assess the quality of the research and its subsequent benefits to society; (2) inform and influence optimal policy and funding allocation; (3) demonstrate accountability, the value of research in terms of efficiency and effectiveness to the government, stakeholders, and society; and (4) maximise impact through better understanding the concept and pathways to impact [4–7].

Measuring and monitoring the impact of healthcare research has become increasingly common in the United Kingdom [5], Australia [5], and Canada [8], as governments, organisations, and higher education institutions seek a framework to allocate funds to projects that are more likely to bring the most benefit to society and the economy [5]. For example, in the UK, the 2014 Research Excellence Framework (REF) has recently been used to assess the quality and impact of research in higher education institutions, through the assessment of impact cases studies and selected qualitative impact metrics [9]. This is the first initiative to allocate research funding based on the economic, societal, and cultural impact of research, although it should be noted that research impact only drives a proportion of this allocation (approximately 20%) [9].

In the UK REF, the measurement of research impact is seen as increasingly important. However, the impact element of the REF has been criticised in some quarters [10,11]. Critics deride the fact that REF impact is determined in a relatively simplistic way, utilising researcher-generated case studies, which commonly attempt to link a particular research outcome to an associated policy or health improvement despite the fact that the wider literature highlights great diversity in the way research impact may be demonstrated [12,13]. This led to the current debate about the optimal method of measuring impact in the future REF [10,14]. The Stern review suggested that research impact should not only focus on socioeconomic impact but should also include impact on government policy, public engagement, academic impacts outside the field, and teaching to showcase interdisciplinary collaborative impact [10,11]. The Higher Education Funding Council for England (HEFCE) has recently set out the proposals for the REF 2021 exercise, confirming that the measurement of such impact will continue to form an important part of the process [15].

With increasing pressure for healthcare research to lead to demonstrable health, economic, and societal impact, there is a need for researchers to understand existing methodological impact frameworks and the means by which impact may be quantified (i.e., impact metrics; see Box 1, 'Definitions') to better inform research activities and funding decisions. From a

Box 1. Definitions

- **Research impact:** 'any identifiable benefit to, or positive influence on, the economy, society, public policy or services, health, the environment, quality of life, or academia' (p. 26) [3].
- **Methodological framework:** 'a body of methods, rules and postulates employed by a particular procedure or set of procedures (i.e., framework characteristics and development)' [18].
- **Pathway:** 'a way of achieving a specified result; a course of action' [19].
- **Quantitative metrics:** 'a system or standard of [quantitative] measurement' [20].
- **Narrative metrics:** 'a spoken or written account of connected events; a story' [21].

researcher's perspective, understanding the optimal pathways to impact can help inform study design aimed at maximising the impact of the project. At the same time, funders need to understand which aspects of impact they should focus on when allocating awards so they can make the most of their investment and bring the greatest benefit to patients and society [2,4,5,16,17].

Whilst previous researchers have summarised existing methodological frameworks and impact case studies [4,22–27], they have not summarised the metrics for use by researchers, funders, and policy makers. The aim of this review was therefore to (1) identify the methodological frameworks used to measure healthcare research impact using systematic methods, (2) summarise common impact themes and metrics in an impact matrix, and (3) provide a simplified consolidated resource for use by funders, researchers, and policy makers.

Methods

Search strategy and selection criteria

Initially, a search strategy was developed to identify the available literature regarding the different methods to measure research impact. The following keywords: 'Impact', 'Framework', and 'Research', and their synonyms, were used during the search of the Medical Literature Analysis and Retrieval System Online (MEDLINE; Ovid) database, the Excerpta Medica Database (EMBASE), the Health Management Information Consortium (HMIC) database, and the Cumulative Index to Nursing and Allied Health Literature (CINAHL+) database (inception to May 2017; see S1 Appendix for the full search strategy). Additionally, the nonindexed Journal of Research Evaluation was hand searched during the same timeframe using the keyword 'Impact'. Other relevant articles were identified through 3 Internet search engines (Google, Google Scholar, and Google Images) using the keywords 'Impact', 'Framework', and 'Research', with the first 50 results screened. Google Images was searched because different methodological frameworks are summarised in a single image and can easily be identified through this search engine. Finally, additional publications were sought through communication with experts.

Following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (see S1 PRISMA Checklist), 2 independent investigators systematically screened for publications describing, evaluating, or utilising a methodological research impact framework within the context of healthcare research [28]. Papers were eligible if they included full or partial methodological frameworks or pathways to research impact; both primary research and systematic reviews fitting these criteria were included. We included any methodological framework identified (original or modified versions) at the point of first occurrence. In addition, methodological frameworks were included if they were applicable to the healthcare discipline with no need of modification within their structure. We defined 'methodological framework' as 'a body of methods, rules and postulates employed by a particular procedure or set of procedures (i.e., framework characteristics and development)' [18], whereas we defined 'pathway' as 'a way of achieving a specified result; a course of action' [19]. Studies were excluded if they presented an existing (unmodified) methodological framework previously available elsewhere, did not explicitly describe a methodological framework but rather focused on a single metric (e.g., bibliometric analysis), focused on the impact or effectiveness of interventions rather than that of the research, or presented case study data only. There were no language restrictions.

Data screening

Records were downloaded into Endnote (version X7.3.1), and duplicates were removed. Two independent investigators (SCR and OLA) conducted all screening following a pilot aimed at

refining the process. The records were screened by title and abstract before full-text articles of potentially eligible publications were retrieved for evaluation. A full-text screening identified the publications included for data extraction. Discrepancies were resolved through discussion, with the involvement of a third reviewer (MJC, DGK, and TJK) when necessary.

Data extraction and analysis

Data extraction occurred after the final selection of included articles. SCR and OLA independently extracted details of impact methodological frameworks, the country of origin, and the year of publication, as well as the source, the framework description, and the methodology used to develop the framework. Information regarding the methodology used to develop each methodological framework was also extracted from framework webpages where available. Investigators also extracted details regarding each framework's impact categories and subgroups, along with their proposed time to impact ('short-term', 'mid-term', or 'long-term') and the details of any metrics that had been proposed to measure impact, which are depicted in an impact matrix. The structure of the matrix was informed by the work of M. Buxton and S. Hanney [2], P. Buykx et al. [5], S. Kuruvila et al. [29], and A. Weiss [30], with the intention of mapping metrics presented in previous methodological frameworks in a concise way. A consensus meeting with MJC, DGK, and TJK was held to solve disagreements and finalise the data extraction process.

Results

Included studies

Our original search strategy identified 359 citations from MEDLINE (Ovid), EMBASE, CINAHL+, HMIC, and the Journal of Research Evaluation, and 101 citations were returned using other sources (Google, Google Images, Google Scholar, and expert communication) (see Fig 1) [28]. In total, we retrieved 54 full-text articles for review. At this stage, 39 articles were excluded, as they did not propose new or modified methodological frameworks. An additional 15 articles were included following the backward and forward citation method. A total of 31 relevant articles were included in the final analysis, of which 24 were articles presenting unique frameworks and the remaining 7 were systematic reviews [4,22–27]. The search strategy was rerun on 15 May 2017. A further 19 publications were screened, and 2 were taken forward to full-text screening but were ineligible for inclusion.

Methodological framework characteristics

The characteristics of the 24 included methodological frameworks are summarised in Table 1, 'Methodological framework characteristics'. Fourteen publications proposed academic-orientated frameworks, which focused on measuring academic, societal, economic, and cultural impact using narrative and quantitative metrics [2,3,5,8,29,31–39]. Five publications focused on assessing the impact of research by focusing on the interaction process between stakeholders and researchers ('productive interactions'), which is a requirement to achieve research impact. This approach tries to address the issue of attributing research impact to metrics [7,40–43]. Two frameworks focused on the importance of partnerships between researchers and policy makers, as a core element to accomplish research impact [44,45]. An additional 2 frameworks focused on evaluating the pathways to impact, i.e., linking processes between research and impact [30,46]. One framework assessed the ability of health technology to influence efficiency of healthcare systems [47]. Eight frameworks were developed in the UK [2,3,29,37,39,42,43,45], 6 in Canada [8,33,34,44,46,47], 4 in Australia [5,31,35,38], 3 in the

