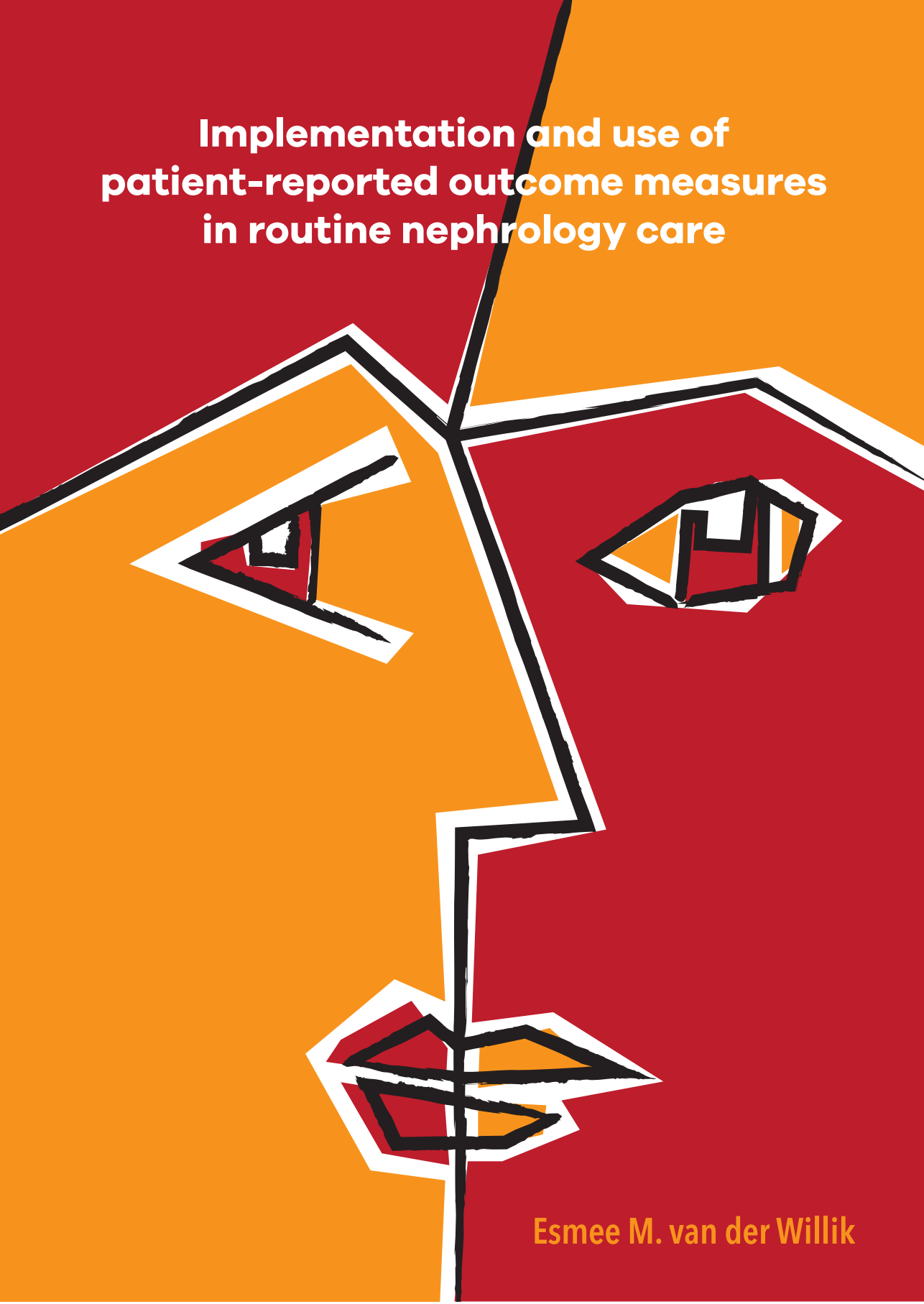
Implementation and use of patient-reported outcome measures in routine nephrology care

Willik, E.M. van der

Implementation and use of patient-reported outcome measures in routine nephrology care



Esmee M. van der Willik

Implementation and use of patient-reported outcome measures in routine nephrology care

Esmee M. van der Willik

Implementation and use of patient-reported outcome measures in routine nephrology care

Esmee M. van der Willik

PhD Dissertation, Leiden University Medical Center, The Netherlands

Cover design and layout: Simone van der Meer

Printed by: Gildeprint

ISBN: 978-94-6419-793-8

Esmee M. van der Willik and other involved researchers declare no conflicts of interest. The work in this dissertation was supported by unrestricted grants from the Dutch Kidney Foundation (A1D1P04 and 18SWO03), the KIDZ program of the Dutch Patient Federation, the Dutch Health Insurers Association and CSL Vifor. Printing of this dissertation was partially sponsored by Chiesi Pharmaceuticals B.V. and CSL Vifor.

Implementation and use of patient-reported outcome measures in routine nephrology care

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op donderdag 8 juni 2023
klokke 15:00 uur

door

Elisabeth Maria van der Willik
geboren te Alkemade
in 1992

Promotores
Prof. Dr. F.W. Dekker
Prof. Dr. M.H. Hemmelder (Maastricht University)

Co-promotor
Dr. Y. Meuleman

Leden promotiecommissie
Prof. Dr. W.J.W. Bos
Prof. Dr. J.I. Rotmans
Prof. Dr. M.W.J.M. Wouters (Antoni van Leeuwenhoek hospital)
Prof. Dr. N.F. de Keizer (Amsterdam University Medical Centers)

Table of Contents

CHAPTER 1 8
General introduction and thesis outline

PART 1. IMPLEMENTATION OF PROMS INTO ROUTINE NEPHROLOGY CARE

CHAPTER 2 24
Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease – a four-phase mixed methods study
BMC Nephrology, 2019

CHAPTER 3 52
Validity and reliability of Patient-Reported Outcomes Measurement Information System (PROMIS®) using Computerized Adaptive Testing (CAT) in patients with advanced chronic kidney disease
Nephrology Dialysis Transplantation, 2022

CHAPTER 4 80
Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures (PROMs)
Clinical Kidney Journal, 2020

PART 2: USE OF PROMS AT POPULATION LEVEL AND IN INDIVIDUAL PATIENTS IN ROUTINE NEPHROLOGY CARE

CHAPTER 5 104
Funnel plots of patient-reported outcomes (PROs) to evaluate healthcare quality: basic principles, pitfalls and considerations
Nephrology (Carlton), 2021

CHAPTER 6 128
Itching in dialysis patients: impact on health-related quality of life and interactions with sleep problems and psychological symptoms – results from the RENINE/PROMs registry
Nephrology Dialysis Transplantation, 2022

CHAPTER 7 164
Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time
Nephrology (Carlton), 2021

CHAPTER 8 184
Discussing results of patient-reported outcome measures (PROMs) between patients and healthcare professionals in routine dialysis care: a qualitative study
BMJ Open, 2022

CHAPTER 9 218
Summary and general discussion

APPENDICES 242
Dutch summary (Nederlandse samenvatting)
Acknowledgements (Dankwoord)
Curriculum Vitae
PhD portfolio
List of publications

1

General introduction and thesis outline

1

General introduction and thesis outline

Healthcare is shifting towards a more person-centred approach.[1-3] More attention is paid to the patients' perspective, aiming at a personalised and holistic treatment that fits the patient's preferences and needs. Insight into patient-reported outcomes (PROs), such as health-related quality of life (HRQOL) and symptom burden, is therefore becoming increasingly important in healthcare.[4, 5] Patient-reported outcome measures (PROMs) systematically assess such outcomes and can facilitate the process of adapting to what is important to the patient.[4-7] But, how to integrate PROMs into a routine care setting and how to use PROMs to achieve this personalised and holistic treatment? This dissertation provides insight into and practical knowledge of the implementation and use of PROMs in routine nephrology care.

Chronic kidney disease

Chronic kidney disease (CKD) is a progressive condition characterized by a decreased kidney function based on a glomerular filtration rate (GFR) of less than 60 mL/min per 1.73 m², or markers of kidney damage, such as albuminuria, present for at least 3 months.[8, 9] Globally, the prevalence of CKD is estimated around 10%[10, 11], and is expected to further increase due to the aging population and the increasing number of people with diabetes and hypertension.[12] Worldwide, about 0.5% of the people has advanced CKD (GFR < 30 mL/min per 1.73 m²) and 0.1% has kidney failure (GFR < 15 mL/min per 1.73 m²).[8, 11] Patients with kidney failure have the choice to receive kidney replacement therapy (KRT) to prolong life, or comprehensive conservative care, which aims at maintaining HRQOL, optimal symptom management and slowing down disease progression.[13, 14] There are two general types of KRT, namely kidney transplantation or dialysis treatment (e.g. peritoneal dialysis or haemodialysis).[8, 14] Peritoneal dialysis treatments are every day or night and are performed from home by the patient (and any caregivers) or with help of a machine.[15] Haemodialysis treatments are usually 3 times a week for approximately 3-5 hours per dialysis session, performed at the dialysis centre or at home.[15] Patients need on average 5-7 hours to fully recover after each haemodialysis session[16, 17], underlining the invasiveness and high impact on people's life.[18] The choice for which treatment fits the patient best is generally based on availability of treatment (e.g. kidney donor and dialysis options at home or in a centre nearby), clinical characteristics (e.g. the patient's health status, medical risks and potential health benefits), and the patient's characteristics and his values, preferences and needs (e.g. what someone finds important in life).[8, 13, 18, 19]

Outcomes in patients with chronic kidney disease

CKD is a growing public health problem causing a high disease burden and healthcare costs.[12, 18, 20, 21] Advanced CKD is associated with a high cardiovascular morbidity, increased mortality and hospitalizations, and has a major impact on people's life.[9, 18, 22] Patients with advanced CKD experience numerous physical and emotional symptoms, such as fatigue, itching, muscle cramps, sleep problems, sexual problems and depressive symptoms, which have a major impact on their HRQOL.[23-26]

Nephrology care traditionally focusses on clinical measures, such as mortality, laboratory values and blood pressure. Although PROs, like HRQOL and symptom burden have been regarded as highly important by patients and healthcare professionals[27-31], these outcomes often remain unknown, undiscussed and undertreated in regular practice.[25, 32] This is partly because patients do not share everything by themselves, for instance because some topics may be difficult to talk about, or because patients assume that their symptoms cannot be treated, or are not related to their CKD or treatment for CKD.[33-35] Additionally, it may be challenging for healthcare professionals to inquire about the wide range of symptoms and needs that patients experience, for example due to time or intervention limitations.[32, 36]

Last decade, healthcare is shifting towards a more person-centred approach, including nephrology care.[1-3, 37] In addition to the traditional clinical measures, there is a stronger focus on the patient's perspective and outcomes that matter to patients.[5, 37, 38] Systematic assessment of PROs can solve the under-recognition of outcomes like HRQOL and symptom burden, and support this personalized and holistic treatment approach.[25, 39] PROs consider experienced health and should thus be assessed from the patient's perspective. PROs can be systematically assessed using PROMs.[40-42]

Patient-reported outcome measures

PROMs are questionnaires that assess aspects of patients' perceived health, such as HRQOL and symptom burden. PROMs are reported by the patients themselves; support may be offered when filling in PROMs, as long as responses reflect the patient's perspective.[40-42]

Many different PROMs exist, using various measurement methods and characteristics. For example, PROMs are often classified as either generic or specific for a certain disease, condition or treatment.[40, 41] Generic PROMs include widely relevant health aspects and are particularly suitable for heterogeneous populations (e.g. multimorbid populations like CKD), and enable comparisons across populations and treatments.[40, 41] A specific PROM is tailored to a certain disease, condi-

tion or treatment, and is particularly suitable for comparisons within a population, as they are usually better able to detect smaller or specific changes.[40, 41] Furthermore, PROMs can be fixed (i.e. nonadaptive) or adaptive. Traditional PROMs are fixed, meaning that it contains the same questions and order for any patient at any timepoint. Adaptive PROMs are relatively novel in healthcare and make use of computerized adaptive tests (CATs), in which the next question is selected based on the answer to previous questions, adapting to the patient's ability.[43] An example of an adaptive PROM is the Patient-Reported Outcomes Measurement Information System (PROMIS).[43] Moreover, PROMs can vary for instance in the underlying measurement method, number of questions, recall period, scoring scale and method, and reference standard.[44] The features of the PROM influence the interpretation of the PROM-scores.[42] In contrast to well-known clinical measures such as blood pressure, healthcare professionals, patients and researchers are often not yet familiar with the interpretation of PROM-scores. Understanding of the PROMs and the interpretation of its PROM-scores are needed for optimal use in clinical practice.

Many different PROMs are available[41, 44-47] and which PROM is suitable for clinical practice does not only depend on the characteristics (e.g. generic or specific, measurement method and scoring) and psychometric quality (e.g. validity and reliability) of the available PROMs, but also on the population and clinical setting.[42] For example: the purpose of measuring the PRO (e.g. use during consultations), the setting (e.g. opportunity to integrate into workflow) and the homogeneity of the population (e.g. variation in experienced health or digital skills). Hence, it is important to deliberately select PROMs, so that they fit routine practice. For nephrology care, the 12-item Short-Form Health Survey (SF-12) to assess generic HRQOL was recommended by an European expert consensus group.[48] Moreover, they underlined the importance of measuring symptom burden in addition to HRQOL, but no consensus was reached on the preferred PROM to assess symptom burden.[48]

The potential of using PROMs in healthcare

PROMs have the potential to contribute to a more person-centred approach.[3, 4, 49-51] PROMs can provide insight into and a more complete picture of how the patient is really doing by incorporating the patient's perspective, complementary to traditional clinical measures. Hence, using PROMs may enhance shared decision making and facilitate personalized treatment.[6, 7, 50, 51] Moreover, literature suggests that the use of PROMs may even result in better health outcomes, for example better symptom management, less hospitalizations and better HRQOL.[5, 52] However,

the majority of existing literature is theoretical and little research has been done in nephrology care.[6, 50, 53] Therefore, research in real-world nephrology care is needed to examine these potential benefits of using PROMs.

Theoretically, the use of PROMs can contribute to clinical practice at multiple levels: at individual patient-level and at aggregated population-level. For example, individual PROM-results can support shared decision making by facilitating patient-professional communication and discussion about patients' experiences and needs.[6, 7, 50, 51] Aggregated PROM-results can inform patients (and healthcare professionals) about prognosis, treatment and factors influencing PROs.[6] In addition, aggregated PROM-results can be used to evaluate healthcare quality.[6, 54, 55] Ideally, PROMs are integrated into routine care in such a way that it provides valuable information at both the individual patient-level and the aggregated population-level.[55, 56] These different purposes must be taken into account and require a structured approach in the implementation of PROMs into routine care.

Implementation of PROMs into routine nephrology care

In nephrology, the importance of PROs is widely recognized and first steps are taken to identify outcomes that matter to patients by the Standardised Outcomes in Nephrology (SONG) initiative[57] and by the International Consortium for Health Outcomes Measurement (ICHOM).[58] However, PROMs have not been widely implemented yet into routine nephrology care.[48, 59] A few examples exist and show that implementation can be challenging, for instance reaching adequate response rates, incorporation into the workflow, and struggles due to lack of knowledge on how to interpret, discuss and intervene on PROM-results.[60-63] Furthermore, literature suggests that the incorporation of PROMs requires engagement from all people involved: patients, healthcare professionals, researchers and policy makers.[61, 62, 64] Patients receiving dialysis treatment have frequent healthcare encounters and dialysis care has a strong infrastructure, which provides a good basis for reaching all people involved and implementation into the existing workflow.[65]

In the Netherlands, we establish a nationwide project to develop and implement PROMs into nephrology care (PROMs-NNL), in close collaboration with all relevant stakeholders: patients (Dutch Kidney Patients Association; NVN), healthcare professionals (Dutch Federation for Nephrology; NFN), researchers (Leiden University Medical Centre; LUMC) and the healthcare quality institute of nephrology care (Nefrovisie Foundation). PROMs will be part of the data collection in RENINE, the Dutch renal registry (www.renine.nl), to ensure nationwide support and minimal burden for healthcare centres.[48, 66] PROMs will be firstly introduced within routine dialysis care, given the relatively easy to reach population and suitable clinical set-

1

ting.[65] The PROMs-NNL project comprises the following four steps to implement PROMs into routine nephrology care in the Netherlands (Figure 1):

Step 1: determine information about which PROs is important and for what purpose.
Step 2: select the best suitable PROMs to measure these PROs, taking into account the aim and setting.
Step 3: pilot test the use of PROMs in clinical practice; are these PROMs suitable and what are feasible methods to collect and provide feedback on PROM-results?
Step 4: make adjustments based on the lessons learned and implement PROMs into routine care at national level. Implementation involves using, evaluating and adjusting iteratively to achieve optimal use of PROMs.