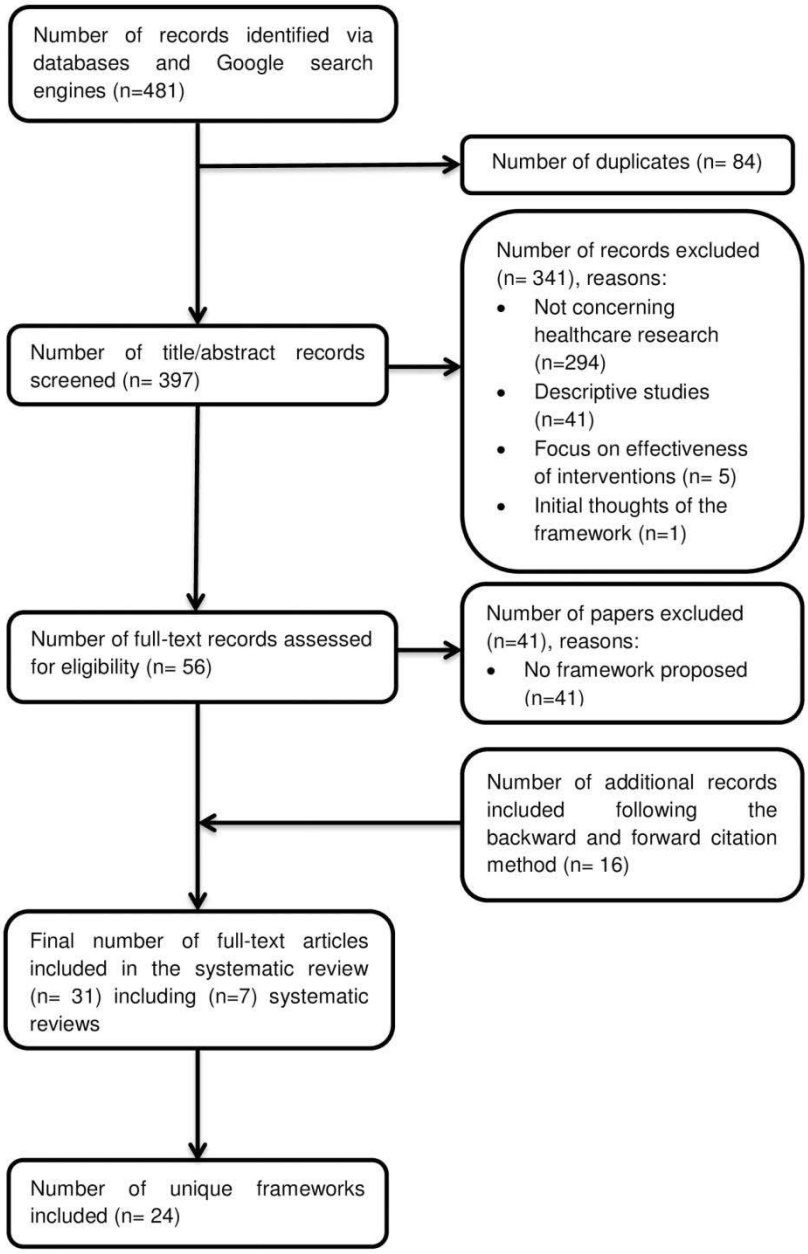


Fig 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

https://doi.org/10.1371/journal.pmed.1002370.g001

Netherlands [7,40,41], and 2 in the United States [30,36], with 1 model developed with input from various countries [32].

Table 1. Methodological framework characteristics.

Framework	Original source for modified framework	Framework description	Impact categories	Methodology followed to develop the framework			Pilot phase
				Literature review	Stakeholder involvement	Methods to incorporate stakeholder views	
Buxton and Hanney **Payback Framework** [2] UK, 1996	• The classic/ purist/knowledge-driven model • The problem-solving/ engineering model • The political model • The enlightenment/ percolation/ limestone model • The interactive/ social interaction model [48–54]	Framework used to assess health research impact through academic outputs and wider societal benefits	• Knowledge benefits • Benefits to future research and research use • Political and administrative benefits • Health sector benefits • Broader economic benefits	Yes, but no systematic approach reported	None reported	None reported	A case study approach was undertaken to determine the effectiveness of the model and exemplify the impact categories
Canadian Institutes of Health Research (CIHR) [8] Canada, 2005	• Payback Framework [2]	Framework designed to gauge the impact of health research and the benefits of investing in health research	• Knowledge production • Research targeting and capacity • Informing policy • Health and health sector benefits • Economic impacts	Yes, but no systematic approach reported	Meetings, presentations, and panel discussions with the Canadian government, academics, health research stakeholders, and international funding agencies	Not specified	No information reported in the main paper or on the webpage
Canadian Academy of Health Sciences (CAHS) [33] Canada, 2009	• Payback Framework [2] • CIHR framework [8]	Framework focused on evaluating how research activity influences decision making in order to improve health, the economy, and social benefits	• Advancing knowledge • Building capacity • Informing decision making • Health • Broad socioeconomic impact	Yes, but no systematic approach reported	Semistructured interviews were conducted to collect stakeholder, sponsor, and external expert feedback	Conceptual cluster analysis was used to analyse the data collected, which helped to capture the perspective of different stakeholders involved in the health system	Six papers were commissioned to test the applicability of the framework. After modifications, the project was sent to external review and changes were made until its approval
K. Graham et al. **Alberta Innovates—Health Solutions (AIHS) impact framework** [34] Canada, 2012	• Payback Framework [2] CAHS model [33]	A framework to measure, assess, and illustrate the relationship between research investment and impact. The ultimate goal is to contribute to the most societal benefit among the people	• Advancing knowledge • Building capacity • Informing decision making • Health • Broad socioeconomic impact • Organisational performance	No systematic approach reported	None reported	None reported	The CAHS model was assessed for applicability and feasibility through retrospective and prospective studies and other tools to map pathways to impact. The findings of the assessment provided feedback to improve the AIHS model
G. Cohen et al. [35] Australia, 2014	• Payback Framework [2] • AIHS framework [34]	To evaluate the policy and practicebenefits of research outputs, which take place later and beyond the research setting	• Scholarly outputs • Translational outputs • Policy or practice impacts • Long-term population outcomes	No systematic approach reported	Two online surveys and semistructured interviews were conducted among the primary chief investigators of the research grants included	Not specified	Case studies were summarised and presented to an experience panel, which scored the relevant impact categories of this framework
S. Kuruvilla et al. **Research Impact Framework (RIF)** [29] UK, 2006	• Payback Framework [2]	Conceptual framework to describe the possible impacts of health research outcomes	• Research-related impacts • Policy impacts • Service impacts: health and intersectorial • Societal impacts	No systematic approach reported	Semistructured interviews with principal investigators of selected projects	Thematic analysis was adopted to analyse the data. The categories of the framework were used as themes	The RIF was validated through consistency with available health research literature and empirical analysis of research projects
L. Kalucy et al. **Primary Health Care Research & Information Service (PHCRIS)** [31,55] Australia, 2007	• Payback Framework [2]	Methodology designed to assess primary healthcare research. Strong collaboration, personal relationships and the participation of practitioners, health care managers, and policy makers in the definition of the research questions and in the research process were identified as the strongest pathways to impact	• Knowledge production • Research targeting, capacity, building, and absorption • Informing policy and product development • Health and health sector benefits • Broader economic benefits • Research transfer	No systematic approach reported	Part 1: Telephone interviews were conducted among the 4 chief investigators of the included projects funded by the National Health and Medical Research Council	Data provided by the interviewees were analysed following the Payback Framework and thematic analysis approach	Parts 2–17: chief investigators completed an online questionnaire to refine the methodology to measure research impact

(*Continued*)

Table 1. (*Continued*)

Framework	Original source for modified framework	Framework description	Impact categories	Methodology followed to develop the framework			
				Literature review	Stakeholder involvement	Methods to incorporate stakeholder views	Pilot phase
J. Guinea et al. **Impact Oriented Monitoring (IOM) [32]** Various countries, 2015	• Payback Framework [2]	Methodology used to identify and assess the impacts of health projects through a set of predefined categories	• Advancing knowledge • Capacity building and research targeting • Informing decision making, practice, and policy • Population health and health sector benefits • Dissemination and knowledge transfer	No systematic approach reported	A project coordinators' survey was conducted to determine and collect all the possible public health research benefits. An end users' survey was conducted to determine the usefulness and practicality of the results of the coordinators' survey to measure research impact. Additionally, a scoring matrix was developed to assess project impacts	Not specified	A small sample of research projects was used to test some of the methods incorporated to measure research impact
J. Lavis et al. **Exchange model [44]** Canada, 2003	–	Assessment tool to measure decision-making impact of health research. Impact measures are categorised according to the level of impact to be measured and the mechanism of research uptake: producer-push, user-pull, and exchange measures	• Producer-push process • User-pull process • Exchange process	No systematic approach reported	None reported	None reported	Two examples were used to demonstrate how the assessment tool can be used
L. Meagher et al. [43] UK, 2008	• Linkage and exchange model [56]	Methodology for assessing research impact of policy and practice	• Primary knowledge producers • Knowledge users, beneficiaries, brokers, and intermediaries • Impacts or research (outcomes) • Research impact processes • Lessons learned and recommendations • Methods for identifying and assessing nonacademic research impact	Yes, but no systematic approach reported	Award holders of the UK's Economic and Social Research Council (ESRC) within the psychology field, heads of departments, ESRC-funded principal researchers, and research users were recruited to conduct a questionnaire and survey; 2 focus group interviews, semistructured interviews, media-related searches, and case studies to determine the level of engagement with research users, impact and processes, activities, and roles leading to impact	Not specified	None
P. Buykx et al. **The Health Services Research Impact Framework [5]** Australia, 2012	• PHC RIS [31] • RIF [29] • Exchange Model [44]	Framework for recording and monitoring the impact of health research. This framework consolidates the most relevant elements of the PHS RIS, RIF, and exchange model	• Knowledge generation and communication • Capacity building, training, and leadership • Informing policy • Improving health and health systems impact • Social and economic benefit impact	No systematic approach reported	None reported	None	Not available in the main paper
Higher Education Funding Council for England (HEFCE) **Research Assessment Exercise (RAE) [39]** UK, 2005	–	The aim of the RAE is to assess the quality of research conducted by academic institutions in the UK, in order to inform funding decisions by higher education funding bodies.	• RAE1: Staff information • RAE2: Research output • RAE3: Research scholarships • RAE4: Attractiveness for external funding • RAE5: Further information on groups of research	None reported	Focus groups, workshops, and meetings with HEFCE institutions, funding body officers, external experts, and stakeholder groups	Not specified	None