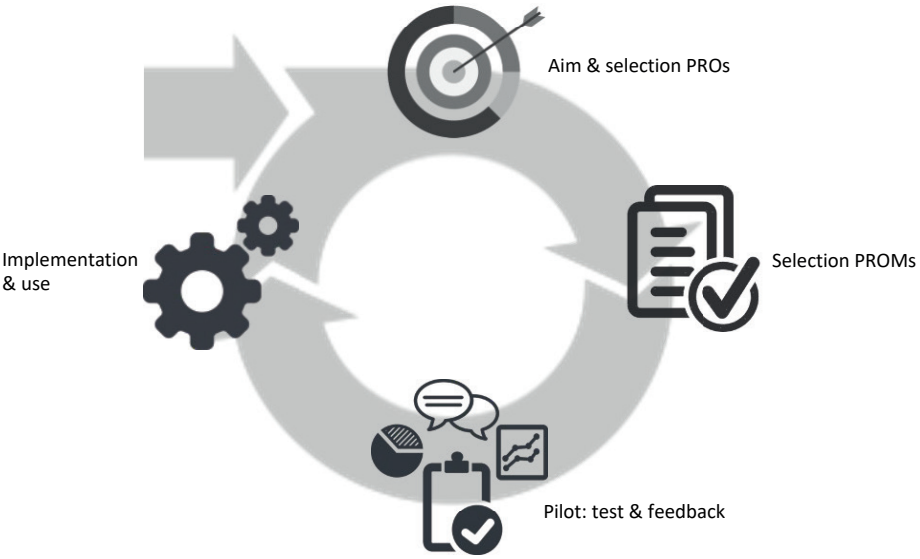


Figure 1. Steps for implementation of PROMs into routine care (PROMs-NNL study).

This dissertation comprises the scientific research performed in each step and aims to provide insight into and practical knowledge of the implementation and use of PROMs in routine nephrology care.

Outline of this thesis

The existing literature shows that HRQOL and symptom burden are highly prioritized by patients and healthcare professionals.[24, 27-29] Information about these PROs can contribute to a personalized treatment both at individual patient-level during consultations and at aggregated level to better inform patients and to evaluate healthcare quality.[6, 50, 54, 55] Therefore, these predetermined aims and PROs are used in the second step.

Chapter 2 describes the selection of the best suitable existing PROM to assess disease-specific symptom burden for routine assessment in nephrology care. We use a four-phase mixed methods approach, including a systematic literature search to identify existing PROMs and symptom clusters, assessment of PROMs based on predefined criteria regarding content validity, and selection based on feedback of two panels with patients and experts. In **Chapter 3**, we examine and compare psychometric properties of two recommended and commonly used generic PROMs to assess HRQOL. This study investigates the content, construct validity and test-retest reliability of seven PROMIS CATs in comparison with the SF-12 in patients with advanced CKD.

Chapter 4 describes the experiences and results of the first introduction of PROMs into Dutch routine nephrology care; the third step. We conduct a pilot study in 16 dialysis centres across the Netherlands, covering a quarter of all Dutch patients receiving dialysis treatment. We use quantitative and qualitative research methods to explore the use and collection of PROMs (e.g. PROM-scores and response rates), and the provision of feedback on PROM-results (e.g. patients' views on individual feedback) as part of routine dialysis care. Building on the findings, the PROMs infrastructure can be optimized for implementation and use of PROMs in routine dialysis care throughout the Netherlands (the fourth step).

At population level, PROM-results can be used to evaluate healthcare quality and to inform patients and professionals about the effects and course of disease or treatment. In **Chapter 5**, we explain how funnel plots can be used to compare healthcare providers on PROs to evaluate healthcare quality. This review provides insight into the use and interpretation of funnel plots by explaining the basic principles, pitfalls and considerations when applied to PROs, using examples of the first year routinely collected PROMs-data from Dutch dialysis care (i.e. RENINE/PROMs registry data). **Chapter 6** shows an example of aggregated PROM-results that can be used to inform patients and healthcare professionals. In this chapter, we use the RENINE/PROMs registry data of 2978 patients to investigate the impact of itching on HRQOL in patients receiving dialysis treatment. The effects of itching on HRQOL and interactions with sleep problems and psychological symp-

toms are examined both cross-sectionally and longitudinally over a 2-year period. For optimal use of PROMs in individual patients, knowledge on how to interpret and discuss PROM-results is needed. In **Chapter 7**, we explain the different types and characteristics of PROMs and provide guidance on how to interpret individual PROM-scores and changes in PROM-scores over time. Concepts such as minimal detectable change, minimal important change and response shift are explained and illustrated with examples from nephrology care. In **Chapter 8**, we investigate how to optimally discuss PROM-results as part of routine care. Individual semi-structured interviews are performed to gain in-depth understanding of patients' and healthcare professionals' experiences with and perspectives on discussing PROM-results in routine dialysis care.

Finally, in **Chapter 9** we summarize and discuss our results, and provide suggestions for future research and clinical implications regarding the implementation and use of PROMs in routine care.

References

1. World Health Organization. Regional Office for the Western P. People-centred health care: technical papers: International Symposium on People-Centred Health Care: Reorienting Health Systems in the 21st Century, The Tokyo International Forum, 25 November 2007. WHO Regional Office for the Western Pacific, Manila: 2008; 2008

2. Porter ME. What is value in health care? *N Engl J Med* 2010;363(26):2477-81

3. Feldthusen C, Forsgren E, Wallström S, et al. Centredness in health care: A systematic overview of reviews. *Health Expect* 2022;25(3):885-901

4. Lewis S. Realizing the PROMise of PROMs. *Healthc Pap* 2011;11(4):20-3; discussion 55-8

5. McLaren S, Jhamb M, Unruh M. Using Patient-Reported Measures to Improve Outcomes in Kidney Disease. *Blood Purif* 2021;50(4-5):649-54

6. Schick-Makaroff K, Levay A, Thompson S, et al. An Evidence-Based Theory About PRO Use in Kidney Care: A Realist Synthesis. *The Patient - Patient-Centered Outcomes Research* 2022;15(1):21-38

7. Noonan VK, Lyddiatt A, Ware P, et al. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 3: patient-reported outcomes can facilitate shared decision-making and guide self-management. *J Clin Epidemiol* 2017;89:125-35

8. KDIGO. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl* 2013;3 (Supl. 1):1-150

9. Webster AC, Nagler EV, Morton RL, Masson P. Chronic Kidney Disease. *The Lancet* 2017;389(10075):1238-52

10. Bikbov B, Purcell CA, Levey AS, et al. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2020;395(10225):709-33

11. 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

12. Jager KJ, Fraser SDS. The ascending rank of chronic kidney disease in the global burden of disease study. *Nephrol Dial Transplant* 2017;32(suppl_2):ii121-ii8

13. Davison SN, Levin A, Moss AH, et al. Executive summary of the KDIGO Controversies Conference on Supportive Care in Chronic Kidney Disease: developing a roadmap to improving quality care. *Kidney Int* 2015;88(3):447-59

14. K/DOQI. Clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002;39(2 Suppl 1):S1-266

15. Abrahams AC, van Jaarsveld BC. [Dialysis in end-stage kidney disease]. *Ned Tijdschr Geneeskd* 2020;164

16. Lindsay RM, Heidenheim PA, Nesrallah G, Garg AX, Suri R. Minutes to Recovery after a Hemodialysis Session: A Simple Health-Related Quality of Life Question That Is Reliable, Valid, and Sensitive to Change. *Clinical Journal of the American Society of Nephrology* 2006;1(5):952-9

17. Elsayed MM, Zeid MM, Hamza OMR, Elkholy NM. Dialysis recovery time: associated factors and its association with quality of life of hemodialysis patients. *BMC Nephrolo-*

gy 2022;23(1):298

18. Bello AK, Okpechi IG, Osman MA, et al. Epidemiology of haemodialysis outcomes. *Nat Rev Nephrol* 2022;18(6):378-95

19. Verberne WR, Dijkers J, Kelder JC, et al. Value-based evaluation of dialysis versus conservative care in older patients with advanced chronic kidney disease: a cohort study. *BMC Nephrol* 2018;19(1):205

20. 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

21. van Oosten MJM, Logtenberg SJJ, Leegte MJH, et al. Age-related difference in health care use and costs of patients with chronic kidney disease and matched controls: analysis of Dutch health care claims data. *Nephrol Dial Transplant* 2020;35(12):2138-46

22. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med* 2004;351(13):1296-305

23. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

24. Fletcher BR, Damery S, Aiyegbusi OL, et al. Symptom burden and health-related quality of life in chronic kidney disease: A global systematic review and meta-analysis. *PLoS Med* 2022;19(4):e1003954

25. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

26. Murtagh FE, Addington-Hall J, Higginson IJ. The prevalence of symptoms in end-stage renal disease: a systematic review. *Adv Chronic Kidney Dis* 2007;14(1):82-99

27. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. *Clin J Am Soc Nephrol* 2014;9(10):1813-21

28. Urquhart-Secord R, Craig JC, Hemmelgarn B, et al. Patient and Caregiver Priorities for Outcomes in Hemodialysis: An International Nominal Group Technique Study. *Am J Kidney Dis* 2016;68(3):444-54

29. de Jong Y, van der Willik EM, Milders J, et al. Person centred care provision and care planning in chronic kidney disease: which outcomes matter? A systematic review and thematic synthesis of qualitative studies : Care planning in CKD: which outcomes matter? *BMC Nephrol* 2021;22(1):309

30. Tommel J, Evers AWM, van Hamersvelt HW, et al. "What matters to you?": The relevance of patient priorities in dialysis care for assessment and clinical practice. *Semin Dial* 2022

31. González AM, Gutman T, Lopez-Vargas P, et al. Patient and Caregiver Priorities for Outcomes in CKD: A Multinational Nominal Group Technique Study. *Am J Kidney Dis* 2020;76(5):679-89

32. Weisbord SD, Fried LF, Mor MK, et al. Renal provider recognition of symptoms in patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2007;2(5):960-7

33. Pugh-Clarke K, Read SC, Sim J. Symptom experience in non-dialysis-dependent chronic kidney disease: A qualitative descriptive study. *J Ren Care* 2017;43(4):197-208

34. Aresi G, Rayner HC, Hassan L, et al. Reasons for Underreporting of Uremic Pruritus in People With Chronic Kidney Disease: A Qualitative Study. *J Pain Symptom Manage* 2019;58(4):578-86.e2

35. Flythe JE, Dorough A, Narendra JH, Forfang D, Hartwell L, Abdel-Rahman E. Perspectives on symptom experiences and symptom reporting among individuals on hemodialysis. *Nephrol Dial Transplant* 2018;33(10):1842-52

36. Feldman R, Berman N, Reid MC, et al. Improving symptom management in hemodialysis patients: identifying barriers and future directions. *J Palliat Med* 2013;16(12):1528-33

37. Kalantar-Zadeh K, Li PK, Tantisattamo E, et al. Living well with kidney disease by patient and care partner empowerment: kidney health for everyone everywhere. *Transpl Int* 2021;34(3):391-7

38. Al Sayah F, Lahtinen M, Bonsel GJ, Ohinmaa A, Johnson JA. A multi-level approach for the use of routinely collected patient-reported outcome measures (PROMs) data in healthcare systems. *Journal of Patient-Reported Outcomes* 2021;5(2):98

39. Kalantar-Zadeh K, Lockwood MB, Rhee CM, et al. Patient-centred approaches for the management of unpleasant symptoms in kidney disease. *Nat Rev Nephrol* 2022;18(3):185-98

40. Weldring T, Smith SM. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health Serv Insights* 2013;6:61-8

41. Tang E, Bansal A, Novak M, Mucsi I. Patient-Reported Outcomes in Patients with Chronic Kidney Disease and Kidney Transplant-Part 1. *Front Med (Lausanne)* 2017;4:254

42. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge university press: 2011;

43. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010;63(11):1179-94

44. Flythe JE, Karlsson N, Sundgren A, et al. Development of a preliminary conceptual model of the patient experience of chronic kidney disease: a targeted literature review and analysis. *BMC Nephrol* 2021;22(1):233

45. 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

46. Flythe JE, Powell JD, Poulton CJ, et al. Patient-Reported Outcome Instruments for Physical Symptoms Among Patients Receiving Maintenance Dialysis: A Systematic Review. *Am J Kidney Dis* 2015;66(6):1033-46

47. Nair D, Wilson FP. Patient-Reported Outcome Measures for Adults With Kidney Disease: Current Measures, Ongoing Initiatives, and Future Opportunities for Incorporation Into Patient-Centered Kidney Care. *Am J Kidney Dis* 2019;74(6):791-802

48. Breckenridge K, Bekker HL, Gibbons E, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. *Nephrol Dial Transplant* 2015;30(10):1605-14

49. Carfora L, Foley CM, Hagi-Diakou P, et al. Patients' experiences and perspectives of patient-reported outcome measures in clinical care: A systematic review and qualitative

meta-synthesis. *PLoS One* 2022;17(4):e0267030

50. 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. *J Patient Rep Outcomes* 2018;2:42

51. Etkind SN, Daveson BA, Kwok W, et al. Capture, transfer, and feedback of patient-centered outcomes data in palliative care populations: does it make a difference? A systematic review. *J Pain Symptom Manage* 2015;49(3):611-24

52. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* 2016;34(6):557-65

53. Johnson JA, Al Sayah F, Buzinski R, et al. A cluster randomized controlled trial for the Evaluation of routinely Measured PATient reported outcomes in HemodialYsis care (EMPATHY): a study protocol. *BMC Health Serv Res* 2020;20(1):731

54. Greenhalgh J, Dalkin S, Gibbons E, et al. How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. *J Health Serv Res Policy* 2018;23(1):57-65

55. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q* 2014;92(4):754-75

56. Bingham CO, 3rd, Noonan VK, Auger C, Feldman DE, Ahmed S, Bartlett SJ. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 4: patient-reported outcomes can inform clinical decision making in chronic care. *J Clin Epidemiol* 2017;89:136-41

57. Standardised Outcomes in Nephrology (SONG). https://songinitiative.org/

58. International Consortium for Health Outcomes Measurement (ICHOM). https://www.ichom.org/

59. Anderson N, Kyte D, McMullan C, et al. Electronic patient-reported outcomes in chronic kidney disease. *Nat Rev Nephrol* 2022;18(12):739-40

60. Nimmo A, Bell S, Brunton C, et al. Collection and determinants of patient reported outcome measures in haemodialysis patients in Scotland. *Qjm* 2018;111(1):15-21

61. Foster A, Croot L, Brazier J, Harris J, O'Cathain A. 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

62. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23(6):508-18

63. Anderson NE, McMullan C, Calvert M, et al. Using patient-reported outcome measures during the management of patients with end-stage kidney disease requiring treatment with haemodialysis (PROM-HD): a qualitative study. *BMJ Open* 2021;11(8):e052629

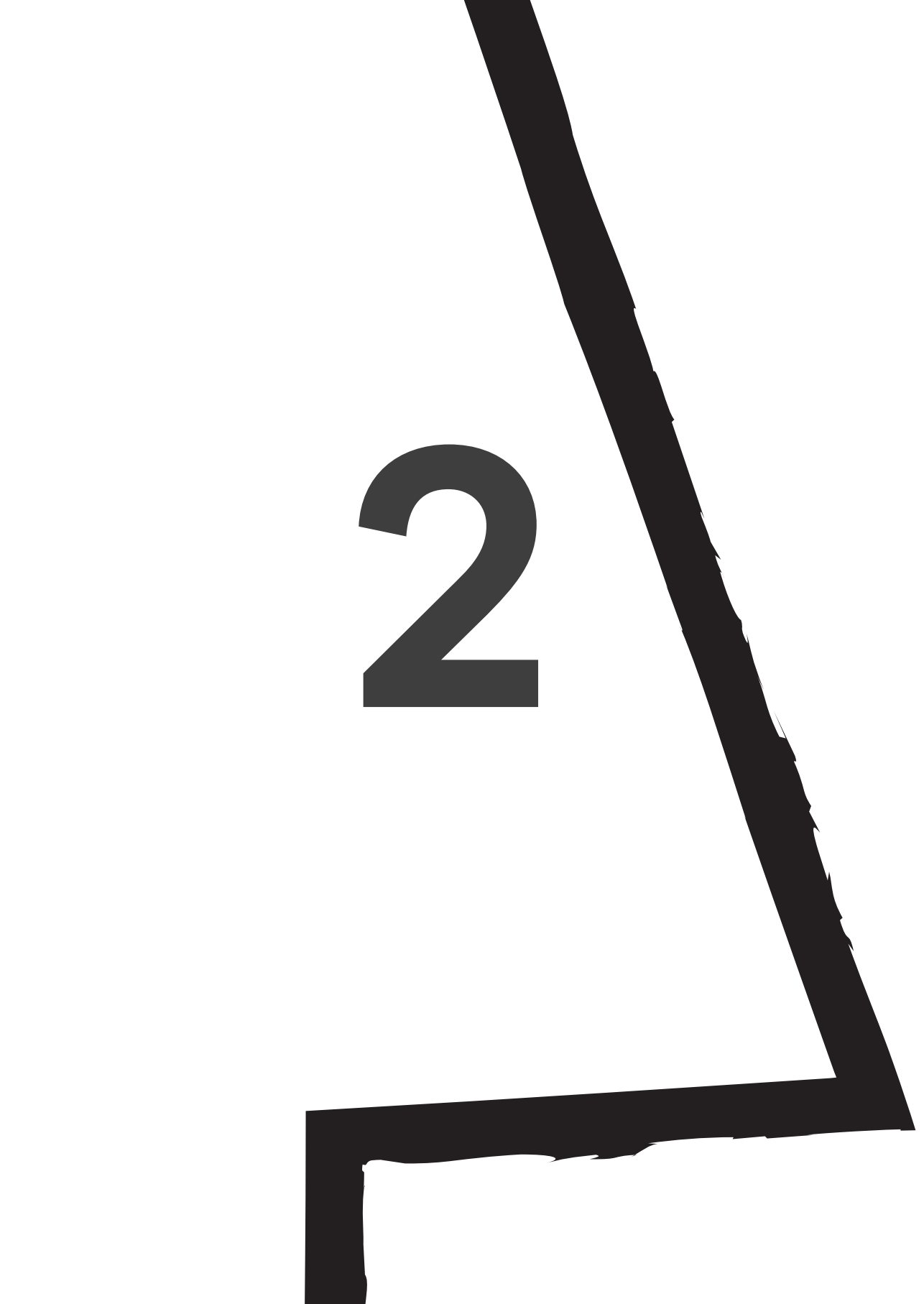
64. Apel C, Hornig C, Maddux FW, Ketchersid T, Yeung J, Guinsburg A. Informed decision-making in delivery of dialysis: combining clinical outcomes with sustainability. *Clin Kidney J* 2021;14(Suppl 4):i98-i113

65. Flythe JE, Narendra JH, Hilliard T, et al. Cultivating a Research-Ready Dialysis Community. *J Am Soc Nephrol* 2019;30(3):375-80

66. van der Veer SN, Couchoud C, Morton RL. The role of kidney registries in expediting large-scale collection of patient-reported outcome measures for people with chronic kidney disease. *Clinical Kidney Journal* 2021;14(6):1495-503

Part 1

Implementation of PROMs into routine nephrology care


2

Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease – a four-phase mixed methods study

Esmee M. van der Willik, Yvette Meuleman, Karen Prantl, Giel van Rijn, Willem Jan W. Bos, Frans J. van Ittersum, Hans A.J. Bart, Marc H. Hemmelder, Friedo W. Dekker

BMC Nephrology. 2019;20(1):344.

Abstract

Background: Patient-reported outcome measures (PROMs) are becoming increasingly important in healthcare. In nephrology, there is no agreement on which chronic kidney disease (CKD) symptom questionnaire to use. Therefore, the aim of this study is to select a valid symptom questionnaire for routine assessment in patients with advanced CKD.

Methods: A four-phase mixed methods approach, using qualitative and quantitative research methods, was applied. First, a systematic literature search was conducted to retrieve existing symptom questionnaires. Second, a symptom list was created including all symptoms in existing questionnaires and symptoms mentioned in interviews with patients with CKD, from which symptom clusters were identified. Next, questionnaires were selected based on predefined criteria regarding content validity. Last, two online feedback panels of patients with CKD (N=151) and experts (N=6) reviewed the most promising questionnaires.

Results: The literature search identified 121 questionnaires, of which 28 were potentially suitable for symptom assessment in patients with advanced CKD. 101 unique symptoms and 10 symptom clusters were distinguished. Based on predefined criteria, the Dialysis Symptom Index (DSI) and Palliative Care Outcome Scale-Renal Version (IPOS-Renal) were selected and reviewed by feedback panels. Patients needed 5.4 and 7.5 minutes to complete the DSI and IPOS-Renal, respectively ($p<0.001$). Patients experienced the DSI as more specific, complete and straightforward compared to the IPOS-Renal.

Conclusions: The DSI was found to be valid and reliable, the most relevant, complete, and comprehensible symptom questionnaire available for routine assessment in patients with advanced CKD. Routine PROMs collection could be of great value to healthcare, both at individual patient and national level. Feedback on scores and involvement of healthcare providers may promote adaptation and implementation in healthcare.

Background

The last decade there has been a shift in focus towards a more patient-centred and value-based healthcare. As described by Michael E. Porter, value in healthcare depends on the outcomes achieved and should be defined around the patient.[1] With this change, patient-reported outcomes (PROs) are becoming increasingly important in healthcare.[1-4] PRO measures (PROMs) can be used to quantify a wide variety of concepts of health that are relevant to the patient, such as quality of life, functional status and symptom burden.[2, 5]

Until recently, PROMs were mainly used in research settings. However, PROMs are increasingly being applied for clinical management in individual patients and evaluation of quality of care.[4, 6] PROMs may enhance understanding of patients' symptoms and needs, and have the potential to improve patient outcomes and engagement in decision making.[4, 7, 8] The use of PROMs is nowadays recommended to be implemented and routinely used in clinical practice.[4, 9, 10]

Broadly a PROM can be classified as a generic or disease specific instrument. Generic PROMs measure general aspects of patients' health status, such as functional status or quality of life. Disease specific PROMs are tailored to a specific condition and address aspects of disease experience and symptoms, making these PROMs in general more sensitive and responsive to change in disease burden.[2, 4, 5, 10] Often, both generic and disease specific PROMs are used to enable comparisons across and within populations.[4, 5, 10]

Also in nephrology, routine collection of PROMs can be of added value.[11] Patients with advanced chronic kidney disease (CKD) experience a poor health-related quality of life (HRQOL) and numerous physical and emotional disease related symptoms.[12-14] Moreover, in patients with advanced CKD, HRQOL levels generally decrease and symptom burden generally increases as the disease progresses.[15] Despite their relevance, many symptoms in patients with advanced CKD remain unnoticed. This may be partly explained by patients being reluctant to share their experienced symptoms, particularly due to feelings of guilt about wasting clinicians' or other patients' time.[16] Additionally, clinicians frequently are not able to identify the full spectrum of experienced symptoms and their severity, resulting in under-recognition and under-treatment of symptoms.[14, 17-19] Routine symptom assessment, using a questionnaire that fits patients' needs, could provide insight and guidance for symptom management.[16, 20] Symptom management has been identified as top priority by patients with advanced CKD.[21]

Although the relevance of patients' perspective is recognized, PROMs have not yet been widely implemented in nephrology.[2, 9, 11] Currently, methods and instruments needed for implementation of PROMs in patients with advanced CKD, including

patients with end-stage kidney disease (ESKD) with and without dialysis, are being explored in the Netherlands. Some generic health questionnaires are considered to be appropriate instruments for this purpose.[9, 22] However, there is no agreement on which questionnaire is most suitable to measure the broad spectrum of symptoms that patients with advanced CKD experience.[9, 23] Therefore, the aim of this study is to systematically select the most suitable CKD-specific symptom questionnaire for routine assessment in patients with advanced CKD and ESKD using a four-phase mixed methods approach.

Methods

Overview

This study is part of the development of a national registry of PROMs, which will be included in the Dutch Renal Registry (Renine) [www.renine.nl]. For now, the PROMs registry is primarily aimed at patients with advanced CKD, including patients with ESKD receiving dialysis or without renal replacement therapy (RRT). Patients will be followed over time across different stages and treatments (e.g. advanced CKD, ESKD, with and without RRT), and therefore, we have chosen not to restrict this study to subpopulations, but to focus on CKD in general, taking all existing CKD-specific symptom questionnaires into consideration.

In this study, the focus was on the content validity of the symptom questionnaire, defined as "the relevance, comprehensiveness, and comprehensibility of the PROM for the construct, target population, and context of use of interest".[24] According to the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) standards, content validity is the most important and first to be considered measurement property in selecting a PROM.[24, 25] Furthermore, since numerous symptom questionnaires are already available [2, 26], it would be preferable to select an existing questionnaire instead of developing a new one. As an alternative for organizing focus groups and interviews with patients to identify domains of symptoms relevant to patients, we searched and used existing CKD symptom questionnaires, assuming that they all have attempted to include the most important domains and items. By combining all these questionnaires, we make use of a much wider variation in patients, methods, clinical settings and countries to gather content-wise relevant domains for CKD.

A four-phase approach, combining qualitative and quantitative research methods, was applied: **1)** conduct a systematic literature search to retrieve all existing symptom questionnaires used in patients with CKD. **2)** Create a complete list of unique symptoms from all symptom questionnaires and interviews with patients with advanced CKD. Cluster these symptoms into relevant symptom groups. **3)** Select

symptom questionnaires based on criteria to ensure content validity, including the completeness, relevance and comprehensibility for the advanced CKD population and context of routine care.[24] **4)** Evaluate the most promising symptom questionnaires using a panel of patients with advanced CKD and experts (i.e. experienced questionnaire assessors). Below the four phases are described in detail.

Systematic literature search – phase 1

A systematic literature search was performed to identify all existing symptom questionnaires developed and/or used in patients with CKD. A query was constructed using numerous synonyms or identifiers for the keywords 'chronic kidney disease', 'symptoms' and 'questionnaires' (Supplementary item S1). The search was restricted to studies published in the English or Dutch language. Studies conducted in individuals <18 years of age were excluded.

The search was executed in PubMed by two independent reviewers (EvdW and GvR). Titles were screened and found to be eligible when describing one or more symptoms or the use of a symptom questionnaire in patients with CKD. Next, the abstracts of articles included by at least one of the reviewers were screened to identify existing symptom questionnaires. Systematic reviews describing the use of questionnaires in patients with CKD were screened full text to make sure that all existing symptom questionnaires were included.

We aim to select a symptom questionnaire addressing the full range of symptoms experienced by the total CKD population. To distinguish such broad symptom questionnaires from in-depth questionnaires addressing only one or two specific symptoms (e.g. depression or fatigue questionnaires), we excluded symptom questionnaires addressing less than four physical or emotional symptoms.[26] Additionally, questionnaires focusing only on transplant-specific symptoms and generic health questionnaires (e.g. HRQOL or activities of daily living questionnaires) were excluded.

Symptom list and clustering – phase 2

Symptoms from questionnaires. A list of symptoms was created from all symptoms included in the questionnaires. To collect only unique symptoms, overlapping symptoms were combined (e.g. 'Tingling in feet or hands' as a combination of 'Tingling in feet' and 'Tingling in hands').

Analysis of videotaped interviews. To assure completeness of the symptom list, 18 videotaped interviews with patients with advanced CKD were analysed to check for missing symptoms. Patients received haemodialysis (n=13), peritoneal dialysis (n=3) or no RRT (n=2), were 20-83 years old, and half of them was male. The

interviews were conducted by two experienced male interviewers (HB and FvdZ), who were not involved in the patients' treatment. The videos were obtained from the Dutch Kidney Patients Association (NVN) [27] and were developed to inform and support patients with CKD in making future choices regarding therapy. During the semi-structured interviews, different aspects of living with CKD were discussed, including aspects about disease, treatment, physical functioning, psychosocial aspects, relationships and quality of life. As a result of patient's answers, additional themes were sometimes introduced including symptoms that patients experience and considered relevant. The NVN and the interviewed patients gave permission to use this material for this research purposes. The videotaped interviews were watched and analysed by two independent researchers (GvR and EvdW). All symptoms mentioned by patients were written down verbatim and subsequently compared to the symptom list derived from the questionnaires (phase 1). Symptoms that were not yet on the list were added.

Clustering of symptoms. The total list of unique symptoms was divided into clusters to identify themes that describe the broad spectrum of symptoms experienced by patients with CKD. Clustering was done by two independent healthcare professionals: a nephrologist (JR) and a nurse practitioner (NBB) specialized in pre-dialysis and dialysis care, both experienced in clinical practice and research. JR and NNB discussed the symptoms and identified clusters inductively by constant comparison and grouping of similar type of symptoms. Clusters and corresponding symptoms were discussed until consensus was reached.

Preliminary selection of symptom questionnaire – phase 3

A set of criteria (Table 1) was applied to make a preliminary selection of symptom questionnaires that are relevant, complete and comprehensible for patients with advanced CKD or ESKD in routine care setting.

Feedback panels – phase 4

Dutch versions of the most promising questionnaires were evaluated by two online feedback panels facilitated by the NVN. One panel consisted of 151 patients receiving different treatments: pre-dialysis (CKD stage 4/5), haemodialysis, peritoneal dialysis and transplantation. The patients in this panel were randomly assigned to one of the selected questionnaires. Patients assessed only one questionnaire so that their judgement on the assigned questionnaire was based on their personal opinion, experiences and needs, and not influenced by the content or structure of another questionnaire. The second panel consisted of six experienced questionnaire assessors, namely NVN patient representatives who

Table 1. Criteria for symptom questionnaires suitable for routine assessment in patients with advanced chronic kidney disease (CKD).

Criterion	Description
A. Symptom clusters	≥ 90% cluster coverage The variety of symptoms experienced in CKD requires a questionnaire addressing a wide range of symptoms. Preferably all, but at least 90% of the clusters should be covered by the questionnaire.
B. Questionnaire length	≤ 90 items The questionnaire needs to have an appropriate length to be suitable for routine assessment. The questionnaires should have a maximum length of 15 minutes to complete [28], which we expect to be exceeded by a questionnaire addressing ≥90 items.[29]
C. Applicable to advanced CKD population	Developed and validated in advanced CKD The questionnaires should be applicable to the advanced CKD population. Preference is given to a questionnaire both developed and validated in patients with advanced CKD.
D. Suitable for use in routine care	Straightforward and clear For a questionnaire addressing more than symptoms only, the symptoms need to be concentrated together (i.e. symptom questions are not mixed with other questions), so that a separate and valid symptom questionnaire can be extracted. Since patient's ability to concentrate and understand difficult items may be impaired, the questionnaire needs to be straightforward with appropriate and easy to interpret items and scales.[29]

advise on research (e.g. questionnaire development). Five of these experts were CKD patient and one person was a relative of a CKD patient. To enable a direct comparison of the questionnaires, this panel of experts compared all questionnaires from the previous phase.

To review the questionnaire, patients were asked to complete the questionnaire and to answer additional questions. Questions concerned the content and structure of the questionnaire, including: time needed for completion, burden of completing the questionnaire, desired frequency of questionnaire assessment, unclear questions, unnecessary questions, missing questions with room to report three additional symptoms, and other suggestions or comments. The time to complete the questionnaire was measured electronically (i.e. objective time). Patients also estimated the time to complete the questionnaire, hereafter referred to as subjective time. Differences between the questionnaires in objective and subjective time to complete were presented as geometric mean.

Statistical analyses were performed using SPSS V.23.0. P<0.05 was considered statistically significant. To evaluate differences in patient, treatment and questionnaire characteristics, Student's t-test and Chi-square tests of association were performed. To test the reliability of the symptom burden score, Cronbach's alpha coefficients were calculated. A sensitivity analysis using one-way ANOVA and Chi-square tests was conducted to determine if the results from the patient panel are the same for transplant patients compared to patients on dialysis or without RRT.

Results

Systematic literature search – phase 1

Figure 1 shows a flow diagram of the literature search and questionnaire selection. The search strategy identified 571 articles, of which 223 articles were included based on title and abstract. From these articles, including two full text reviews, and through snowballing, 121 unique symptom questionnaires were identified. Of these questionnaires, 93 were excluded (mainly because less than four symptoms were addressed, see Figure 1), resulting in 28 symptom questionnaires for further investigation.

Symptom list and clustering – phase 2

A complete symptom list was created from the 28 symptom questionnaires. One hundred unique symptoms were identified from these questionnaires. Analysis of the videotaped interviews with patients with advanced CKD resulted in one additional symptom (Supplementary table S2). From this symptom list, two healthcare professionals distinguished the following ten clusters: general symptoms, night's rest, gastroenterology, cardiopulmonary, central nervous system, musculoskeletal, skin, head/throat, psychosocial and sex. The total symptom list categorized into ten clusters is available in Supplementary table S2.