(*Continued*)

Table 1. (*Continued*)

Framework	Original source for modified framework	Framework description	Impact categories	Methodology followed to develop the framework			Pilot phase
				Literature review	Stakeholder involvement	Methods to incorporate stakeholder views	
C. Donovan **Australian Research Quality Framework (RQF)** [38] **Substituted in 2015 by the Excellence in Research for Australia (ERA) framework** Australia, 2008	–	A 5-point rating scale to evaluate research excellence and societal returns of publicly funded research. This framework highlights the importance of end-users' interactions to enhance the use of research	The impact rating scale: • Research has produced an outstanding social, economic, environmental, or cultural benefit. • Research has produced a significant social, economic, environmental, or cultural benefit. • Research has produced new policies, products, attitudes, behaviours, or outlooks in the end-user communities. • Research has engaged with the end-user community to address a social, economic, environmental, or cultural issue. • Research has had limited or no identifiable social, economic, environmental, or cultural outcome.	None reported	The technical working group on research impact (senior university managers, representatives from business and industry, experts in impact evaluation, and members of the development advisory group) and the Australian higher education sector participated during different phases of the development of the framework. The Australian higher education sector was consulted. The technical working group was in charge of further development of the framework characteristics	Not specified	None
HEFCE **Research Excellence Framework (REF 2014)** [3] UK, 2011	• RQF [38,39]	Framework for assessing the quality of research of UK higher education institutions	• REF1: Staff details • REF2: Research outputs • REF3: Impact template and case studies • REF4: Environmental data • REF5: Environmental template	Scoping study, no systematic approach reported	An initial consultation on the REF was conducted among HEFCE institutions and other stakeholders to determine the potential of introducing bibliometric indicators in the REF. A second consultation was conducted based on the results of the first initiative, which included proposals on how to assess the impact of research	Not specified	A pilot exercise was conducted to test and develop bibliometric indicators to measure research quality. A second pilot was conducted to test the proposals to assess research impact
R. Jacob and M. McGregor **Health Technology Assessment (HTA) Organisation Assessment Framework** [47] Canada, 2007	–	The impact of health technology assessment is measured by the ability to influence the efficiency of the healthcare system	Levels of importance according to the type of policy involved: • Level 1: General statements of ministerial policy • Level 2: Planning guidelines for health services organisation • Level 3: Practice norms prescribed by the professional corporation • Level 4: Ministry rules concerning coverage of health services • Level 5: Hospital rules concerning utilisation of services • Level 6: Ministry decisions on the organisation of specific health services • Level 7: Ministry decisions concerning the allocation of resources	No, a case study approach was undertaken to determine the impact of health technology assessment on policy decision	Interviews and questionnaires were conducted with stakeholders affiliated with the Canadian Ministry of Health; those who were involved in defining policy were contacted to corroborate the supporting documentation collected	Not specified	None

(*Continued*)

Table 1. (*Continued*)

Framework	Original source for modified framework	Framework description	Impact categories	Methodology followed to develop the framework			Pilot phase
				Literature review	Stakeholder involvement	Methods to incorporate stakeholder views	
R. Landry et al. **Research Utilisation Ladder** [46] Canada, 2001	• Knott and Wildavsky Model [57]	Assessment of the pathway in which research progresses towards its utilisation by decision makers and practitioners	• Transmission • Cognition • Reference • Effort • Influence • Application	Yes, but no systematic approach reported	A mail survey and telephone calls were used to identify potential participants. Once the participants were recruited, a questionnaire focused on utilisation knowledge was distributed. The stakeholders involved were faculty members of different Canadian universities	Not specified	A modified version of the Knott and Wildavsky model was used to measure knowledge utilization from the data collected. A quantitative approach was adopted to analyse the data of the model
C. Sarli et al. **The Becker Medical Library Model for Assessment of Research Impact** [36] US, 2010	• W. K. Kellogg Foundation Model [58]	Methodology beyond citation counts, to assess research impact as a result of research output. This model allows the interaction between researchers and institutions to document and quantify the impact of research	• Research output • Knowledge transfer • Clinical implementations • Community benefit	Yes, but no systematic approach reported	Authors consulted expert opinion, researchers, clinicians, and librarians to identify indicators to measure research impact	Not specified	The evidence available to measure the impact of the Ocular Hypertension Treatment Study was analysed using the preliminary framework. The outcomes refined the framework leading to a tool to assess research impact
V. C. Brueton et al. [37] UK, 2014	• The Becker Model [36]	Framework used to better report and assess impact of methodological research	• Advancement of knowledge • Implementation	No systematic approach reported	Two semistructured interviews and email queries to the methodologists for the included projects were used to examine other indicators and analyse evidence of research implementation	Not specified	None
Royal Netherlands Academy of Arts and Sciences **The social impact of applied research** [41] The Netherlands, 2002	–	Measuring the societal impact of applied research, as an incentive for researchers to improve their performance within this field	• Science and certified knowledge • Education and training • Innovation and professionalism • Public policy • Collaboration and visibility	Yes, but not systematic approach reported	None	None	None
A. Weiss **Logic Model for Medical Research** [30] US, 2007	• United Way Model [59]	Measurement outcomes (awareness, implementation, and patient benefit) provide the basis to assess the quality of the investment in research	• Initial outcome: awareness • Intermediate outcome: implementation • Long-term outcome: patient benefit	No systematic approach reported	None	None	Qualitative methods and a case study were conducted to assess the impact of the model developed
J. Canavan et al. [45] UK, 2009	• Models of research impact: A cross sector review of literature and practice [60]	This approach proposes 6 recommendations that can increase the likelihood of an effective and practical measurement of research impact policy	Recommendations: • Commissioner-researcher partnerships • Embedded researcher model • Planning for impact from the outset • Adopting facilitative research methodological strategies • Following good practice towards research impact • Connecting research measurement to performance management	Yes, but no systematic approach reported	None	None	A case study was developed to exemplify the framework proposed

(*Continued*)

Table 1. (*Continued*)

Framework	Original source for modified framework	Framework description	Impact categories	Methodology followed to develop the framework			Pilot phase
				Literature review	Stakeholder involvement	Methods to incorporate stakeholder views	
J. Spaapen and L. van Drooge **Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society (SIAMPI)** [40] The Netherlands, 2011	–	Learning tool to better understand how research interactions lead to social impact. This approach focuses on productive interactions, especially exchanges between researchers and stakeholders	Types of productive interactions: • Direct interactions • Indirect interactions • Financial interactions	Yes, but no systematic approach reported	The European commission, research organisations, science policy makers, research councils, academies, and other bodies involved in research were involved in meetings and discussions	A case study approach was used to test the framework. The feedback of the case study representatives was used to refine the framework [61]	Not specified
M. O. Kok and A. Schuit **Contribution Mapping** [7] The Netherlands, 2012	–	An approach to monitor and evaluate contributions to determine how the utilisation of research can contribute to better action for health. The method focuses on processes, actors, and efforts to enhance contributions and enable alignment of efforts.	Research related contributions categories: • Changes in abilities and actions of involved and linked actors • Contributed knowledge products • Contributions through linked utilisation • Indications of utilisation at a distance	Yes, but no systematic approach reported	Investigators of the projects were included. Potential key users and other potential informants were interviewed to draft the model and understand contributions. Stakeholders were consulted to refine the model	Preliminary results were shared with stakeholders for feedback and validation. Once discrepancies were solved, results were shared with stakeholders	Not specified
S. Morton **Research Contribution Framework** [42] UK, 2015	–	Assessment of research impact using contribution analysis to explain the influence in policy and practice	• Final outcome • Policy or practice change • Capacity, knowledge and skill • Awareness, reaction • Engagement, participation • Activities and outputs • Inputs	No systematic approach reported	Semistructured interviews were conducted among project partners that took part in the process of the research (i.e., conferences and workshops) and research users (i.e., practitioners)	A thematic analysis approach was used to analyse the data collected and refine the framework. The themes that emerged were tested against other sources	A case study was used to illustrate how the Research Contribution Framework can assess impact

https://doi.org/10.1371/journal.pmed.1002370.t001

Methodological framework development

The included methodological frameworks varied in their development process, but there were some common approaches employed. Most included a literature review [2,5,7,8,31,33,36,37,40–46], although none of them used a recognised systematic method. Most also consulted with various stakeholders [3,8,29,31,33,35–38,43,44,46,47] but used differing methods to incorporate their views, including quantitative surveys [32,35,43,46], face-to-face interviews [7,29,33,35,37,42,43], telephone interviews [31,46], consultation [3,7,36], and focus groups [39,43]. A range of stakeholder groups were approached across the sample, including principal investigators [7,29,43], research end users [7,42,43], academics [3,8,39,40,43,46], award holders [43], experts [33,38,39], sponsors [33,39], project coordinators [32,42], and chief investigators [31,35]. However, some authors failed to identify the stakeholders involved in the development of their frameworks [2,5,34,41,45], making it difficult to assess their appropriateness. In addition, only 4 of the included papers reported using formal analytic methods to interpret stakeholder responses. These included the Canadian Academy of Health Sciences framework, which used conceptual cluster analysis [33]. The Research Contribution [42], Research Impact [29], and Primary Health Care & Information Service [31] used a thematic analysis approach. Finally, some authors went on to pilot their framework, which shaped refinements on the methodological frameworks until approval. Methods used to pilot the frameworks included a case study approach [2,3,30,32,33,36,40,42,44,45], contrasting results against available literature [29], the use of stakeholders' feedback [7], and assessment tools [35,46].