Preliminary selection of symptom questionnaire – phase 3

In the third phase, the previous two steps were combined: the 28 symptom questionnaires were judged on their coverage of at least nine out of ten symptom clusters. Fifteen questionnaires were excluded based on this criterion (criterion A). An additional three questionnaires were excluded due to their extensive length (criterion B), leaving ten questionnaires for further examination.

Table 2 shows the characteristics of the ten questionnaires on which the questionnaires were compared and evaluated. Six out of ten questionnaires were both developed and validated in an advanced CKD population, meeting criterion C. Two of these six are derivatives of a third questionnaire; the two questionnaires include

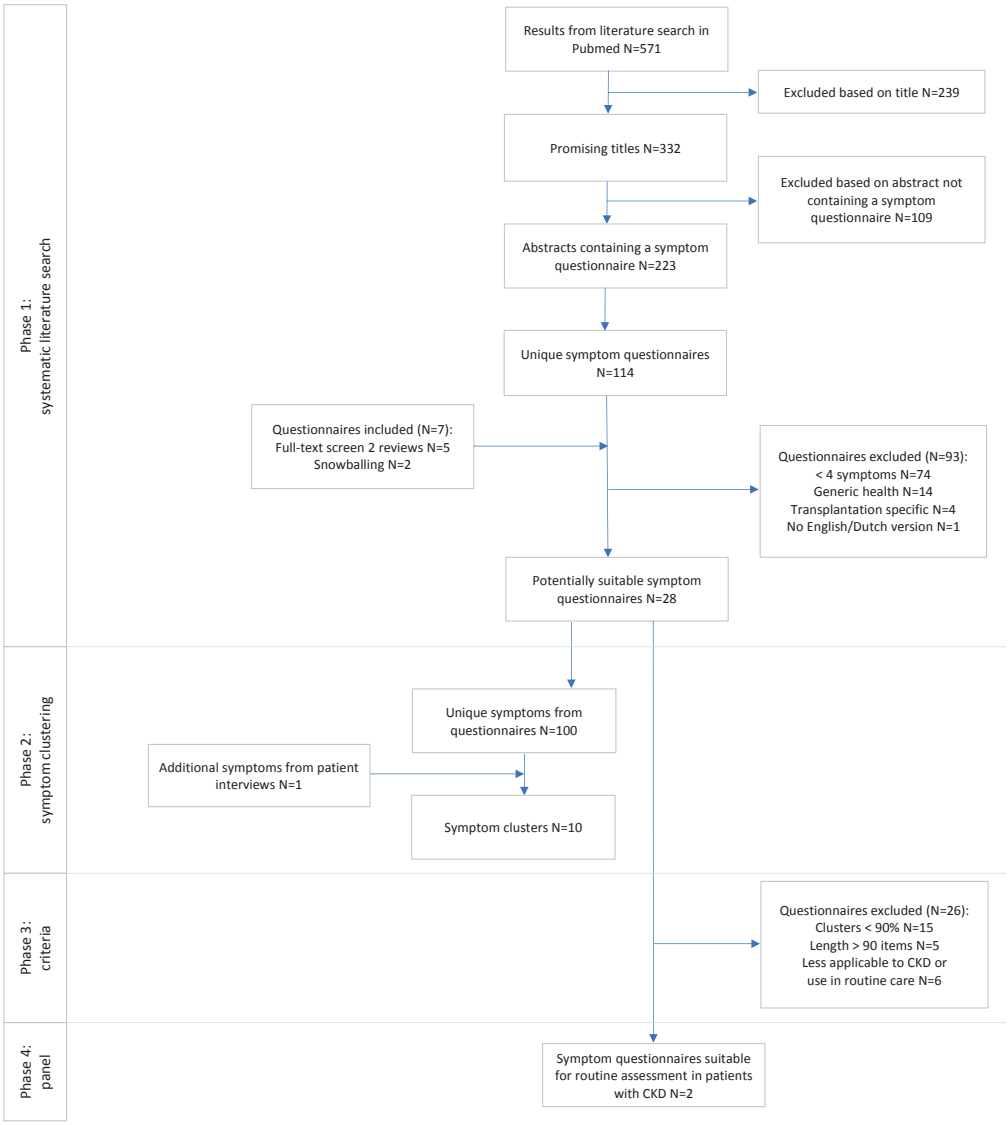


Figure 1. Flow chart of the selection of a valid CKD-specific symptom questionnaire.

exactly the same symptoms but also distinguish how much a symptom bothers, the severity and the frequency of symptoms and hereby exceed the determined maximum length (criterion B). Another two questionnaires address a broader perspective than symptoms only. The questions regarding symptoms are spread across the questionnaire, which does not satisfy criterion D. Based on the criteria two questionnaires were selected for further consideration in the next phase.

Table 2. Characteristics of 10 suitable symptom questionnaires for routine assessment in patients with advanced CKD based on cluster coverage and suitable length.

	CHEQ	CKD-SBI$^{\Delta}$	Curtin	DFSSBI$^{\Delta}$	DSI$^{\Delta}$	IPOS-Renal	KDQOL-SF	MSAS	MSAS-SF/ + renal symptoms	RSCL
Development population	Dialysis	CKD 4-5	Dialysis	Dialysis	Dialysis	CKD 4-5	Dialysis	*Cancer*	*Cancer* / CKD 5	*Cancer*
Validated for advanced CKD	Yes	Yes	*No*	Yes	Yes	Yes	Yes	*No*	Yes / *No*	*No*
Total number of items	80	*33**	47	*31**	30	21	82	*33**	33 / 39	39
Separate symptom component	*No*	Yes	Yes	Yes	Yes	Yes	*No*	Yes	Yes	Yes
Number of items per cluster										
General symptoms	3	2	4	2	2	2	6	5	5	3
Night's rest	3	2	4	2	2	2	4	2	2	1
Gastroenterology	4	5	8	5	5	5	2	7	7	7
Cardiopulmonary	2	5	3	4	4	1	3	3	3 / 4	1
Central nervous system	3	4	7	4	4	1	3	3	3 / 5	4
Musculoskeletal	1	3	5	3	3	1	2	0	0 / 3	2
Skin	2	2	3	2	2	2	2	4	4 / 5	1
Head and throat	1	1	4	1	1	1	0	3	3	3
Psychosocial	12	6	6	6	5	4	12	4	4	7
Sex	3	2	3	2	2	0	2	1	1	1
Open questions	1	1	0	0	0	2	0	1	1	0
Burden rating scale	2 till 7-point scale	11-point Likert scale	5-point Likert scale	5- and 10-point scale	5-point Likert scale	5-point Likert scale	2 till 10-point scale	4- and 5-point scale	4- and 5-point scale	4-point scale
Recall$^{\#}$	4 weeks / 3 months / in general	4 weeks	4 weeks	1-3 days^	1 week	1 week	4 weeks / in general	1 week	1 week	1 week

Italic marks reasons for exclusion.
[Δ] CKD-SBI and DFSSBI are derivatives of the DSI and include the same symptoms (with 1 or 2 additional symptom(s)).
[*] For each item, the patient was expected to report the frequency, severity and bothersome.
[#] The time period addressed, e.g. the recall period is 1 week for the question "did you experience this symptom during the past week?"
[^] The time period was defined as "since your last dialysis treatment" and thus depended on the frequency of dialysis treatment, which was estimated at 3-7 times per week.
Abbreviations and questionnaires: CKD, Chronic Kidney Disease, CHEQ, CHOICE Health Experience Questionnaire [30]; CKD-SBI, CKD Symptom Burden Index [31]; Curtin [32]; DFSSBI, Dialysis Frequency, Severity, and Symptom Burden Index [33]; DSI, Dialysis Symptom Index [34]; IPOS-Renal, Palliative Care Outcome Scale - Renal Version [35]; KDQOL-SF, Kidney Disease Quality of Life instrument - Short Form [36]; MSAS, Memorial Symptom Assessment Scale [37]; MSAS-SF, Memorial Symptom Assessment Scale-Short Form [38]; MSAS-SF with additional renal symptoms [39]; RSCL, Rotterdam Symptom Checklist [40].

Feedback panels – phase 4

Feedback panels. The Dialysis Symptom Index (DSI) [34] and Palliative Care Outcome Scale – Renal Version (IPOS-Renal) [35] were judged by two online panels of patients and experts. Patients were randomly assigned to a questionnaire (Table 3). In total 127 patients (84.1%) received RRT, of which 27 dialysis (17.9%) and 100 transplantation (66.2%). The second panel of six experts evaluated and compared both questionnaires.

Table 3. Comparison of two CKD-specific symptom questionnaires based on feedback of the patient panel (N=151).

	DSI (N=76)	IPOS-Renal (N=75)	*p*-value
Age (years)	60.6 (12.5)	60.2 (10.4)	0.8
Treatment modality			0.5
Pre-dialysis	6 (7.9)	13 (17.3)	
Haemodialysis	8 (10.5)	9 (12.0)	
Peritoneal dialysis	6 (7.9)	4 (5.3)	
Transplant	53 (69.7)	47 (62.7)	
Other	3 (3.9)	2 (2.7)	
Objective time to complete* (minutes)	5.4 (1.6)	7.5 (1.8)	<0.001
Subjective time to complete* (minutes)	3.2 (1.8)	4.8 (1.6)	<0.001
Number of symptoms reported^	12.0 (6.5)	8.0 (4.1)	<0.001
Additional 1-3 symptoms reported#	21 (27.6)	25 (33.3)	0.5
Burdensome of questionnaire (yes)	4 (5.3)	2 (2.7)	0.4
Appropriate frequency of submission (times per year)	2.7 (1.8)	2.9 (2.2)	0.6

Values are shown in N (%) or mean (SD).
The DSI and IPOS-Renal questionnaires showed good reliability for symptom burden score with Cronbach's alpha values of 0.90 and 0.86, respectively.
**Objective time to complete was defined as the difference in minutes between the start and completion of the online questionnaire. Subjective time to complete is the time to complete estimated by the patient. Values shown as geometric mean (SD).*
^The number of symptoms reported is based on the symptoms defined in the questionnaire and rated by the patient as bothering a little bit to very much (or affecting slightly to overwhelmingly).
#The number of patients reporting an additional 1 to 3 symptoms not mentioned in de questionnaire.
Abbreviations: CKD, Chronic Kidney Disease; DSI, Dialysis Symptom Index; IPOS-Renal, Palliative Care Outcome Scale - Renal Version.

Time to complete. The patient panel needed on average (standard deviation; SD) 5.4 (1.6) minutes to complete the DSI and 7.5 (1.8) minutes to complete the IPOS-Renal ($p<0.001$). Also subjectively the DSI was less time consuming than the IPOS-Renal, with a difference in geometric mean of 1.6 minutes ($p<0.001$). The time to complete estimated by experts ranged from 2-15 and 3-20 minutes for the DSI and IPOS-Renal, respectively.

Burden and frequency. Four and two patients of the online patient panel experienced, respectively, the DSI and IPOS-renal as burdensome. All experts indicated that both questionnaires were not burdensome. For both questionnaires, most patients prefer to complete the questionnaire two or four times per year. Most experts (4 out of 6) desired four times per year. In both panels, participants noted that the questionnaire should be filled in prior to each consultation with the nephrologist.

Questions. Both panels indicated that, overall, the questions in both questionnaires were clear. All experts fully agreed that both questionnaires were easy to interpret and one expert added that the questionnaires were also comprehensible for patients with low literacy. For the IPOS-Renal some patients and one expert noted that the questions might be too generally formulated, which can cause confusion or difficulties to interpret a question. Also, some patients indicated that some questions might not be applicable to all patients or treatment modalities. For the DSI some patients mentioned that questions about sexual problems may be not applicable to all patients. Two experts indicated that, in comparison to the DSI, some symptoms may be missed when using the IPOS-Renal, but these experts did not mention which symptoms were lacking. For both questionnaires patients reported additional symptoms, which were all covered by the defined clusters. Patients reported more symptoms using the DSI than using the IPOS-Renal, with 12.0 and 8.0 experienced symptoms, respectively ($p<0.001$) (Table 3).

Comments. Patients made comments similar to the answers described above (see Questions). About the DSI several patients reported that they experienced the questionnaire as pleasant, clear and enlightening. For both questionnaires patients suggested to add questions on treatment and "how patients experience their lives". Additionally, the patients pointed out that feedback on their results and involvement of the nephrologist are highly important.

Preference. The experts compared both questionnaires. Five out of six experts preferred the DSI. They qualified the DSI as more specific and complete, and believed that the questions were more clear and easier to fill in than the IPOS-Renal. Two experts, however, also mentioned that the lay-out of the IPOS-Renal was visually more attractive than the DSI.

Sensitivity analysis. Results of the patient panel stratified for transplant and non-transplant patients are available in Supplementary table S3. Similar differences between the DSI and IPOS-Renal were found in transplant and non-transplant patients. Transplant patients completed both questionnaires faster (p = n.s.) and reported less symptoms (p = n.s.) compared to non-transplant patients. However, both transplant and non-transplant patients needed less time to complete the DSI ($p < 0.05$) and reported more symptoms using the DSI compared to the IPOS-Renal ($p < 0.05$). Also comments regarding the content and structure of the questionnaires were similar in transplant and non-transplant patients.

Discussion

The aim of this study was to select a valid CKD specific symptom questionnaire for routine assessment in patients with advanced CKD or ESKD. The first two phases, the literature search and symptom clustering, resulted in 28 potentially suitable symptom questionnaires and ten symptom clusters. During the third phase, two questionnaires were selected based on their relevance, completeness and comprehensibility to routine assessment in patients with CKD: the DSI and IPOS-Renal. These two questionnaires were reviewed by panels of patients and experts in the fourth phase. The results of the panel reviews showed that the DSI was the most complete, specific and comprehensible symptom questionnaire. Therefore, the DSI was considered to be the most suitable symptom questionnaire currently available for routine assessment in patients with advanced CKD or ESKD.

Previous literature and current findings support the completeness and straightforwardness of the DSI. First, the patient panel reported 12 symptoms using the DSI, which is 1.5 times the number of symptoms reported when using the IPOS-Renal. We believe that this increased score is due to differences in completeness of the questionnaires rather than differences in characteristics between patients. Similar numbers of symptoms are also presented in previous literature.[41] Furthermore, a recent study showed that symptoms of insomnia, fatigue, cramping, anxiety, depression and frustration were considered top-priority by dialysis patients. Such physical and emotional symptoms are also included in the DSI.[42] Still, additional symptoms were mentioned by the patients assessing the DSI. Therefore, we propose to retain the possibility to report additional symptoms as this may favour the completeness and patient satisfaction.[43] Besides this, the time to complete the questionnaire reflects the straightforwardness of the DSI. Although the DSI contains more items than the IPOS-Renal, patients needed less time to complete the questionnaire. This might suggest that the DSI is more clear and easier to complete for patients.

We believe that routine symptom assessment can contribute to a more patient-centred healthcare system and improvement in quality of care. Routine assessment enables patients and healthcare professionals to track changes in symptom burden over time, which may result in a more complete and better understanding of patients' symptoms and needs. Routine assessment may also yield valuable information for the evaluation of effectiveness of treatment and the progression of symptoms.

Herein, the provision of feedback on PROM score to patients and healthcare providers, both on individual and on aggregated level, may be of great importance.[44] At the individual patient level, feedback may enhance communication between patients and healthcare professionals, which is considered highly important by patients with advanced CKD.[21] Moreover, results of similar patients could provide insight in what to expect in the future and may promote patient engagement in decision making.[3, 45] Additional to the provision of feedback, the involvement of clinicians was considered very important by several patients and is expected to contribute to a successful implementation and a more patient-centred healthcare.[46]

At centre or national level, performance and variation in outcomes between centres can be mapped out and may promote initiatives to improve quality of care. Besides, patient outcomes are of great value to the already available clinical performance measures, which mainly consider structure and process of care.[47, 48] So far, PROMs have been mainly used in scientific research and less often for nationwide assessment in clinical practice.[4, 9] Consequently, little is known about how PROMs can be best deployed to achieve quality improvement.[49, 50]

Further research is needed to investigate how PROMs can be best used in clinical practice to improve symptom management, shared decision making and to address patients' needs. We propose to assess and discuss symptoms using the DSI twice per year, in order to gain insight into symptom development with a minimal burden to patients and to healthcare professionals. In addition to a suitable questionnaire, successful implementation of PROMs into routine care requires planning, facilities (e.g. electronic system to collect and report PROM scores) and involvement of all stakeholders (e.g. patients, healthcare professionals and researchers).[51] Furthermore, barriers may be encountered when implementing PROMs into routine care, including low response rates, organizational struggles or low commitment from patients or healthcare professionals.[11] To facilitate implementation and sustainability, it is vital to take these barriers into account, by, for example, providing information and communication systems to adequately collect data and discuss PROM-scores. We suggest to test the PROMs in collaboration

with all stakeholders so that it fits the workflow and priorities in routine care.