Major impact categories

1. Primary research-related impact. A number of methodological frameworks advocated the evaluation of 'research-related impact'. This encompassed content related to the generation of new knowledge, knowledge dissemination, capacity building, training, leadership, and the development of research networks. These outcomes were considered the direct or primary impacts of a research project, as these are often the first evidenced returns [30,62].

A number of subgroups were identified within this category, with frameworks supporting the collection of impact data across the following constructs: 'research and innovation outcomes'; 'dissemination and knowledge transfer'; 'capacity building, training, and leadership'; and 'academic collaborations, research networks, and data sharing'.

1.1. Research and innovation outcomes. Twenty of the 24 frameworks advocated the evaluation of 'research and innovation outcomes' [2,3,5,7,8,29–39,41,43,44,46]. This subgroup included the following metrics: number of publications; number of peer-reviewed articles (including journal impact factor); citation rates; requests for reprints, number of reviews, and meta-analysis; and new or changes in existing products (interventions or technology), patents, and research. Additionally, some frameworks also sought to gather information regarding 'methods/methodological contributions'. These advocated the collection of systematic reviews and appraisals in order to identify gaps in knowledge and determine whether the knowledge generated had been assessed before being put into practice [29].

1.2. Dissemination and knowledge transfer. Nineteen of the 24 frameworks advocated the assessment of 'dissemination and knowledge transfer' [2,3,5,7,29–32,34–43,46]. This comprised collection of the following information: number of conferences, seminars, workshops, and presentations; teaching output (i.e., number of lectures given to disseminate the research findings); number of reads for published articles; article download rate and number of journal webpage visits; and citations rates in nonjournal media such as newspapers and mass and social media (i.e., Twitter and blogs). Furthermore, this impact subgroup considered the measurement of research uptake and translatability and the adoption of research findings in

technological and clinical applications and by different fields. These can be measured through patents, clinical trials, and partnerships between industry and business, government and non-governmental organisations, and university research units and researchers [29].

1.3. Capacity building, training, and leadership. Fourteen of 24 frameworks suggested the evaluation of 'capacity building, training, and leadership' [2,3,5,8,29,31–35,39–41,43]. This involved collecting information regarding the number of doctoral and postdoctoral studentships (including those generated as a result of the research findings and those appointed to conduct the research), as well as the number of researchers and research-related staff involved in the research projects. In addition, authors advocated the collection of 'leadership' metrics, including the number of research projects managed and coordinated and the membership of boards and funding bodies, journal editorial boards, and advisory committees [29]. Additional metrics in this category included public recognition (number of fellowships and awards for significant research achievements), academic career advancement, and subsequent grants received. Lastly, the impact metric 'research system management' comprised the collection of information that can lead to preserving the health of the population, such as modifying research priorities, resource allocation strategies, and linking health research to other disciplines to maximise benefits [29].

1.4. Academic collaborations, research networks, and data sharing. Lastly, 10 of the 24 frameworks advocated the collection of impact data regarding 'academic collaborations (internal and external collaborations to complete a research project), research networks, and data sharing' [2,3,5,7,29,34,37,39,41,43].

2. Influence on policy making. Methodological frameworks addressing this major impact category focused on measurable improvements within a given knowledge base and on interactions between academics and policy makers, which may influence policy-making development and implementation. The returns generated in this impact category are generally considered as intermediate or midterm (1 to 3 years). These represent an important interim stage in the process towards the final expected impacts, such as quantifiable health improvements and economic benefits, without which policy change may not occur [30,62]. The following impact subgroups were identified within this category: 'type and nature of policy impact', 'level of policy making', and 'policy networks'.

2.1. Type and nature of policy impact. The most common impact subgroup, mentioned in 18 of the 24 frameworks, was 'type and nature of policy impact' [2,7,29–38,41–43,45–47]. Methodological frameworks addressing this subgroup stressed the importance of collecting information regarding the influence of research on policy (i.e., changes in practice or terminology). For instance, a project looking at trafficked adolescents and women (2003) influenced the WHO guidelines (2003) on ethics regarding this particular group [17,21,63].

2.2. Level of policy impact. Thirteen of 24 frameworks addressed aspects surrounding the need to record the 'level of policy impact' (international, national, or local) and the organisations within a level that were influenced (local policy makers, clinical commissioning groups, and health and wellbeing trusts) [2,5,8,29,31,34,38,41,43–47]. Authors considered it important to measure the 'level of policy impact' to provide evidence of collaboration, coordination, and efficiency within health organisations and between researchers and health organisations [29,31].

2.3. Policy networks. Five methodological frameworks highlighted the need to collect information regarding collaborative research with industry and staff movement between academia and industry [5,7,29,41,43]. A policy network emphasises the relationship between policy communities, researchers, and policy makers. This relationship can influence and lead to incremental changes in policy processes [62].

3. Health and health systems impact. A number of methodological frameworks advocated the measurement of impacts on health and healthcare systems across the following impact subgroups: 'quality of care and service delivering', 'evidence-based practice', 'improved information and health information management', 'cost containment and effectiveness', 'resource allocation', and 'health workforce'.

3.1. Quality of care and service delivery. Twelve of the 24 frameworks highlighted the importance of evaluating 'quality of care and service delivery' [2,5,8,29–31,33–36,41,47]. There were a number of suggested metrics that could be potentially used for this purpose, including health outcomes such as quality-adjusted life years (QALYs), patient-reported outcome measures (PROMs), patient satisfaction and experience surveys, and qualitative data on waiting times and service accessibility.

3.2. Evidence-based practice. 'Evidence-based practice', mentioned in 5 of the 24 frameworks, refers to making changes in clinical diagnosis, clinical practice, treatment decisions, or decision making based on research evidence [5,8,29,31,33]. The suggested metrics to demonstrate evidence-based practice were adoption of health technologies and research outcomes to improve the healthcare systems and inform policies and guidelines [29].

3.3. Improved information and health information management. This impact subcategory, mentioned in 5 of the 24 frameworks, refers to the influence of research on the provision of health services and management of the health system to prevent additional costs [5,29,33, 34,38]. Methodological frameworks advocated the collection of health system financial, nonfinancial (i.e., transport and sociopolitical implications), and insurance information in order to determine constraints within a health system.

3.4. Cost containment and cost-effectiveness. Six of the 24 frameworks advocated the subcategory 'cost containment and cost-effectiveness' [2,5,8,17,33,36]. 'Cost containment' comprised the collection of information regarding how research has influenced the provision and management of health services and its implication in healthcare resource allocation and use [29]. 'Cost-effectiveness' refers to information concerning economic evaluations to assess improvements in effectiveness and health outcomes—for instance, the cost-effectiveness (cost and health outcome benefits) assessment of introducing a new health technology to replace an older one [29,31,64].

3.5. Resource allocation. 'Resource allocation', mentioned in 6frameworks, can be measured through 2 impact metrics: new funding attributed to the intervention in question and equity while allocating resources, such as improved allocation of resources at an area level; better targeting, accessibility, and utilisation; and coverage of health services [2,5,29,31,45,47]. The allocation of resources and targeting can be measured through health services research reports, with the utilisation of health services measured by the probability of providing an intervention when needed, the probability of requiring it again in the future, and the probability of receiving an intervention based on previous experience [29,31].

3.6. Health workforce. Lastly, 'health workforce', present in 3 methodological frameworks, refers to the reduction in the days of work lost because of a particular illness [2,5,31].

4. Health-related and societal impact. Three subgroups were included in this category: 'health literacy'; 'health knowledge, attitudes, and behaviours'; and 'improved social equity, inclusion, or cohesion'.

4.1. Health knowledge, attitudes, and behaviours. Eight of the 24 frameworks suggested the assessment of 'health knowledge, attitudes, behaviours, and outcomes', which could be measured through the evaluation of levels of public engagement with science and research (e.g., National Health Service (NHS) Choices end-user visit rate) or by using focus groups to analyse changes in knowledge, attitudes, and behaviour among society [2,5,29,33–35,38,43].

4.2. Improved equity, inclusion, or cohesion and human rights. Other methodological frameworks, 4 of the 24, suggested capturing improvements in equity, inclusion, or cohesion and human rights. Authors suggested these could be using a resource like the United Nations Millennium Development Goals (MDGs) (superseded by Sustainable Development Goals [SDGs] in 2015) and human rights [29,33,34,38]. For instance, a cluster-randomised controlled trial in Nepal, which had female participants, has demonstrated the reduction of neonatal mortality through the introduction of maternity health care, distribution of delivery kits, and home visits. This illustrates how research can target vulnerable and disadvantaged groups. Additionally, this research has been introduced by the World Health Organisation to achieve the MDG 'improve maternal health' [16,29,65].

4.3. Health literacy. Some methodological frameworks, 3 of the 24, focused on tracking changes in the ability of patients to make informed healthcare decisions, reduce health risks, and improve quality of life, which were demonstrably linked to a particular programme of research [5,29,43]. For example, a systematic review showed that when HIV health literacy/knowledge is spread among people living with the condition, antiretroviral adherence and quality of life improve [66].

5. Broader economic impacts. Some methodological frameworks, 9 of 24, included aspects related to the broader economic impacts of health research—for example, the economic benefits emerging from the commercialisation of research outputs [2,5,29,31,33,35,36,38,67]. Suggested metrics included the amount of funding for research and development (R&D) that was competitively awarded by the NHS, medical charities, and overseas companies. Additional metrics were income from intellectual property, spillover effects (any secondary benefit gained as a repercussion of investing directly in a primary activity, i.e., the social and economic returns of investing on R&D) [33], patents granted, licences awarded and brought to the market, the development and sales of spinout companies, research contracts, and income from industry.

The benefits contained within the categories 'health and health systems impact', 'health-related and societal impact', and 'broader economic impacts' are considered the expected and final returns of the resources allocated in healthcare research [30,62]. These benefits commonly arise in the long term, beyond 5 years according to some authors, but there was a recognition that this could differ depending on the project and its associated research area [4].

Data synthesis

Five major impact categories were identified across the 24 included methodological frameworks: (1) 'primary research-related impact', (2) 'influence on policy making', (3) 'health and health systems impact', (4) 'health-related and societal impact', and (5) 'broader economic impact'. These major impact categories were further subdivided into 16 impact subgroups. The included publications proposed 80 different metrics to measure research impact. This impact typology synthesis is depicted in 'the impact matrix' (Fig 2 and Fig 3).

Commonality and differences across frameworks

The 'Research Impact Framework' and the 'Health Services Research Impact Framework' were the models that encompassed the largest number of the metrics extracted. The most dominant methodological framework was the Payback Framework; 7 other methodological framework models used the Payback Framework as a starting point for development [8,29,31–35]. Additional methodological frameworks that were commonly incorporated into other tools included the CIHR framework, the CAHS model, the AIHS framework, and the Exchange model [8,33,34,44]. The capture of 'research-related impact' was the most widely advocated concept

across methodological frameworks, illustrating the importance with which primary short-term impact outcomes were viewed by the included papers. Thus, measurement of impact via number of publications, citations, and peer-reviewed articles was the most common. 'Influence on policy making' was the predominant midterm impact category, specifically the subgroup 'type and nature of policy impact', in which frameworks advocated the measurement of (i) changes to legislation, regulations, and government policy; (ii) influence and involvement in decision-making processes; and (iii) changes to clinical or healthcare training, practice, or guidelines. Within more long-term impact measurement, the evaluations of changes in the 'quality of care and service delivery' were commonly advocated.

In light of the commonalities and differences among the methodological frameworks, the 'pathways to research impact' diagram (Fig 4) was developed to provide researchers, funders, and policy makers a more comprehensive and exhaustive way to measure healthcare research impact. The diagram has the advantage of assorting all the impact metrics proposed by previous frameworks and grouping them into different impact subgroups and categories. Prospectively, this global picture will help researchers, funders, and policy makers plan strategies to achieve multiple pathways to impact before carrying the research out. The analysis of the data extraction and construction of the impact matrix led to the development of the 'pathways to research impact' diagram (Fig 4). The diagram aims to provide an exhaustive and comprehensive way of tracing research impact by combining all the impact metrics presented by the different 24 frameworks, grouping those metrics into different impact subgroups, and grouping these into broader impact categories.

Discussion

This review has summarised existing methodological impact frameworks together for the first time using systematic methods (Fig 4). It allows researchers and funders to consider pathways to impact at the design stage of a study and to understand the elements and metrics that need to be considered to facilitate prospective assessment of impact. Users do not necessarily need to cover all the aspects of the methodological framework, as every research project can impact on different categories and subgroups. This review provides information that can assist researchers to better demonstrate impact, potentially increasing the likelihood of conducting impactful research and reducing research waste. Existing reviews have not presented a methodological framework that includes different pathways to impact, health impact categories, subgroups, and metrics in a single methodological framework.

Academic-orientated frameworks included in this review advocated the measurement of impact predominantly using so-called 'quantitative' metrics—for example, the number of peer-reviewed articles, journal impact factor, and citation rates. This may be because they are well-established measures, relatively easy to capture and objective, and are supported by research funding systems. However, these metrics primarily measure the dissemination of research finding rather than its impact [30,68]. Whilst it is true that wider dissemination, especially when delivered via world-leading international journals, may well lead eventually to changes in healthcare, this is by no means certain. For instance, case studies evaluated by Flinders University of Australia demonstrated that some research projects with non-peer-reviewed publications led to significant changes in health policy, whilst the studies with peer-reviewed publications did not result in any type of impact [68]. As a result, contemporary literature has tended to advocate the collection of information regarding a variety of different potential forms of impact alongside publication/citations metrics [2,3,5,7,8,29–47], as outlined in this review.

Time frame	Impact categories	Framework / Impact subgroups	Payback Framework	HTA Organisation Assessment Framework	Research Utilisation Ladder	The social impact of applied research	Exchange model	RAE	CIHR	RIF	PHC RIS	Logic Model	Meagher et al.	RQF
Short-term	1. Research-related impact	Research and innovation outcomes	✔		✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
		Dissemination and knowledge transfer	✔		✔	✔		✔		✔	✔	✔	✔	✔
		Capacity building, training and leadership	✔			✔		✔	✔	✔	✔		✔	
		Academic collaborations, research networks and data sharing	✔			✔		✔					✔	
Mid-term	2. Influencing and involvement in policy making	Level of policy- making	✔	✔	✔	✔	✔		✔	✔	✔		✔	✔
		Type and nature of policy impact	✔		✔	✔				✔	✔	✔	✔	✔
		Policy networks	✔		✔	✔				✔			✔	
Long-term	3. Health and health systems impact	Evidence-based practice							✔	✔	✔			
		Quality of care and service delivery	✔			✔			✔	✔	✔	✔		
		Improved information and health services management								✔				✔
		Cost containment and effectiveness	✔						✔	✔				
		Resource allocation	✔	✔						✔	✔			
		Health workforce	✔								✔			
	4. Health related and societal impacts	Health literacy								✔			✔	
		Health knowledge, attitudes, behaviour and outcomes	✔							✔			✔	✔
		Improved social equity, inclusion or cohesion								✔				✔
	5. Broader economic impacts		✔							✔	✔			✔

Fig 2. The impact matrix (1). CIHR, Canadian Institutes of Health Research; HTA, Health Technology Assessment; PHC RIS, Primary Health Care Research & Information Service; RAE, Research Assessment Exercise; RQF, Research Quality Framework.

https://doi.org/10.1371/journal.pmed.1002370.g002

The 2014 REF exercise adjusted UK university research funding allocation based on evidence of the wider impact of research (through case narrative studies and quantitative metrics), rather than simply according to the quality of research [12]. The intention was to ensure funds were directed to high-quality research that could demonstrate actual realised benefit. The inclusion of a mixed-method approach to the measurement of impact in the REF (narrative and quantitative metrics) reflects a widespread belief—expressed by the majority of authors of the included methodological frameworks in the review—that individual quantitative impact metrics (e.g., number of citations and publications) do not necessary capture the complexity of the relationships involved in a research project and may exclude measurement of specific aspects of the research pathway [10,12].