A unique feature of this study is the four-phase mixed methods approach with both qualitative and quantitative research methods. Especially with the combination of these methods we believe to have selected a valid and reliable symptom questionnaire that is relevant, complete and appropriate for the population and context of interest. First, this method addressed all criteria for evaluation of content validity as established in the COSMIN standard.[24] Second, with the use of all existing symptom questionnaires, we believe to have reached completeness and to have identified the domains that are most relevant to the patient, more so than would be possible when conducting a single study. This conclusion is also supported by patients' input: the analysis of the interviews with patients with CKD resulted in only one additional symptom and no new symptoms or domains were mentioned by the patient panel. Third, patients, healthcare professionals and experts were involved in this study. Particularly patient involvement was considered highly important, because patients' perspective helps to select the questionnaire that is most complete, comprehensible and relevant to them. This might increase the probability of completing the questionnaire when implemented in daily practice.[52] By using this mixed methods design, a symptom questionnaire that was preferred by experts and very positively assessed by patients was selected.

On the downside, the patient panel might be not representative of the entire advanced CKD population. First, patients participating in an online panel may be more health conscious, familiar with online questionnaires and involved in healthcare compared to those who do not participate (i.e. healthy responder bias). Second, most participants in the patient panel received a kidney transplant. However, the results and comments on the questionnaires of the transplant patients did not differ from those of the patients on dialysis or without RRT. Besides this, within the context of interest, patients will be followed over time, through different stages and treatments. Many patients receive (pre-)dialysis care prior to their transplantation, and thus, we do not expect that the inclusion of patients who received a kidney transplant affected the evaluation of the questionnaires.

With the method used, the focus was on the most important PROM property, namely the content validity of the symptom questionnaire. Additionally, the DSI showed good reliability: excellent internal consistency of the symptom burden score in this current study and good test-retest reliability in the development-study.[34] However, more research is needed to further explore the reliability and validity of this questionnaire.[34] Additionally, further research is needed to investigate if the DSI detects (clinically relevant) changes in symptom burden (i.e. responsiveness). Moreover, the smallest detectable change and the minimal important change need to be investigated for the interpretation of changes in symptom burden over time.[53]

Conclusion

In conclusion, the DSI was found to be valid and reliable, the most relevant, complete, and comprehensible symptom questionnaire currently available for routine assessment in patients with advanced CKD or ESKD. The use of PROMs could be of great added value to healthcare, both at the individual patient and national level. Feedback on results and involvement of healthcare providers may promote adaptation and implementation of PROMs into healthcare.

References

1. Porter ME. What is value in health care? *N Engl J Med* 2010;363(26):2477-81

2. Tang E, Bansal A, Novak M, Mucsi I. Patient-reported outcomes in patients with chronic kidney disease and kidney transplant-part 1. *Front Med (Lausanne)* 2017;4:254

3. Rotenstein LS, Huckman RS, Wagle NW. Making patients and doctors happier - the potential of patient-reported outcomes. *N Engl J Med* 2017;377(14):1309-12

4. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167

5. Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the Swedish National Quality Registers. *J Intern Med* 2016;279(2):141-53

6. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. *BMJ* 2010;340:c186

7. Etkind SN, Daveson BA, Kwok W, et al. Capture, transfer, and feedback of patient-centered outcomes data in palliative care populations: does it make a difference? A systematic review. *J Pain Symptom Manage* 2015;49(3):611-24

8. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? *Qual Life Res* 2009;18(1):115-23

9. Breckenridge K, Bekker HL, Gibbons E, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. *Nephrol Dial Transplant* 2015;30(10):1605-14

10. Devlin NJ, Appleby J: Getting the most out of PROMs: putting health outcomes at the heart of NHS decision-making, 2010. Available at: https://www.kingsfund.org.uk/sites/default/files/Getting-the-most-out-of-PROMs-Nancy-Devlin-John-Appleby-Kings-Fund-March-2010.pdf. Accessed March 14, 2018

11. Nimmo A, Bell S, Brunton C, et al. Collection and determinants of patient reported outcome measures in haemodialysis patients in Scotland. *QJM* 2018;111(1):15-21

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

13. Fukuhara S, Lopes AA, Bragg-Gresham JL, et al. Health-related quality of life among dialysis patients on three continents: the Dialysis Outcomes and Practice Patterns Study. *Kidney Int* 2003;64(5):1903-10

14. Raj R, Ahuja KDK, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

15. de Goeij MC, Ocak G, Rotmans JI, Eijgenraam JW, Dekker FW, Halbesma N. Course of symptoms and health-related quality of life during specialized pre-dialysis care. *PLoS One* 2014;9(4):e93069

16. Pugh-Clarke K, Read SC, Sim J. Symptom experience in non-dialysis-dependent chronic kidney disease: a qualitative descriptive study. *J Ren Care* 2017;43(4):197-208

17. Stephens RJ, Hopwood P, Girling DJ, Machin D. Randomized trials with quality of life endpoints: are doctors' ratings of patients' physical symptoms interchangeable with patients' self-ratings? *Qual Life Res* 1997;6(3):225-36

18. Weisbord SD, Fried LF, Mor MK, et al. Renal provider recognition of symptoms in

patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2007;2(5):960-7

19. van der Veer SN, Aresi G, Gair R. Incorporating patient-reported symptom assessments into routine care for people with chronic kidney disease. *Clin Kidney J* 2017;10(6):783-7

20. Davison SN, Levin A, Moss AH, et al. Executive summary of the KDIGO controversies conference on supportive care in chronic kidney disease: developing a roadmap to improving quality care. *Kidney Int* 2015;88(3):447-59

21. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. *Clin J Am Soc Nephrol* 2014;9(10):1813-21

22. Chronic kidney disease data collection reference guide, International Consortium for Health Outcomes Measurement (ICHOM), 2017

23. Thong MS, van Dijk S, Noordzij M, et al. Symptom clusters in incident dialysis patients: associations with clinical variables and quality of life. *Nephrol Dial Transplant* 2009;24(1):225-30

24. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27(5):1159-70

25. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27(5):1147-57

26. Flythe JE, Powell JD, Poulton CJ, et al. Patient-reported outcome instruments for physical symptoms among patients receiving maintenance dialysis: a systematic review. *Am J Kidney Dis* 2015;66(6):1033-46

27. Kidney Decision Aid, Dutch Kidney Patients Association [Nierwijzer, Nierpatienten Vereniging Nederland (NVN)], 2017. www.nierwijzer.nl. Accessed 26 February, 2018

28. Basch E, Abernethy AP, Mullins CD, et al. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. *J Clin Oncol* 2012;30(34):4249-55

29. Lloyd GG, Guthrie E. Handbook of Liaison Psychiatry. Cambridge University Press: 2007; 96-8.

30. 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. *Am J Kidney Dis* 2001;37(1):11-21

31. Almutary H, Bonner A, Douglas C. Which patients with chronic kidney disease have the greatest symptom burden? A comparative study of advanced CKD stage and dialysis modality. *J Ren Care* 2016;42(2):73-82

32. Curtin R, C Bultman D, Thomas-Hawkins C, A J Walters B, Schatell D. Hemodialysis patients' symptom experiences: effects on physical and mental functioning. *Nephrol Nurs J* 2003;29(6):562-74

33. Danquah FV, Zimmerman L, Diamond PM, Meininger J, Bergstrom N. Frequency, severity, and distress of dialysis-related symptoms reported by patients on hemodialysis. *Nephrol Nurs J* 2010;37(6):627-38; quiz 39

34. Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *J Pain Symptom Manage* 2004;27(3):226-40

35. Murphy EL, Murtagh FE, Carey I, Sheerin NS. Understanding symptoms in patients with advanced chronic kidney disease managed without dialysis: use of a short patient-completed assessment tool. *Nephron Clin Pract* 2009;111(1):c74-80

36. Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Development of the kidney disease quality of life (KDQOL) instrument. *Qual Life Res* 1994;3(5):329-38

37. Portenoy RK, Thaler HT, Kornblith AB, et al. The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. *Eur J Cancer* 1994;30a(9):1326-36

38. Chang VT, Hwang SS, Feuerman M, Kasimis BS, Thaler HT. The Memorial Symptom Assessment Scale Short Form (MSAS-SF). *Cancer* 2000;89(5):1162-71

39. Murtagh FE, Addington-Hall JM, Edmonds PM, et al. Symptoms in advanced renal disease: a cross-sectional survey of symptom prevalence in stage 5 chronic kidney disease managed without dialysis. *J Palliat Med* 2007;10(6):1266-76

40. de Haes JCJM, Olschewski M, Fayers P, et al.: Measuring the quality of life of cancer patients with the Rotterdam Symptom Checklist (RSCL): a manual, Groningen, Northern Centre for Healthcare Research (NCH), 2012

41. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

42. Flythe JE, Hilliard T, Castillo G, et al. Symptom prioritization among adults receiving in-center hemodialysis: a mixed methods study. *Clin J Am Soc Nephrol* 2018;13(5):735-45

43. Marcinowicz L, Chlabicz S, Grebowski R. Open-ended questions in surveys of patients' satisfaction with family doctors. *J Health Serv Res Policy* 2007;12(2):86-9

44. Bingham CO, 3rd, Noonan VK, Auger C, Feldman DE, Ahmed S, Bartlett SJ. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 4: patient-reported outcomes can inform clinical decision making in chronic care. *J Clin Epidemiol* 2017;89:136-41

45. O'Hare AM, Armistead N, Schrag WL, Diamond L, Moss AH. Patient-centered care: an opportunity to accomplish the "Three Aims" of the National Quality Strategy in the Medicare ESRD program. *Clin J Am Soc Nephrol* 2014;9(12):2189-94

46. O'Hare AM. Patient-centered care in renal medicine: five strategies to meet the challenge. *Am J Kidney Dis* 2018;71(5):732-6

47. Kliger AS. Quality measures for dialysis: time for a balanced scorecard. *Clin J Am Soc Nephrol* 2016;11(2):363-8

48. Nissenson AR. Improving outcomes for ESRD patients: shifting the quality paradigm. *Clin J Am Soc Nephrol* 2014;9(2):430-4

49. Howell D, Molloy S, Wilkinson K, et al. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. *Ann Oncol* 2015;26(9):1846-58

50. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 2006;12(5):559-68

51. Anderson NE, Calvert M, Cockwell P, Dutton M, Kyte D. The Use of Patient-Reported Outcomes in Patients Treated With Maintenance Hemodialysis: A Perspective. *Am J Kidney Dis* 2019

52. Wiering B, de Boer D, Delnoij D. Patient involvement in the development of patient-reported outcome measures: the developers' perspective. *BMC Health Serv Res* 2017;17(1):635

53. De Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: A Practical Guide. Cambridge University Press, United Kingdom: 2011; 338 p

Supplemental Material for Chapter 2

Supplementary item S1. Search string for systematic literature search for symptom questionnaires used in patients with chronic kidney disease.

2

(((("Chronic Kidney Disease"[ti] OR "Chronic Kidney Diseases"[ti] OR "Chronic Renal Disease"[ti] OR "Chronic Renal Diseases"[ti] OR "CKD"[ti] OR "End-Stage Renal Disease"[ti] OR "End-Stage Renal Diseases"[ti] OR "ESRD"[ti] OR "End-Stage Kidney Disease"[ti] OR "End-Stage Kidney Diseases"[ti] OR "Advanced Renal Disease"[ti] OR "Advanced Kidney Disease"[ti] OR "Renal Insufficiency, Chronic"[majr] OR "Chronic Renal Insufficiency"[ti] OR "Chronic Renal Failure"[ti] OR "Chronic Kidney Failure"[ti] OR (("Kidney Diseases"[majr] OR "kidney disease"[ti] OR "renal disease"[ti] OR "kidney diseases"[ti] OR "renal diseases"[ti] OR "kidney failure"[ti] OR "renal failure"[ti] OR "renal insufficiency"[ti] OR "kidney insufficiency"[ti]) AND ("Chronic Disease"[majr] OR "chronic"[ti] OR chronic*[ti])) OR (("pre-dialysis"[ti] OR pre-dialy*[ti] OR "predialysis"[ti] OR predial*[ti] OR "chronic renal"[ti] OR "chronic kidney"[ti] OR "Renal Insufficiency, Chronic"[majr] OR "Kidney Failure, Chronic"[majr] OR "end stage renal"[ti] OR "end stage kidney"[ti]) AND ("3"[ti] OR "4"[ti] OR "5"[ti] OR "three"[ti] OR "four"[ti] OR "five"[ti] OR "iii"[ti] OR "iv"[ti] OR "v"[ti]) AND ("stage"[ti] OR "stages"[ti] OR "late"[ti])) OR "Renal Replacement Therapy"[majr] OR "Renal Replacement Therapy"[ti] OR "RRT"[ti] OR "hemodialysis"[ti] OR "haemodialysis"[ti] OR "peritoneal dialysis"[ti] OR "Kidney Transplantation"[ti] OR "Renal Transplantation"[ti] OR Kidney Transplant*[ti] OR Renal Transplant*[ti] OR "Dialysis"[majr] OR "Dialysis"[ti] OR "hemodiafiltration"[ti] OR "haemodiafiltration"[ti])

AND ("Signs and Symptoms"[Mesh:noexp] OR "Signs and Symptoms"[majr] OR "Symptom"[ti] OR "symptoms"[ti] OR "Symptom burden"[tw] OR symptom*[ti])

AND ("Surveys and Questionnaires"[Mesh:noexp] OR "Patient Reported Outcome Measures"[Mesh] OR "Questionnaire"[tw] OR "Questionnaires"[tw] OR Questionnair*[tw] OR "Patient-Reported Outcome Measure"[tw] OR "Patient-Reported Outcome Measures"[tw] OR "PROMs"[tw] OR "PROM"[tw] OR "Self Report"[tw] OR "assessment instrument"[tw] OR "assessment system"[tw] OR "assessment method"[tw] OR "assessment instruments"[tw] OR "assessment systems"[tw] OR "assessment methods"[tw] OR "Assessment Scale"[tw] OR "Assessment Scales"[tw] OR "instrument"[ti] OR "scale"[ti] OR "checklist"[ti] OR "score"[ti] OR "instruments"[ti] OR "scales"[ti] OR "checklists"[ti] OR "scores"[ti] OR "inventory"[ti] OR "inventories"[ti] OR "Symptom Burden Index"[tw] OR "symptom burden instrument"[tw] OR "symptom burden measures"[tw] OR "symptom burden score"[tw] OR "symptom burden scores"[tw])

AND (english[la] OR dutch[la])

NOT (("Adolescent"[mesh] OR Adolescen*[ti] OR "Child"[mesh] OR "child"[ti] OR "children"[ti] OR "girl"[ti] OR "girls"[ti] OR "boy"[ti] OR "boys"[ti]) NOT ("Adult"[mesh] OR

"adult"[ti] OR "adults"[ti]))**)**

OR

(("Chronic Kidney Disease"[ti] OR "Chronic Kidney Diseases"[ti] OR "Chronic Renal Disease"[ti] OR "Chronic Renal Diseases"[ti] OR "CKD"[ti] OR "End-Stage Renal Disease"[ti] OR "End-Stage Renal Diseases"[ti] OR "ESRD"[ti] OR "End-Stage Kidney Disease"[ti] OR "End-Stage Kidney Diseases"[ti] OR "Advanced Renal Disease"[ti] OR "Advanced Kidney Disease"[ti] OR "Renal Insufficiency, Chronic"[majr] OR "Chronic Renal Insufficiency"[ti] OR "Chronic Renal Failure"[ti] OR"Chronic Kidney Failure"[ti] OR (("Kidney Diseases"[majr] OR "kidney disease"[ti] OR "renal disease"[ti] OR "kidney diseases"[ti] OR "renal diseases"[ti] OR "kidney failure"[ti] OR "renal failure"[ti] OR "renal insufficiency"[ti] OR "kidney insufficiency"[ti]) AND ("Chronic Disease"[majr] OR "chronic"[ti] OR chronic*[ti])) OR (("pre-dialysis"[ti] OR pre-dialy*[ti] OR "predialysis"[ti] OR predial*[ti] OR "chronic renal"[ti] OR "chronic kidney"[ti] OR "Renal Insufficiency, Chronic"[majr] OR "Kidney Failure, Chronic"[majr] OR "end stage renal"[ti] OR "end stage kidney"[ti]) AND ("3"[ti] OR "4"[ti] OR "5"[ti] OR "three"[ti] OR "four"[ti] OR "five"[ti] OR "iii"[ti] OR "iv"[ti] OR "v"[ti]) AND ("stage"[ti] OR "stages"[ti] OR "late"[ti])) OR "Renal Replacement Therapy"[majr] OR "Renal Replacement Therapy"[ti] OR "RRT"[ti] OR "hemodialysis"[ti] OR "haemodialysis"[ti] OR "peritoneal dialysis"[ti] OR "Kidney Transplantation"[ti] OR "Renal Transplantation"[ti] OR Kidney Transplant*[ti] OR Renal Transplant*[ti] OR "Dialysis"[majr] OR "Dialysis"[ti] OR "hemodiafiltration"[ti] OR "haemodiafiltration"[ti])

AND ("Signs and Symptoms"[Mesh:noexp] OR "Signs and Symptoms"[majr] OR "Symptom"[tw] OR "symptoms"[tw] OR "Symptom burden"[tw] OR symptom*[tw])

AND ("Surveys and Questionnaires"[majr:noexp] OR "Patient Reported Outcome Measures"[majr] OR "Questionnaire"[ti] OR "Questionnaires"[ti] OR Questionnair*[ti] OR "Patient-Reported Outcome Measure"[ti] OR "Patient-Reported Outcome Measures"[ti] OR "PROMs"[ti] OR "PROM"[ti] OR "Self Report"[ti] OR "assessment instrument"[ti] OR "assessment system"[ti] OR "assessment method"[ti] OR "assessment instruments"[ti] OR "assessment systems"[ti] OR "assessment methods"[ti] OR "Assessment Scale"[ti] OR "Assessment Scales"[ti] OR "instrument"[ti] OR "scale"[ti] OR "checklist"[ti] OR "score"[ti] OR "instruments"[ti] OR "scales"[ti] OR "checklists"[ti] OR "scores"[ti] OR "inventory"[ti] OR "inventories"[ti] OR "Symptom Burden Index"[ti] OR "symptom burden instrument"[ti] OR "symptom burden measures"[ti] OR "symptom burden score"[ti] OR "symptom burden scores"[ti])

AND (english[la] OR dutch[la])

NOT (("Adolescent"[mesh] OR Adolescen*[ti] OR "Child"[mesh] OR "child"[ti] OR "children"[ti] OR "girl"[ti] OR "girls"[ti] OR "boy"[ti] OR "boys"[ti]) NOT ("Adult"[mesh] OR "adult"[ti] OR "adults"[ti]))**))**

Supplementary table S2. Unique symptoms identified from questionnaires and interviews with patients with chronic kidney disease, divided into ten symptom clusters.