Many of the frameworks included in this review advocated the collection of a range of academic, societal, economic, and cultural impact metrics; this is consistent with recent recommendations from the Stern review [10]. However, a number of these metrics encounter research 'lag': i.e., the time between the point at which the research is conducted and when the actual benefits arise [69]. For instance, some cardiovascular research has taken up to 25 years to generate impact [70]. Likewise, the impact may not arise exclusively from a single piece of research. Different processes (such as networking interactions and knowledge and research

Time frame	Impact categories	Framework / Impact subgroups	Canavan et al.	CAHS	The Becker Model	REF	SIAMPI	The Health Services Research Impact Framework	AIHS	Contribution Mapping	Cohen et al.	Brueton et al.	IOM	Research Contribution Framework
Short-term	1. Research-related impact	Research and innovation outcomes		✔	✔	✔		✔	✔	✔	✔	✔	✔	
		Dissemination and knowledge transfer				✔	✔	✔	✔	✔	✔	✔	✔	
		Capacity building, training and leadership		✔	✔	✔	✔	✔	✔		✔		✔	✔
		Academic collaborations, research networks and data sharing				✔		✔	✔	✔		✔		
Mid-term	2. Influencing and involvement in policy making	Level of policy- making	✔					✔	✔					
		Type and nature of policy impact	✔	✔	✔				✔	✔	✔	✔	✔	✔
		Policy networks						✔		✔				
Long-term	3. Health and health systems impact	Evidence-based practice		✔				✔						
		Quality of care and service delivery		✔	✔			✔						
		Improved information and health services management		✔				✔						
		Cost containment and effectiveness		✔	✔			✔						
		Resource allocation		✔				✔						
		Health workforce						✔						
	4. Health related and societal impacts	Health literacy						✔						
		Health knowledge, attitudes, behaviour and outcomes		✔				✔	✔		✔			
		Improved social equity, inclusion or cohesion		✔					✔					
	5. Broader economic impacts			✔	✔			✔	✔		✔			

Fig 3. The impact matrix (2). AIHS, Alberta Innovates—Health Solutions; CAHS, Canadian Institutes of Health Research; IOM, Impact Oriented Monitoring; REF, Research Excellence Framework; SIAMPI, Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society.

https://doi.org/10.1371/journal.pmed.1002370.g003

translation) and multiple individuals and organisations are often involved [4,71]. Therefore, attributing the contribution made by each of the different actors involved in the process can be a challenge [4]. An additional problem associated to attribution is the lack of evidence to link research and impact. The outcomes of research may emerge slowly and be absorbed gradually. Consequently, it is difficult to determine the influence of research in the development of a new policy, practice, or guidelines [4,23].

A further problem is that impact evaluation is conducted 'ex post', after the research has concluded. Collecting information retrospectively can be an issue, as the data required might not be available. 'ex ante' assessment is vital for funding allocation, as it is necessary to determine the potential forthcoming impact before research is carried out [69]. Additionally, ex ante evaluation of potential benefit can overcome the issues regarding identifying and capturing evidence, which can be used in the future [4]. In order to conduct ex ante evaluation of potential benefit, some authors suggest the early involvement of policy makers in a research project coupled with a well-designed strategy of dissemination [40,69].

Providing an alternate view, the authors of methodological frameworks such as the SIAMPI, Contribution Mapping, Research Contribution, and the Exchange model suggest that the problems of attribution are a consequence of assigning the impact of research to a particular impact metric [7,40,42,44]. To address these issues, these authors propose focusing on the contribution of research through assessing the processes and interactions between

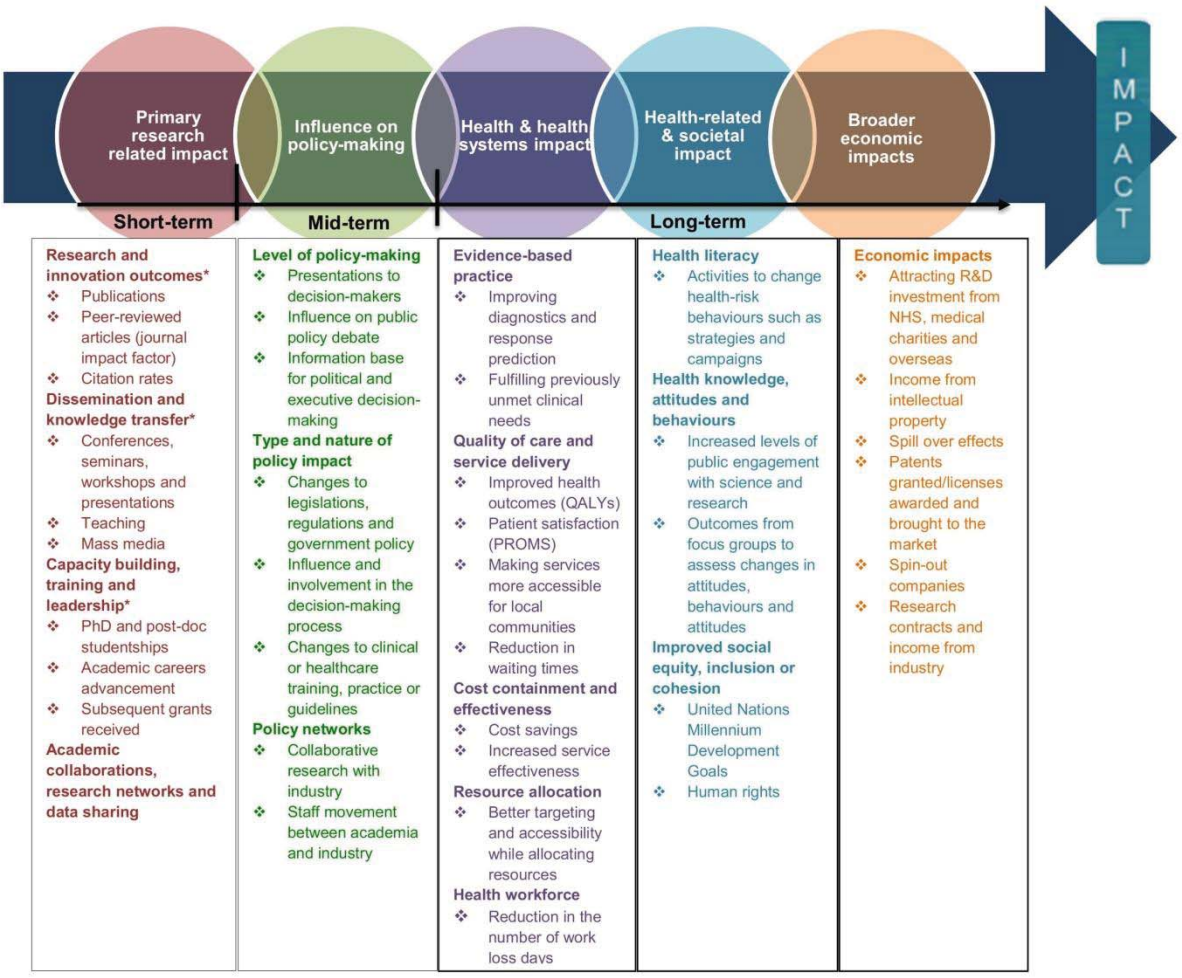


Fig 4. Pathways to research impact. NHS, National Health Service; PROM, patient-reported outcome measure; QALY, quality-adjusted life year; R&D, research and development.

https://doi.org/10.1371/journal.pmed.1002370.g004

stakeholders and researchers, which arguably take into consideration all the processes and actors involved in a research project [7,40,42,43]. Additionally, contributions highlight the importance of the interactions between stakeholders and researchers from an early stage in the research process, leading to a successful ex ante and ex post evaluation by setting expected impacts and determining how the research outcomes have been utilised, respectively [7,40,42,43]. However, contribution metrics are generally harder to measure in comparison to academic-orientated indicators [72].

Currently, there is a debate surrounding the optimal methodological impact framework, and no tool has proven superior to another. The most appropriate methodological framework

for a given study will likely depend on stakeholder needs, as each employs different methodologies to assess research impact [4,37,41]. This review allows researchers to select individual existing methodological framework components to create a bespoke tool with which to facilitate optimal study design and maximise the potential for impact depending on the characteristic of their study (Fig 2 and Fig 3). For instance, if researchers are interested in assessing how influential their research is on policy making, perhaps considering a suite of the appropriate metrics drawn from multiple methodological frameworks may provide a more comprehensive method than adopting a single methodological framework. In addition, research teams may wish to use a multidimensional approach to methodological framework development, adopting existing narratives and quantitative metrics, as well as elements from contribution frameworks. This approach would arguably present a more comprehensive method of impact assessment; however, further research is warranted to determine its effectiveness [4,69,72,73].

Finally, it became clear during this review that the included methodological frameworks had been constructed using varied methodological processes. At present, there are no guidelines or consensus around the optimal pathway that should be followed to develop a robust methodological framework. The authors believe this is an area that should be addressed by the research community, to ensure future frameworks are developed using best-practice methodology.

For instance, the Payback Framework drew upon a literature review and was refined through a case study approach. Arguably, this approach could be considered inferior to other methods that involved extensive stakeholder involvement, such as the CIHR framework [8]. Nonetheless, 7 methodological frameworks were developed based upon the Payback Framework [8,29,31–35].