General symptoms

Fatigue / feeling tired / lack of energy
Change in weight
Difficulty concentrating
Feeling sick
Pain (in general)
Changes in appearance

Night's rest

Trouble falling asleep
Trouble staying asleep
Changes in amount of sleep
Drowsiness

Gastroenterology

Constipation
Nausea
Vomiting
Diarrhoea
Decreased appetite / lack of appetite
Feeling of fullness or bloating
Abdominal pain / stomach cramps
Heartburn
Stomach or bowel problems
Overeating / food cravings

Cardiopulmonary

Chest pain
Heart palpitations / arrhythmia
Easy bruising
Slow-healing sores
Shortness of breath / dyspnoea
Coughing
Wheezing
Swelling in legs / feet
Chest tightness
Nycturia

Central nervous system

Light-headedness or dizziness
Numbness in feet or hands
Tingling in feet or hands
Headache
Restless legs or difficulty keeping legs still
Shivering / hot or cold spells
Trembling
Trouble remembering things / memory loss
Sluggish / react slowly
Difficulty keeping attention
Inadequate / having to (double-)check what you do
Pain / burning / frequency of urination

Musculoskeletal

Muscle loss
Muscle cramps
Stiffening of joints
Bone or joint pain / pain in arms, legs or joints
Muscle soreness
Back pain
Muscle spasm

Muscle weakness
Poor mobility
Pelvic pain
Humps in muscles*

Skin

Dry skin
Itching / pruritus
Changes in skin
Loss of hair
Sweating

Head/throat

Dry mouth
Thirst
Change in taste
Pain when swallowing
Sore throat
Burning / sore eyes
Sore mouth
Hearing loss
Ringing in your ears
Impaired visual ability

Psychosocial

Feeling nervous
Feeling irritable
Feeling sad
Feeling anxious
Confusion
Worrying
Depressed mood
Restless
Tension / feeling tense or keyed up
Feeling blue
Feeling frustrated
Feeling angry
Feeling bored
Feeling lonely
Lack of vitality
Decreased motivation
Feel worn out
Difficulty making decisions
Feeling everything is an effort
Feeling of being trapped or caught
Feelings of guilt
Thoughts of ending your life
Difficulties with family life and social contacts
Feeling critical of others
Difficulties to trust others
Intrusive thoughts
Personality changes
Emotional swings
Despairing about the future
A lump in your throat

Sex

Decreased interest in sex
Difficulty becoming sexually aroused
Inability to relax and enjoy sex

**Symptom retrieved from videotaped interviews with patients with chronic kidney disease.*

Supplementary table S3. Comparison of two CKD-specific symptom questionnaires based on feedback of transplant and non-transplant patients.

	DSI (N=76)		IPOS-Renal (N=75)		
	Transplant (N=53)	**Non-transplant (N=23)**	**Transplant (N=47)**	**Non-transplant (N=28)**	***p*-value**
Age (years)	60.4 (11.6)	61.2 (14.6)	59.7 (10.9)	61.1 (9.7)	0.941
Objective time to complete* (minutes)	5.0 (1.6)	6.3 (1.4)	7.4 (1.9)	7.6 (1.6)	**0.001**[a]
Subjective time to complete* (minutes)	3.0 (1.8)	3.9 (1.7)	5.0 (1.7)	4.4 (1.5)	**<0.001**[b]
Number of symptoms reported^	11.0 (6.7)	14.4 (5.5)	7.4 (4.1)	9.0 (3.9)	**<0.001**[c]
Additional 1-3 symptoms reported[#]	12 (22.6)	9 (39.1)	18 (38.1)	7 (25.0)	0.251
Burdensome of questionnaire (yes)	4 (7.5)	0 (0.0)	1 (2.2)	1 (3.7)	0.387
Appropriate frequency of submission (times per year)	2.7 (2.0)	2.9 (1.6)	2.4 (1.4)	3.8 (2.9)	0.072

Values are shown in N (%) or mean (SD).

[a] *Post hoc tests: differences are statistically significant between DSI transplant and IPOS-Renal transplant, and between DSI transplant and IPOS-Renal non-transplant groups.* [b] *Post hoc tests: differences are statistically significant between DSI transplant and IPOS-Renal transplant, and between DSI transplant and IPOS-Renal non-transplant groups.* [c] *Post hoc tests: differences are statistically significant between DSI transplant and IPOS-Renal transplant, between DSI non-transplant and IPOS-Renal transplant, and between DSI non-transplant and IPOS-Renal non-transplant groups.*

**Objective time to complete was defined as the difference in minutes between the start and completion of the online questionnaire. Subjective time to complete is the time to complete estimated by the patient. Values shown as geometric mean (SD).*

^*The number of symptoms reported is based on the symptoms defined in the questionnaire and rated by the patient as bothering a little bit to very much (or affecting slightly to overwhelmingly).*

[#]*The number of patients reporting an additional 1 to 3 symptoms not mentioned in de questionnaire.*

CKD, Chronic Kidney Disease; PROMs, Patient Reported Outcome Measures; DSI, Dialysis Symptom Index; IPOS-Renal, Palliative Care Outcome Scale - Renal Version.

3

Validity and reliability of Patient-Reported Outcomes Measurement Information System (PROMIS®) using Computerized Adaptive Testing (CAT) in patients with advanced chronic kidney disease

Esmee M. van der Willik, Fenna van Breda, Brigit C. van Jaarsveld, Marlon van de Putte, Isabelle W. Jetten, Friedo W. Dekker, Yvette Meuleman, Frans J. van Ittersum, Caroline B. Terwee

Nephrology Dialysis Transplantation. 2022;gfac231.

Abstract

Background: The Patient-Reported Outcomes Measurement Information System (PROMIS®) has been recommended for computerized adaptive testing (CAT) of health-related quality of life (HRQOL). This study compared the content, validity and reliability of seven PROMIS CATs to the 12-item Short-Form Health Survey (SF-12) in patients with advanced chronic kidney disease (CKD).

Methods: Adult CKD patients with an eGFR<30 ml/min.1.73m^2 not receiving dialysis treatment completed seven PROMIS CATs (assessing physical function, pain interference, fatigue, sleep disturbance, anxiety, depression, and ability to participate in social roles and activities), the SF-12 and, additionally, the PROMIS Pain Intensity single item and Dialysis Symptom Index (DSI) at inclusion and 2-weeks. A content comparison was performed between PROMIS CATs and SF-12. Construct validity of PROMIS CATs was assessed using Pearson's correlations. Test-retest reliability of all patient-reported outcome measures (PROMs) was assessed by calculating the intra-class correlation coefficient (ICC) and minimal detectable change (MDC).

Results: In total, 207 patients participated in the study. A median of 45 items (10 minutes) was completed for PROMIS CATs. All PROMIS CATs showed evidence for sufficient construct validity. PROMIS CATs, most SF-12 domains and summary scores, and DSI showed sufficient test-retest reliability (ICC≥0.70). PROMIS CATs had a lower MDC compared to the SF-12 (5.7-7.4 compared to 11.2-21.7 across domains, respectively).

Conclusion: PROMIS CATs showed sufficient construct validity and test-retest reliability in patients with advanced CKD. PROMIS CATs required more items but showed better reliability than the SF-12. Future research is needed to investigate the feasibility of PROMIS CATs for routine nephrology care.

Introduction

Patients with advanced chronic kidney disease (CKD) experience numerous physical and emotional disease-related symptoms, which are associated with a decreased health-related quality of life (HRQOL).[1-4] Although several symptoms and the impact on physical, mental, and social functioning have been considered of great importance by patients and healthcare professionals[5, 6], these patient-relevant outcomes may still be regularly underrecognized and therefore insufficiently managed in routine nephrology care.[4, 7] Patient-reported outcome measures (PROMs) can be used to improve insight into these important outcomes. PROMs have been incorporated into Dutch routine dialysis care[3] and are now also being implemented into the care for Dutch patients with advanced CKD and kidney transplant recipients[8].

Many different generic and disease-specific PROMs are being used within and across countries.[9, 10] In Dutch nephrology care, the 12-item Short-Form Health Survey (SF-12) and the Dialysis Symptom Index (DSI) are used to assess generic HRQOL and disease-related symptom burden, respectively.[3] A major advantage of using the same PROMs is that this enables comparison and monitoring of outcomes across CKD stages and treatments.

Recently, the Patient-Reported Outcomes Measurement Information System (PROMIS®) was selected as one of the recommended PROMs to measure generic HRQOL in patients with CKD by a consensus group of the International Consortium of Health Outcomes Measurement (ICHOM).[11] Additionally, PROMIS was recommended by the Linnean initiative, a nationwide network of stakeholders in The Netherlands, for all patient populations, to standardize outcome measurement across medical conditions.[12] PROMIS consists of a collection of item banks (i.e. large sets of questions), developed to measure commonly relevant domains across patient conditions, such as physical function, fatigue and anxiety. Because PROMIS item banks were developed using item response theory (IRT) models, they can also be administered as Computerized Adaptive Test (CAT). The use of CAT is relatively novel in healthcare and has several advantages compared to traditional fixed (i.e. non-adaptive) PROMs. In a CAT, the computer selects questions from an item bank based on the answers to previous questions. With this method, the PROM is adapted to the patient, resulting in questions that are likely more relevant to the patient. In addition, on average less questions will be required to obtain similar or even more precise measurements compared to fixed PROMs.[13, 14] Sufficient validity and reliability of fixed PROMIS measures was found in several disease populations[15-17], including patients with CKD.[18, 19] However, the psychometric properties of PROMIS CATs have not yet been studied in patients with CKD.

Therefore, this study aimed to examine and compare the content, construct validity and test-retest reliability (including minimal detectable change) of seven PROMIS CATs (assessing physical function, pain interference, fatigue, sleep disturbance, anxiety, depression, and ability to participate in social roles and activities) with the SF-12 in patients with advanced CKD. Additionally, we assessed test-retest reliability of the PROMIS Pain Intensity single item and the DSI, as these PROMs are often used together with the PROMIS CATs and SF-12.

Methods

Study design and population

This observational study included adult patients with advanced CKD with an estimated glomerular filtration rate (eGFR) <30 ml/min.1.73m^2, not receiving dialysis treatment. Exclusion criteria were start with kidney replacement therapy (KRT; dialysis or kidney transplantation) planned within 4 weeks, rapid deterioration of kidney function (i.e. decrease in eGFR of >20 ml/min.1.73 m^2 during the last 6 months), not able to complete PROMs due to cognitive impairment, poor knowledge of the Dutch language, or no informed consent. Patients were recruited between November 2020 and August 2021 by their nephrologist at the outpatient clinics of Amsterdam University Medical Centre in Amsterdam and Niercentrum aan de Amstel in Amstelveen, the Netherlands. Eligible patients received written information by mail and were, if needed, approached by telephone after 2 weeks for further information. After providing written informed consent, patients were invited by e-mail to complete the PROMs digitally at the KLIK ('Kwaliteit van Leven In Kaart'; www.hetklikt.nu) research platform at inclusion (i.e. baseline), after 2 weeks and after 6 months. If necessary, two reminders were sent by e-mail or patients were contacted by telephone. Patients without access to an electronic device with internet connection could participate by telephone. In this study, the baseline and 2-weeks measurements were used (Figure 1).

The study was reviewed by the Medical Ethics Review Committee of VU University Medical Centre in the Netherlands, which confirmed that the Medical Research Involving Human Subjects Act (WMO) does not apply.

Measures

Demographic and clinical characteristics, including age, sex, primary kidney disease according to European Renal Association codes[20], body mass index (BMI), smoking status, comorbidities (hypertension, diabetes mellitus, cardiovascular disease, lung disease, liver disease and malignancy) as defined by ICHOM[11], eGFR

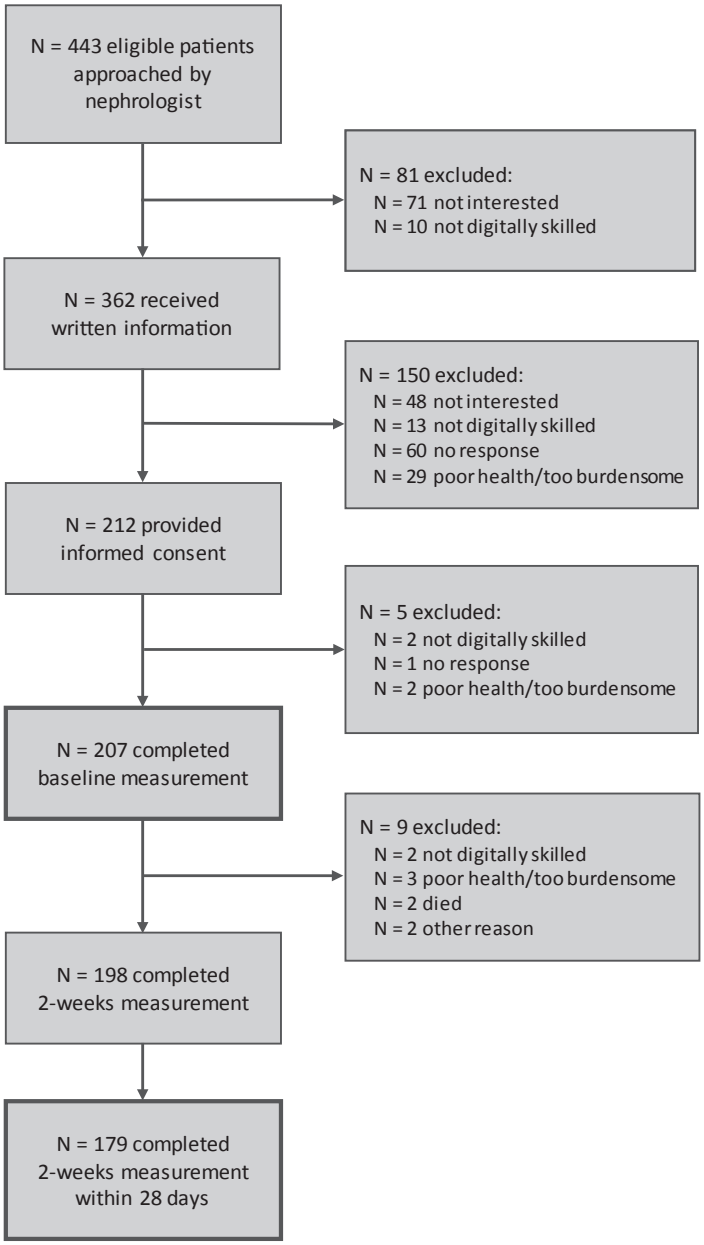


Figure 1: Flow diagram of patient inclusion for baseline and 2-weeks measurements.
All patients that completed the baseline measurement constitute the study sample for validity analyses. All patients that completed the 2-weeks measurement within 28 days after baseline are included for reliability analyses.
The reason for exclusion was indicated by the patient. Patients who were not digitally skilled were offered participation by telephone, but were not willing to participate in that manner.