Limitations

The present review is the first to summarise systematically existing impact methodological frameworks and metrics. The main limitation is that 50% of the included publications were found through methods other than bibliographic databases searching, indicating poor indexing. Therefore, some relevant articles may not have been included in this review if they failed to indicate the inclusion of a methodological impact framework in their title/abstract. We did, however, make every effort to try to find these potentially hard-to-reach publications, e.g., through forwards/backwards citation searching, hand searching reference lists, and expert communication. Additionally, this review only extracted information regarding the methodology followed to develop each framework from the main publication source or framework webpage. Therefore, further evaluations may not have been included, as they are beyond the scope of the current paper. A further limitation was that although our search strategy did not include language restrictions, we did not specifically search non-English language databases. Thus, we may have failed to identify potentially relevant methodological frameworks that were developed in a non-English language setting.

Conclusion

In conclusion, the measurement of research impact is an essential exercise to help direct the allocation of limited research resources, to maximise benefit, and to help minimise research waste. This review provides a collective summary of existing methodological impact frameworks and metrics, which funders may use to inform the measurement of research impact and researchers may use to inform study design decisions aimed at maximising the short-, medium-, and long-term impact of their research.

Supporting information

S1 Appendix. Search strategy.
(TIF)

S1 PRISMA Checklist. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.
(PDF)

Acknowledgments

We would also like to thank Mrs Susan Bayliss, Information Specialist, University of Birmingham, and Mrs Karen Biddle, Research Secretary, University of Birmingham.

Author Contributions

Conceptualization: Samantha Cruz Rivera, Derek G. Kyte, Melanie J. Calvert.

Data curation: Samantha Cruz Rivera, Olalekan Lee Aiyegbusi.

Formal analysis: Samantha Cruz Rivera, Derek G. Kyte, Olalekan Lee Aiyegbusi, Thomas J. Keeley, Melanie J. Calvert.

Funding acquisition: Samantha Cruz Rivera, Derek G. Kyte, Melanie J. Calvert.

Investigation: Samantha Cruz Rivera.

Methodology: Samantha Cruz Rivera, Derek G. Kyte, Olalekan Lee Aiyegbusi, Thomas J. Keeley, Melanie J. Calvert.

Project administration: Derek G. Kyte, Melanie J. Calvert.

Supervision: Derek G. Kyte, Thomas J. Keeley, Melanie J. Calvert.

Validation: Samantha Cruz Rivera, Derek G. Kyte, Olalekan Lee Aiyegbusi, Thomas J. Keeley, Melanie J. Calvert.

Writing – original draft: Samantha Cruz Rivera.

Writing – review & editing: Samantha Cruz Rivera, Derek G. Kyte, Olalekan Lee Aiyegbusi, Thomas J. Keeley, Melanie J. Calvert.

References

1. Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis J, et al. Biomedical research: increasing value, reducing waste. Lancet. 2014; 383(9912):101–4. https://doi.org/10.1016/S0140-6736(13)62329-6 PMID: 24411643
2. Buxton M, Hanney S. How can payback from health services research be assessed? J Health Serv Res Policy. 1996; 1(1):35–43. PMID: 10180843.
3. HEFCE. REF 2014: Assessment framework and guidance on submissions 2011 [cited 2016 15 Feb]. Available from: http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf.
4. Penfield T, Baker MJ, Scoble R, Wykes MC. Assessment, evaluations, and definitions of research impact: A review. Res Eval. 2014; 23(1):21–32. https://doi.org/10.1093/reseval/rvt021
5. Buykx P, Humphreys J, Wakerman J, Perkins D, Lyle D, McGrail M, et al. Making evidence count: A framework to monitor the impact of health services research. Aust J Rural Health. 2012; 20(2):51–8. https://doi.org/10.1111/j.1440-1584.2012.01256.x PMID: 22435764
6. Martin BR. The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? Res Eval. 2011; 20(3):247–54. https://doi.org/10.3152/095820211x13118583635693

7. Kok MO, Schuit AJ. Contribution mapping: a method for mapping the contribution of research to enhance its impact. Health Res Policy Syst. 2012; 10. https://doi.org/10.1186/1478-4505-10-21 PMID: 22748169
8. Canadian Institutes of Health Research. Developing a CIHR framework to measure the impact of health research 2005 [cited 2016 26 Feb]. Available from: http://publications.gc.ca/collections/Collection/MR21-65-2005E.pdf.
9. HEFCE. HEFCE allocates £3.97 billion to universities and colleges in England for 2015–1 2015. Available from: http://www.hefce.ac.uk/news/newsarchive/2015/Name,103785,en.html.
10. Stern N. Building on Success and Learning from Experience—An Independent Review of the Research Excellence Framework 2016 [cited 2016 05 Aug]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/541338/ind-16-9-ref-stern-review.pdf.
11. Matthews D. REF sceptic to lead review into research assessment: Times Higher Education; 2015 [cited 2016 21 Apr]. Available from: https://www.timeshighereducation.com/news/ref-sceptic-lead-review-research-assessment.
12. HEFCE. The Metric Tide—Report of the Independent Review of the Role of Metrics in Research Assessment and Management 2015 [cited 2016 11 Aug]. Available from: http://www.hefce.ac.uk/media/HEFCE,2014/Content/Pubs/Independentresearch/2015/The,Metric,Tide/2015_metric_tide.pdf.
13. Bainbridge R, Tsey K, McCalman J, Kinchin I, Saunders V, Lui FW, et al. No one's discussing the elephant in the room: contemplating questions of research impact and benefit in Aboriginal and Torres Strait Islander Australian health research. BMC Public Health. 2015; 15. https://doi.org/10.1186/s12889-015-2052-3 PMID: 26202429
14. LSE Public Policy Group. Maximizing the impacts of your research: A handbook for social scientists. http://www.lse.ac.uk/government/research/resgroups/LSEPublicPolicy/Docs/LSE_Impact_Handbook_April_2011.pdf. London: LSE; 2011.
15. HEFCE. Consultation on the second Research Excellence Framework. 2016.
16. Ovseiko PV, Oancea A, Buchan AM. Assessing research impact in academic clinical medicine: a study using Research Excellence Framework pilot impact indicators. BMC Health Serv Res. 2012; 12. https://doi.org/10.1186/1472-6963-12-478 PMID: 23259467
17. Kuruvilla S, Mays N, Walt G. Describing the impact of health services and policy research. J Health Serv Res Policy. 2007; 12. https://doi.org/10.1258/135581907780318374 PMID: 17411504
18. Merriam-Webster Dictionary 2017. Available from: https://www.merriam-webster.com/dictionary/methodology.
19. Oxford Dictionaries—pathway 2016 [cited 2016 19 June]. Available from: http://www.oxforddictionaries.com/definition/english/pathway.
20. Oxford Dictionaries—metric 2016 [cited 2016 15 Sep]. Available from: https://en.oxforddictionaries.com/definition/metric.
21. WHO. WHO Ethical and Safety Guidelines for Interviewing Trafficked Women 2003 [cited 2016 29 July]. Available from: http://www.who.int/mip/2003/other_documents/en/Ethical_Safety-GWH.pdf.
22. Greenhalgh T, Raftery J, Hanney S, Glover M. Research impact: a narrative review. BMC Med. 2016; 14(1):1–16. https://doi.org/10.1186/s12916-016-0620-8 PMID: 27211576
23. Banzi R, Moja L, Pistotti V, Facchini A, Liberati A. Conceptual frameworks and empirical approaches used to assess the impact of health research: an overview of reviews. Health Res Policy Syst. 2011; 9. https://doi.org/10.1186/1478-4505-9-26 PMID: 21702930
24. Yazdizadeh B, Majdzadeh R, Salmasian H. Systematic review of methods for evaluating healthcare research economic impact. Health Res Policy Syst. 2010; 8. https://doi.org/10.1186/1478-4505-8-6 PMID: 20196839
25. Milat AJ, Bauman AE, Redman S. A narrative review of research impact assessment models and methods. Health Res Policy Syst. 2015; 13. https://doi.org/10.1186/s12961-015-0003-1 PMID: 25884944
26. Bornmann L. What is societal impact of research and how can it be assessed? a literature survey. American Society for Information Science and Technology. 2013; 64(2):217–33. https://doi.org/10.1002/asi.22803
27. Patel VM, Ashrafian H, Ahmed K, Arora S, Jiwan S, Nicholson JK, et al. How has healthcare research performance been assessed? A systematic review. J R Soc Med. 2011; 104(6):251–61. https://doi.org/10.1258/jrsm.2011.110005 PMID: 21659400
28. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009; 151(4):264–9. PMID: 19622511

29. Kuruvilla S, Mays N, Pleasant A, Walt G. Describing the impact of health research: a Research Impact Framework. BMC Health Serv Res. 2006; 6:134. PMID: 17049092; PubMed Central PMCID: PMCPMC1635046. https://doi.org/10.1186/1472-6963-6-134

30. Weiss AP. Measuring the impact of medical research: moving from outputs to outcomes. Am J Psychiatry. 2007; 164(2):206–14. PMID: 17267781. https://doi.org/10.1176/ajp.2007.164.2.206

31. Kalucy L, et al. Primary Health Care Research Impact Project: Final Report Stage 1 Adelaide: Primary Health Care Research & Information Service; 2007 [cited 2016 26 Feb]. Available from: http://www.phcris.org.au/phplib/filedownload.php?file=/elib/lib/downloaded_files/publications/pdfs/phcris_pub_3338.pdf.

32. Guinea J, Sela E, Gómez-Núñez AJ, Mangwende T, Ambali A, Ngum N, et al. Impact oriented monitoring: A new methodology for monitoring and evaluation of international public health research projects. Res Eval. 2015; 24(2):131–45. https://doi.org/10.1093/reseval/rvu034

33. Canadian Academy of Health Sciences. Making an impact—A preferred framework and indicators to measure returns on investment in health research 2009 [cited 2016 26 Feb]. Available from: http://www.cahs-acss.ca/wp-content/uploads/2011/09/ROI_FullReport.pdf.

34. Graham KER, Chorzempa HL, Valentine PA, Magnan J. Evaluating health research impact: Development and implementation of the Alberta Innovates–Health Solutions impact framework. Res Eval. 2012; 21(5):354–67. https://doi.org/10.1093/reseval/rvs027

35. Cohen G, Schroeder J, Newson R, King L, Rychetnik L, Milat AJ, et al. Does health intervention research have real world policy and practice impacts: testing a new impact assessment tool. Health Res Policy Syst. 2014; 13(1):3–1p. https://doi.org/10.1186/1478-4505-13-3 PMID: 109774007. Language: English. Entry Date: 20150923. Revision Date: 20150923. Publication Type: Journal Article.

36. Sarli CC, Dubinsky EK, Holmes KL. Beyond citation analysis: a model for assessment of research impact. J Med Libr Assoc. 2010; 98(1):17–23. doi: 10.3163/1536-5050.98.1.008. PMID: 20098647

37. Brueton VC, Vale CL, Choodari-Oskooei B, Jinks R, Tierney JF. Measuring the impact of methodological research: a framework and methods to identify evidence of impact. Trials. 2014; 15(1):464–1p. https://doi.org/10.1186/1745-6215-15-464 PMID: 109768210. Language: English. Entry Date: 20150313. Revision Date: 20150923. Publication Type: Journal Article.

38. Donovan C. The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research. New Directions for Evaluation. 2008; 2008(118):47–60. https://doi.org/10.1002/ev.260

39. HEFCE. RAE 2008—Guidance in submissions 2005 [cited 2016 15 Feb]. Available from: http://www.rae.ac.uk/pubs/2005/03/rae0305.pdf.

40. Spaapen J, van Drooge L. Introducing 'productive interactions' in social impact assessment. Res Eval. 2011; 20(3):211–8. https://doi.org/10.3152/095820211x12941371876742

41. Royal Netherlands Academy of Arts and Sciences. The societal impact of applied health research—Towards a quality assessment system 2002 [cited 2016 29 Feb]. Available from: https://www.knaw.nl/en/news/publications/the-societal-impact-of-applied-health-research/@@download/pdf_file/20021098.pdf.

42. Morton S. Progressing research impact assessment: A 'contributions' approach. Res Eval. 2015; 24(4):405–19. https://doi.org/10.1093/reseval/rvv016

43. Meagher L, Lyall C, Nutley S. Flows of knowledge, expertise and influence: a method for assessing policy and practice impacts from social science research. Res Eval. 2008; 17(3):163–73. https://doi.org/10.3152/095820208x331720

44. Lavis J, Ross S, McLeod C, Gildiner A. Measuring the impact of health research. J Health Serv Res Policy. 2003; 8(3):165–70. https://doi.org/10.1258/135581903322029520 PMID: 12869343.

45. Canavan J, Gillen A, Shaw A. Measuring research impact: developing practical and cost-effective approaches. Evidence & Policy. 2009; 5(2):167–77.

46. Landry R, Amara N, Lamari M. Climbing the ladder of research utilization—Evidence from social science research. Science Communication. 2001; 22(4):396–422. https://doi.org/10.1177/1075547001022004003

47. Jacob R, McGregor M. Assessing the impact of health technology assessment. Int J Technol Assess Health Care. 1997; 13(1):68–80. PMID: 9119625

48. Weiss CH. Using social research in public policy making: Lexington Books; 1977.

49. Weiss CH. The Many Meanings of Research Utilization. Public Adm Rev. 1979; 39(5):426–31. https://doi.org/10.2307/3109916

50. Kogan M, Henkel M. Government and research: the Rothschild experiment in a government department: Heinemann Educational Books; 1983.

51. Thomas P. The Aims and Outcomes of Social Policy Research. Croom Helm; 1985.

52. Bulmer M. Social Science Research and Government: Comparative Essays on Britain and the United States: Cambridge University Press; 2010.

53. Booth T. Developing Policy Research. Aldershot, Gower1988.

54. Yin R, Moore G. Lessons on the utilization of research from nine case experiences in the natural hazards field. Knowledge in Society. 1988; 1.

55. Kalucy L, et al Exploring the impact of primary health care research Stage 2 Primary Health Care Research Impact Project Adelaide: Primary Health Care Research & Information Service (PHCRIS); 2009 [cited 2016 26 Feb]. Available from: http://www.phcris.org.au/phplib/filedownload.php?file=/elib/lib/downloaded_files/publications/pdfs/phcris_pub_8108.pdf.

56. CHSRF. Canadian Health Services Research Foundation 2000. Health Services Research and Evidence-based Decision Making [cited 2016 February]. Available from: http://www.cfhi-fcass.ca/migrated/pdf/mythbusters/EBDM_e.pdf.

57. Knott J, Wildavsky A. If dissemination is the answer, what is the problem? Knowledge: Creation, Diffusion, Utilization. 1980; 1.

58. W.K. Kellogg Foundation. Logic Model Development Guide 2004 [cited 2016 19 July]. Available from: http://www.smartgivers.org/uploads/logicmodelguidepdf.pdf.

59. United Way of America. Measuring Program Outcomes: A Practical Approach 1996 [cited 2016 19 July]. Available from: https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf.

60. Nutley S, Percy-Smith J and Solesbury W. Models of research impact: a cross sector review of literature and practice. London: Learning and Skills Research Centre 2003.

61. Spaapen J, van Drooge L. SIAMPI final report [cited 2017 Jan]. Available from: http://www.siampi.eu/Content/SIAMPI_Final%20report.pdf.

62. Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. Health Res Policy Syst. 2003; 1(1):1–28. https://doi.org/10.1186/1478-4505-1-2

63. LSHTM. The Health Risks and Consequences of Trafficking in Women and Adolescents—Findings from a European Study 2003 [cited 2016 29 July]. Available from: http://www.oas.org/atip/global%20reports/zimmerman%20tip%20health.pdf.

64. Drummond M, Cooke J, Walley T. Economic evaluation under managed competition: evidence from the U.K. Soc Sci Med. 1997; 45.

65. Manandhar DS, Osrin D, Shrestha BP, Mesko N, Morrison J, Tumbahangphe KM, et al. Effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster-randomised controlled trial. Lancet. 364(9438):970–9. https://doi.org/10.1016/S0140-6736(04)17021-9 PMID: 15364188

66. Wawrzyniak AJ, Ownby RL, McCoy K, Waldrop-Valverde D. Health Literacy: Impact on the Health of HIV-Infected Individuals. Curr HIV/AIDS Rep. 2013; 10(4):295–304. doi: 10.1007/s11904-013-0178-4. PMID: 24222474

67. Graham KER, Chorzempa HL, Valentine PA, Magnan J. Evaluating health research impact: Development and implementation of the Alberta Innovates–Health Solutions impact framework. Res Eval. 2012; 21. https://doi.org/10.1093/reseval/rvs027

68. van Driel ML, Maier M, De Maeseneer J. Measuring the impact of family medicine research: scientific citations or societal impact? Fam Pract. 2007; 24(5):401–2. PMID: 17998264. https://doi.org/10.1093/fampra/cmm061

69. Hansen J, Muscat N, Keskimaki L. Measuring and improving the societal impact of health care research. Eurohealth. 2013; 19(3):32–5.

70. Russell G. Response to second HEFCE consultation on the Research Excellence Framework 2009 [cited 2016 04 Apr]. Available from: http://russellgroup.ac.uk/media/5262/ref-consultation-response-final-dec09.pdf.

71. Cohen G, Schroeder J, Newson R, King L, Rychetnik L, Milat AJ, et al. Does health intervention research have real world policy and practice impacts: testing a new impact assessment tool. Health Res Policy Syst. 2015; 13. https://doi.org/10.1186/1478-4505-13-3 PMID: 25552272

72. Bornmann L. Measuring the societal impact of research. Research is less and less assessed on scientific impact alone-we should aim to quantify the increasingly important contributions of science to society. EMBO Reports. 2012; 13(8):673–6. PMID: 22777497. https://doi.org/10.1038/embor.2012.99

73. Erno-Kjolhede E, Hansson F. Measuring research performance during a changing relationship between science and society. Res Eval. 2011; 20(2):131–43. https://doi.org/10.3152/095820211x12941371876544