(ml/min/1.73m^2), KRT in medical history, start of KRT and death during follow-up were collected from medical records. Educational level and ethno-cultural background were self-reported at baseline.

The PROMs included in this study are seven PROMIS CATs, the SF-12, one PROMIS single item and the DSI. The SF-12 and DSI have demonstrated validity within patients with CKD.[10, 21-24] PROMs were presented in random order across patients, but with fixed order within patients during follow-up. The research platform to complete PROMs did not allow for any missing values within a PROM.

Seven Dutch-Flemish PROMIS CATs[25] were administered: v1.2 Physical Function, v1.1 Pain Interference, v1.0 Fatigue, v1.0 Sleep Disturbance, v1.0 Anxiety, v1.0 Depression, and v2.0 Ability to Participate in Social Roles and Activities. All items have 5 response options (e.g. ranging from 'never' to 'always' or from 'not at all' to 'very much'). PROMIS CATs are presented as T-scores where 50 (SD: 10) represents the average score of the US general population. A difference of >2 points was considered relevant.[26] Higher scores indicate more of the construct (e.g. a higher Depression score means more depression, a higher Physical Function score means more [better] function). Within each PROMIS CAT, questions were selected one-by-one from an underlying item bank. The starting item is the item with the highest information value for the average level of the domain in the general population. The next items are subsequently selected from the item bank based on the respondent's answers to previous items. For example: a respondent reports to have difficulties with doing two hours of physical labor (first item). Then the second item will be a more 'easy' activity, e.g. a question about ability to do chores such as vacuuming. The respondent is not asked about more 'difficult' activities (e.g. running five miles) that (s)he is assumably not able to do. By tailoring the next item to the person's ability, questions are more often relevant to that person and on average less questions need to be completed. (See Supplement A for a visual illustration of CAT). After each item, the score and standard error (SE) are estimated based on all items completed so far. In this study, the CAT stopped when a SE of 2.2 on the T-score metric was reached (comparable to a reliability of approximately 0.95) or when a maximum of 12 items per CAT was administered. We used a lower SE compared to the standard stopping rule (i.e. SE: 3.0)[13], because a higher reliability may be preferable for routine care and by using this setting, the optimal performance of PROMIS CATs could be investigated. PROMIS CATs were administered using CAT software of the Dutch-Flemish Assessment Center, part of the Dutch-Flemish PROMIS National Center[27].

The SF-12 v2[28, 29] is a 12-item generic PROM assessing 8 domains of HRQOL: physical functioning, role-physical, bodily pain, general health, vitality, social function-

ing, role-emotional and mental health. Additionally, a physical component summary score (including physical functioning, role-physical, bodily pain and general health) and a mental component summary score (including vitality, social functioning, role-emotional and mental health) can be calculated. Domain and summary scores range from 0 to 100 and the US general population is used as reference with an average score of 50 (SD: 10). Higher scores indicate a better HRQOL.
The PROMIS item v1.0 Numerical Rating Scale Pain Intensity 1a is a single item with a 0-10 scale, with higher scores indicating more pain.
The DSI[21] is a 30-item disease-specific PROM to assess physical and emotional symptom burden. Patients report the presence of 30 symptoms (yes/no) during the past week and, if present, the burden of each symptom on a 5-point Likert scale ranging from 1 'not at all' to 5 'very much' bothersome. Two overall scores were calculated: 1) total number of symptoms (0-30 symptoms), and 2) total symptom burden score, which is the sum of burden on individual symptoms ranging from 0 (no symptoms) to 150 (all 30 symptoms are very much bothersome).[3, 30] The DSI items 'feeling tired or lack of energy', 'feeling anxious', 'trouble falling asleep' and 'trouble staying asleep' (hereafter combined as 'sleep problems') were used as comparison items in the construct validity analyses since these items intend to measure constructs comparable to the PROMIS CATs Fatigue, Anxiety and Sleep Disturbance.

Content comparison

To provide insight into the comparability of PROMIS CATs and SF-12, their content was compared by providing 1) an overview of the PROM characteristics (e.g. domains, number of items, recall period, scoring and interpretation), and 2) a visual comparison of the domain score distributions using an interpretative color indication (from green [better] to red [worse] HRQOL), in line with the use in routine care.[31, 32]

Construct validity

Construct validity of PROMIS CATs was assessed using Pearson's correlations. Hypotheses were formulated a priori about the expected correlations between PROMIS CATs and SF-12 and DSI, based on literature[15-18] and expert judgement (EvdW and CT). We expect strong correlations ($r \geq 0.7$) between PROMIS CATs and comparable SF-12 domains and similar DSI items, moderate correlations ($r = 0.5$-0.7) between PROMIS CATs and largely related SF-12 domains, and no strong correlations for other comparisons ($r \leq 0.6$) (see Table 1). Construct validity was considered sufficient if ≥75% of the results was in accordance with the hypotheses.

Table 1. Hypotheses for construct validity

PROMIS CAT	Strong correlation Pearson's r ≥ 0.7	Moderate correlation Pearson's r 0.5-0.7	No strong correlation Pearson's r ≤ 0.6
Physical Function	SF-12 physical functioning SF-12 physical component summary*	SF-12 general health SF-12 bodily pain	All other SF-12 domains DSI total number of symptoms and symptom burden score
Pain Interference	SF-12 bodily pain	SF-12 physical functioning SF-12 physical component summary	All other SF-12 domains DSI total number of symptoms and symptom burden score
Fatigue	SF-12 vitality DSI feeling tired or lack of energy (1 item)		All other SF-12 domains DSI total number of symptoms and symptom burden score
Sleep Disturbance	DSI sleep problems (2 items)#		All other SF-12 domains DSI total number of symptoms and symptom burden score
Anxiety	SF-12 mental health SF-12 mental component summary* DSI feeling anxious (1 item)		All other SF-12 domains DSI total number of symptoms and symptom burden score
Depression	SF-12 mental health SF-12 mental component summary*		All other SF-12 domains DSI total number of symptoms and symptom burden score
Ability to Participate in Social Roles and Activities	SF-12 social functioning	SF-12 role physical SF-12 role emotional	All other SF-12 domains DSI total number of symptoms and symptom burden score

** SF-12 physical component summary includes the domains physical functioning, role-physical, bodily pain and general health; SF-12 mental component summary includes the domains vitality, social functioning, role-emotional and mental health.*

 # DSI Sleep problems were defined as trouble falling asleep and/or trouble staying asleep.

Test-retest reliability

Test-retest reliability of PROMIS CATs, SF-12, PROMIS Pain Intensity single item, and DSI was assessed by calculating the intra-class correlation coefficient (ICC) in patients with valid baseline and 2-weeks measurements (Figure 1). The ICC was calculated using a two-way random-effects model for absolute agreement: *ICC agreement* = $\sigma^2_p / (\sigma^2_p + \sigma^2_m + \sigma^2_e)$, whereby σ^2_p is the variation between patients, σ^2_m is the variation between measurements and σ^2_e is random error variance. An ICC ≥0.70 was considered sufficient.[33]

The ICC was computed for each PROMIS CAT and SF-12 domain separately. Additionally, the ICC was calculated for the PROMIS Pain Intensity single item and for the DSI total number of symptoms and symptom burden score. Although the DSI was not designed to be interpreted as an overall score (as it measures 30 different symptoms), the total number of symptoms and symptom burden score are often used within healthcare, and insight into the reliability of these scores is therefore of clinical relevance.

The minimal detectable change (MDC) was also calculated for each domain of the PROMIS CATs and SF-12, the PROMIS Pain Intensity single item, and the DSI total number of symptoms and symptom burden score. The MDC is a parameter of measurement error and is defined as the "smallest change in score that can be detected beyond measurement error", with 95% confidence.[33] Two different methods were applied to calculate the MDC, in line with the underlying measurement theories, namely classical test theory (CTT) or IRT, that assume a constant or varying standard error of measurement (SEM) across the PROM-scale, respectively.[34, 35]

The MDC, based on CTT, of the SF-12 domains, PROMIS Pain Intensity single item, and the DSI total number of symptoms and symptom burden score was calculated using the formula: $1.96 * \sqrt{2} * SEM$, whereby SEM was calculated as: $\sqrt{(\sigma^2_m + \sigma^2_e)}$.

The MDC, based on IRT, of each PROMIS CAT varies per patient (because with IRT the SE of each score is different) and was calculated using the following formula: $1.96 * \sqrt{(SE_1^2 + SE_2^2)}$, whereby SE_1 is the patient's IRT estimated standard error of the T-score at baseline and SE_2 at the 2-weeks measurement. A mean MDC of each PROMIS CAT was subsequently calculated for the whole group.

Data analyses were performed using SPSS V.25.0 (IBM Corp., Armonk, NY, USA).

Table 2. Characteristics of study sample at baseline and 2-weeks measurements.

	Study sample at baseline* (n=207)	Study sample at 2 weeks* (n=179)
Sex, male	124 (59.9)	107 (59.8)
Age, years	65.5 (13.8)	66.1 (13.1)
Ethno-cultural group[$], Dutch	176 (85.0)	152 (84.9)
Educational level[#]		
Low	85 (41.0)	74 (41.3)
Middle	49 (23.7)	43 (24.0)
High	73 (35.3)	62 (34.6)
Primary kidney disease		
Glomerulonephritis/sclerosis	34 (16.6)	33 (18.6)
Pyelonephritis	7 (3.4)	7 (4.0)
Polycystic kidney disease	16 (7.8)	15 (8.5)
Other congenital/hereditary kidney diseases	15 (7.3)	13 (7.3)
Hypertension/renal vascular disease	46 (22.5)	42 (23.7)
Diabetes mellitus	14 (6.8)	12 (6.8)
Miscellaneous	63 (30.7)	49 (27.7)
Unknown	10 (4.9)	6 (3.4)
Kidney function, eGFR	21.4 (6.7)	21.6 (6.6)
KRT in medical history[£], yes	35 (17.0)	30 (16.9)
BMI	26.8 (5.2)	26.9 (5.2)
Smoking		
Yes	25 (13.2)	19 (11.7)
No, stopped	94 (49.7)	82 (50.6)
No, never smoked	70 (37.0)	61 (37.7)
Comorbidities		
Hypertension, yes	164 (79.2)	140 (78.2)
Diabetes mellitus, yes	62 (30.0)	53 (29.6)
Cardiovascular disease, yes	53 (25.6)	43 (24.0)
Lung disease, yes	30 (14.5)	28 (15.6)
Liver disease, yes	11 (5.3)	8 (4.5)
Malignancy, yes	50 (24.2)	43 (24.0)

Values are shown in n (%) or mean (SD).
** Study sample at baseline was used for validity analyses. Study sample at 2-weeks measurement was used for reliability analyses.*
$ Self-reported ethno-cultural group: "What ethnic group do you consider yourself to belong to?"
 # Educational level according to International Standard Classification of Education (ISCED) levels 2011, classified as low: primary, lower secondary or lower vocational education; middle: upper secondary or upper vocational education; high: tertiary education (college/university).
£ KRT in medical history includes patients who have undergone (temporary) dialysis treatment or a kidney transplant in the past. At study inclusion, all patients had an eGFR<30 and did not require dialysis treatment, in accordance with inclusion criteria.
Missing values at baseline: primary kidney disease: n=2 (1.0%); KRT in medical history: n=1 (0.5%); BMI: n=11 (5.3%); smoking: n=18 (8.7%). Missing values at 2 weeks: primary kidney disease: n=2 (1.1%); KRT in medical history: n=1 (0.6%); BMI: n=9 (5.0%); smoking: n=17 (9.5%).
Abbreviations: eGFR, estimated glomerular filtration rate; KRT, kidney replacement therapy; BMI, body mass index.

Results

Study participants

Almost half of the patients that were approached provided written informed consent. In total, 207 participants completed the baseline measurement and were included in current analyses. Of them, 179 (86.5%) participants completed the 2-weeks measurement within 28 days and were eligible for reliability analyses (Figure 1). The average time between the baseline and 2-weeks measurement was 14.1 (SD: 3.7) days. Eleven patients participated by telephone. Sociodemographic and clinical characteristics of the participants at baseline and 2-weeks measurements are shown in Table 2. The baseline and 2-weeks study samples were comparable. About 60% was male, mean (SD) age was 65.5 (13.8) years and the majority (85%) had a Dutch ethno-cultural background. Mean (SD) eGFR was 21.4 (6.7) and 17% has had KRT in the past.

Content comparison

Table 3 shows the similarities and differences in characteristics of PROMIS CATs and SF-12. Although assessing the same patient-relevant outcome (i.e. generic HRQOL), PROMIS CATs and SF-12 include related but slightly different domains. The PROMs have similarities in scoring (e.g. score range and US reference), but use a different underlying measurement method and score interpretation. In PROMIS CATs, the (number of) items varies from person to person, depending on the severity of symptoms or the function level on the domain being measured and the consistency of the answers. Our study sample of advanced CKD patients completed a median (IQR) of 45 (38-55) items for all seven PROMIS CATs, which took them a median (IQR) of 10.2 (8.3-12.6) minutes. The median (IQR) time to complete the SF-12 was 3.3 (2.4-4.6) minutes.

Table 3. Content comparison of PROMIS CAT with SF-12^

	PROMIS CAT	SF-12
Type of PROM	Generic	Generic
PRO	HRQOL	HRQOL
Domains	Physical Function Pain Interference Fatigue Sleep Disturbance Anxiety Depression Ability to Participate in Social Roles and Activities	Physical functioning Bodily pain Vitality Role-physical Role-emotional Mental health Social functioning General health *Composite summary scores*:* Physical component summary Mental component summary
Number of items	All PROMIS domains[$] median (IQR): 45 (38-55) items Physical Function median (IQR): 4 (3-6) items Pain Interference median (IQR): 4 (2-12) items Fatigue median (IQR): 5 (4-6) items Sleep Disturbance median (IQR): 10 (8-12) items Anxiety median (IQR): 7 (6-10) items Depression median (IQR): 8 (5-12) items Ability to Participate in Social Roles and Activities median (IQR): 5 (4-6) items	All SF-12 items 12 items Physical functioning 2 items Bodily pain 1 item Vitality 1 item Role-physical 2 items Role-emotional 2 items Mental health 2 items Social functioning 1 item General health 1 item *Composite summary scores*:* Physical component summary 6 items Mental component summary 6 items
Recall period	In general/1 week	In general/4 weeks
Rating scale	5-point scale	3- and 5-point scales
Score (range)	Norm-based scoring T-score (roughly 0–100)	Norm-based scoring (roughly 0-100)
Norm or reference standard	General US population: mean 50, SD 10	General US population: mean 50, SD 10

Score interpretation	Higher scores represent more of the HRQOL domain being measured. E.g. a higher score on fatigue means a worse fatigue, and a higher score on physical function means a better physical function.	Higher scores represent a more favourable HRQOL. E.g. a higher score on bodily pain means less bodily pain, and a higher score on physical functioning means a better physical functioning.
Measurement method	Item Response Theory	Classical Test Theory
Completion options	Electronic only	Electronic or paper-based
Time to complete[#]	All PROMIS CATs median (IQR): 10.2 (8.3-12.6) min. Physical Function median (IQR): 1.3 (0.8-1.7) min. Pain Interference median (IQR): 1.2 (0.8-1.8) min. Fatigue median (IQR): 1.3 (1.0-2.0) min. Sleep Disturbance median (IQR): 2.0 (1.5-2.6) min. Anxiety median (IQR): 1.4 (1.0-1.9) min. Depression median (IQR): 1.3 (1.0-1.8) min. Ability to Participate in Social Roles and Activities median (IQR): 1.2 (1.0-1.6) min.	All SF-12 items median (IQR): 3.3 (2.4-4.6) min.

[^] The Dialysis Symptom Index (DSI) aims to measure a different patient-relevant outcome and is therefore not included in this table. For characteristics of the DSI, see Weisbord 2004[21] and Van der Willik 2021[36].

[*] SF-12 physical component summary includes the domains physical functioning, role-physical, bodily pain and general health; SF-12 mental component summary includes the domains vitality, social functioning, role-emotional and mental health.

[$] Number of items used as observed in current study sample at baseline. Additional item details, including the top 3 most frequently used items of PROMIS CATs are provided in Supplement B.

[#] Time to complete the PROMs as observed in current study sample at baseline.

Abbreviations: PROM, patient-reported outcome measure; PROMIS, Patient-Reported Outcomes Measurement Information System; CAT, Computerized Adaptive Test; SF-12, 12-item Short Form Health Survey; PRO, patient-reported outcome; IQR, interquartile range; SD, standard deviation.

Table 4 and Figure 2 show the PROM-scores in our study sample of patients with advanced CKD. Less variation (i.e. lower SDs) was observed in PROMIS CATs compared to SF-12 domains and summary scores. Overall, PROMIS CATs showed 'better' (towards the green area) HRQOL scores compared to the SF-12; only two PROMIS CATs showed worse HRQOL scores than the general US population (Physical Function [mean ± SD: 43.4±8.3] and Fatigue [53.2±8.7]), compared to six SF-12 domains and one summary score (physical functioning [40.5±11.3], role-physical [40.1±10.3], bodily pain [46.9±11.3], general health [36.3±10.9], social functioning [43.4±12.1], role-emotional [44.2±11.3] and physical component summary [39.2±10.7]).

Construct validity

All PROMIS CATs showed evidence for sufficient construct validity as ≥75% of the results were in accordance with the hypotheses (Table 5). For Pain Interference, Sleep Disturbance and Depression, all correlations were in accordance with the hypotheses. For Physical Function 14 out of 15 hypotheses were met. For Fatigue and Ability to Participate in Social Roles and Activities, 13 out of 15 correlations, and for Anxiety, 12 out of 15 correlations were in accordance with the hypotheses.

Table 4. Baseline scores on PROMIS CATs, SF-12, PROMIS Pain Intensity, and DSI in patients with chronic kidney disease (n=207)

	N (%)[$]	Mean (SD) or Median (IQR)	Range (min-max)
PROMIS CATs			
Physical Function	205 (99.0)	43.4 (8.3)	24.1 - 67.6
Pain Interference	203 (98.1)	51.9 (9.1)	41.0 - 74.9
Fatigue	203 (98.1)	53.2 (8.7)	28.8 - 70.7
Sleep Disturbance	203 (98.1)	49.3 (7.9)	30.0 - 71.6
Anxiety	203 (98.1)	51.2 (7.7)	35.9 - 70.3
Depression	204 (98.6)	49.8 (7.5)	37.1 - 70.0
Ability to Participate in Social Roles and Activities	203 (98.1)	49.2 (8.6)	29.9 - 64.9
SF-12			
Physical functioning	204 (98.6)	40.5 (11.3)	22.1 - 56.5
Role-physical	204 (98.6)	40.1 (10.3)	20.3 - 57.2
Bodily pain	204 (98.6)	46.9 (11.3)	16.7 - 57.4
General health	204 (98.6)	36.3 (10.9)	18.9 - 62.0
Vitality	204 (98.6)	48.5 (10.2)	27.6 - 67.9

Social functioning	204 (98.6)	43.4 (12.1)	16.2 - 56.6
Role-emotional	204 (98.6)	44.2 (11.3)	11.3 - 56.1
Mental health	204 (98.6)	50.1 (9.3)	28.0 - 64.5
Physical component summary*	204 (98.6)	39.2 (10.7)	11.1 - 61.4
Mental component summary*	204 (98.6)	49.3 (9.7)	23.4 - 69.0
PROMIS single item			
Pain Intensity (0-10)	204 (98.6)	1 (0-5)	0 - 10
DSI			
Number of symptoms (0-30)	203 (98.1)	9.4 (5.6)	0 - 28
Symptom burden score (0-150)	203 (98.1)	22 (12-36)	0 - 96
Feeling tired or lack of energy (0-5)^	203 (98.1)	2.0 (1.6)	0 - 5
Sleep problems (0-10)^#	203 (98.1)	2 (0-3)	0 - 10
Feeling anxious (0-5)^	203 (98.1)	0 (0-0)	0 - 5

$ In total, four people did not finish the measurement and only completed part of the PROMs.
** SF-12 physical component summary includes the domains physical functioning, role-physical, bodily pain and general health; SF-12 mental component summary includes the domains vitality, social functioning, role-emotional and mental health.*
^ Prevalence of feeling tired or lack of energy: 70.0%, sleep problems: 52.7%, feeling anxious: 18.7%.
 # Sleep problems were defined as trouble falling asleep and/or trouble staying asleep.
Abbreviations: PROMIS, Patient-Reported Outcomes Measurement Information System; CAT, Computerized Adaptive Test; SF-12, 12-item Short Form Health Survey; DSI, Dialysis Symptom Index; SD, standard deviation; IQR, interquartile range.

Test-retest reliability

The reliability measures – ICC agreement, SEM and MDC – of the PROMIS CATs, SF-12, PROMIS Pain Intensity single item, and DSI are shown in Table 6. All PROMIS CATs showed sufficient test-retest reliability (ICCs between 0.77 and 0.92). The SF-12 domains physical functioning, role-physical, bodily pain, general health, mental health, and the physical and mental component summary scores also showed sufficient reliability (ICCs between 0.70 and 0.85). For SF-12 role-emotional, social functioning and vitality the ICC was between 0.48 and 0.67. The PROMIS Pain Intensity single item showed an ICC of 0.68. The DSI total number of symptoms and symptom burden score showed sufficient reliability (ICC 0.85 and 0.88, respectively).

The SEM and MDC of PROMIS CATs ranged from 2.1 to 2.7, and from 5.7 to 7.4, respectively, across domains. For the SF-12, the SEM and MDC ranged from 4.1 to 7.8, and from 11.2 to 21.7, respectively, across domains.

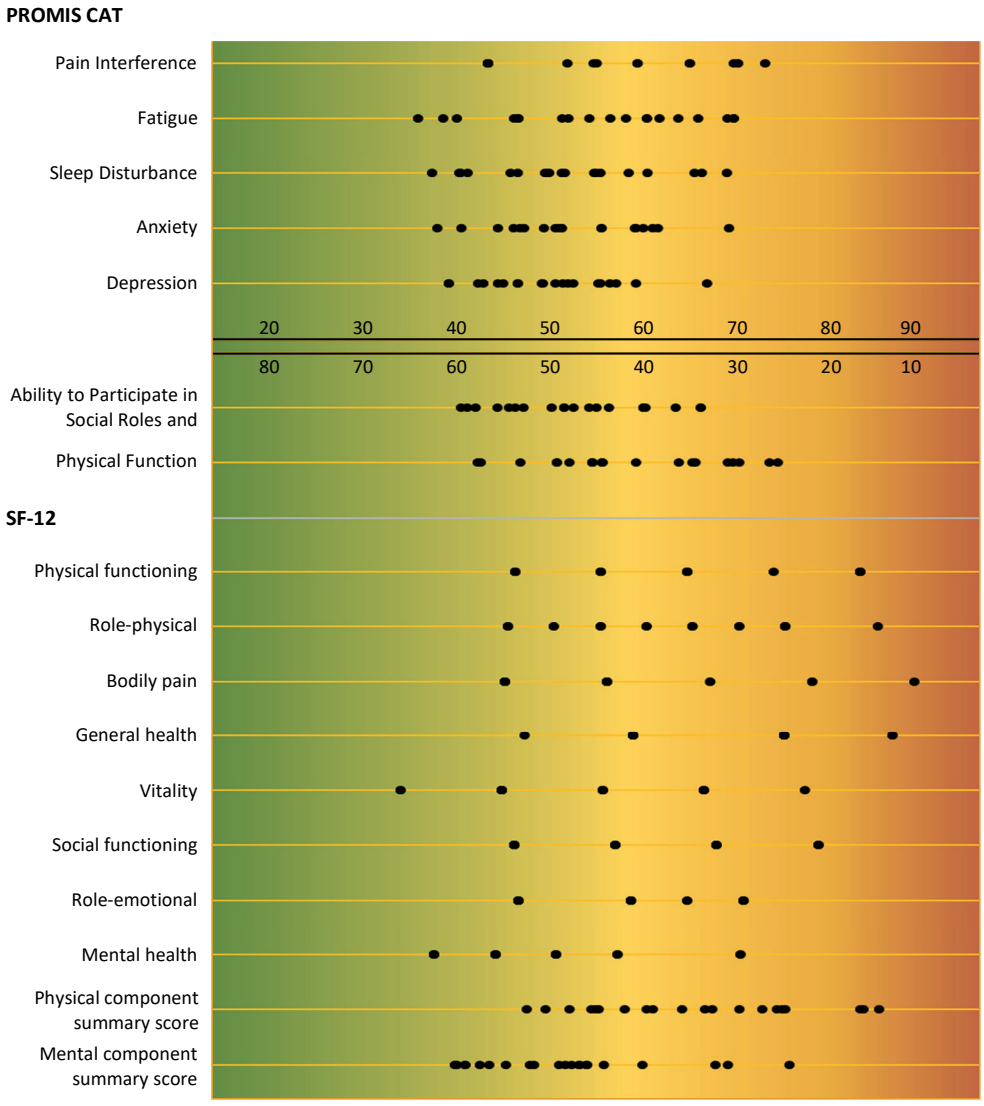


Figure 2. Score distributions of PROMIS CATs and SF-12 domains and summary scores.

The figure's background color gives an indication of the interpretation of scores, ranging from good (green) to worse (red) HRQOL.[31] *Note that the first five PROMIS CATs use a reverse scale compared to the other PROMIS CATs and SF-12.*

Table 5. Pearson's r for correlations between PROMIS CATs and SF-12 and DSI scores (n=207)

		PROMIS CATs						
		Physical Function	**Pain Interference**	**Fatigue**	**Sleep Disturbance**	**Anxiety**	**Depression**	**Ability to Participate**
SF-12	Physical functioning	**0.80**	*-0.52*	-0.49	-0.17	-0.19	-0.25	0.52
	Role-physical	0.65	-0.49	-0.59	-0.26	-0.24	-0.36	*0.59*
	Bodily pain	*0.59*	**-0.79**	-0.47	-0.35	-0.33	-0.33	0.47
	General health	*0.52*	-0.36	-0.53	-0.27	-0.24	-0.32	0.52
	Vitality	0.52	-0.39	**-0.66**	-0.31	-0.32	-0.43	0.59
	Social functioning	0.54	-0.49	-0.54	-0.34	-0.54	-0.58	**0.66**
	Role-emotional	0.30	-0.34	-0.41	-0.26	-0.40	-0.49	*0.39*
	Mental health	0.22	-0.33	-0.46	-0.33	**-0.66**	**-0.73**	0.40
	Physical component summary score*	**0.80**	*-0.63*	-0.55	-0.24	-0.13	-0.20	0.58
	Mental component summary score*	0.20	-0.29	-0.49	-0.35	**-0.64**	**-0.72**	0.47
DSI	Number of symptoms	-0.45	0.53	0.59	0.48	0.54	0.54	-0.48
	Symptom burden score	-0.48	0.55	0.60	0.51	0.51	0.52	-0.49
	Feeling tired or lack of energy	-0.41	0.41	**0.76**	0.35	0.36	0.46	-0.50
	Sleep problems#	-0.29	0.32	0.30	**0.79**	0.27	0.23	-0.27
	Feeling anxious	-0.04	0.16	0.16	0.25	**0.56**	0.48	-0.13
Hypotheses confirmed (%)		**93**	**100**	**87**	**100**	**80**	**100**	**87**

Correlations in **bold** *were expected to be strong (≥0.7 or ≤-0.7), correlations in italic were expected to be moderate (±0.5-0.7), other correlations were expected not to be strong (≤ 0.6 or ≥-0.6).*

** SF-12 physical component summary includes the domains physical functioning, role-physical, bodily pain and general health; SF-12 mental component summary includes the domains vitality, social functioning, role-emotional and mental health.*

 # DSI Sleep problems were defined as trouble falling asleep and/or trouble staying asleep.

Abbreviations: PROMIS, Patient-Reported Outcomes Measurement Information System; CAT, Computerized Adaptive Test; SF-12, 12-item Short Form Health Survey; DSI, Dialysis Symptom Index.

Discussion

This study examined the validity and reliability of seven PROMIS CATs in comparison to the SF-12 in patients with advanced CKD in The Netherlands. This is the first study investigating the psychometric performance of the Dutch-Flemish version of these PROMIS domains using CATs. All PROMIS CATs demonstrated evidence for sufficient construct validity and test-retest reliability. Overall, PROMIS CATs showed better reliability, with higher ICCs and lower MDCs, but required more items compared to the SF-12.

The observed average HRQOL scores are in line with scores that would be expected from existing literature in CKD patients for both the PROMIS CATs[18, 37, 38] and SF-12[2, 3, 39, 40]. However, comparison of the domain scores revealed a slightly better HRQOL in advanced CKD patients based on PROMIS CATs compared to SF-12. This demonstrates that the scores are not directly comparable in contrast to what one might intuitively expect based on the corresponding characteristics of both PROMs (0-100 scale, mean 50 with SD 10, US reference population). This can be explained by the fact that PROMIS CAT and SF-12 scores are on a different metric because they originate from different (calibration) samples[41], which is reflected in the smaller SDs for PROMIS CATs compared to SF-12. By means of linking[42], the scores of comparable PROMIS CAT and SF-12 domains could be converted into each other; this has been done for many other PROMs[43] and would be a valuable next step, as it facilitates harmonization of data across studies or healthcare organizations (e.g. when both instruments are used across different healthcare specialties) and comparison to historical data if one changes from one PROM to the other.[42]

All PROMIS CATs showed sufficient test-retest reliability with better ICCs and small MDCs compared to the SF-12. Small MDCs allow for small changes to be distinguished from measurement error with 95% confidence, and are therefore desirable especially when the minimal important change (MIC) is small.[36] For PROMIS, the MIC has been estimated at 2-6 points[26], which is slightly smaller but close to the MDC of 6-7 points. Information about the MIC for SF-12 domains is limited, which makes it difficult to say to what extent SF-12 can distinguish important changes from measurement error.[36] Our reliability results were better than results found in other research using PROMIS short forms (e.g. PROMIS-29 and -57, including 4 and 8 fixed items per domain, respectively).[18] This was expected given the underlying method of CAT and the stopping rule including a low SE to achieve high reliability. A downside of the higher precision stopping rule is the relatively large number of 45 questions asked (i.e. six to seven items per domain and three to four times the length of the SF-12). This number of items might raise some feasibility concerns for use in routine clinical practice. If fewer items are preferred, alternative stopping

Table 6. Reliability measures of PROMIS CAT, SF-12, PROMIS Pain Intensity, and DSI in patients with chronic kidney disease (n=179)

	ICC agreement (95%CI)	SEM	MDC
PROMIS CAT			
Physical Function	0.92 (0.89-0.94)	2.06	5.72
Pain Interference	0.78 (0.71-0.83)	2.65	7.43
Fatigue	0.81 (0.75-0.86)	2.06	5.71
Sleep Disturbance	0.84 (0.79-0.88)	2.22	6.15
Anxiety	0.78 (0.71-0.83)	2.29	6.36
Depression	0.81 (0.76-0.86)	2.35	6.53
Ability to Participate in Social Roles and Activities	0.77 (0.71-0.83)	2.09	5.80
SF-12			
Physical functioning	0.76 (0.69-0.82)	5.27	14.61
Role-physical	0.73 (0.65-0.79)	5.10	14.13
Bodily pain	0.70 (0.62-0.77)	6.02	16.67
General health	0.75 (0.68-0.81)	5.23	14.50
Vitality	0.67 (0.58-0.75)	5.72	15.85
Social functioning	0.64 (0.54-0.72)	7.20	19.96
Role-emotional	0.48 (0.36-0.58)	7.82	21.67
Mental health	0.78 (0.82-0.83)	4.32	11.98
Physical component summary score*	0.85 (0.81-0.89)	4.07	11.29
Mental component summary score*	0.72 (0.65-0.79)	5.09	14.11
PROMIS single item			
Pain Intensity (0-10)	0.68 (0.59-0.76)	1.53	4.24
DSI			
Number of symptoms (0-30)	0.85 (0.80-0.88)	2.12	5.87
Total symptom burden score (0-150)	0.88 (0.85-0.91)	5.75	15.94

* *SF-12 physical component summary includes the domains physical functioning, role-physical, bodily pain and general health; SF-12 mental component summary includes the domains vitality, social functioning, role-emotional and mental health.*

Abbreviations: PROMIS, Patient-Reported Outcomes Measurement Information System; CAT, Computerized Adaptive Test; SF-12, 12-item Short Form Health Survey; DSI, Dialysis Symptom Index; ICC, intra-class correlation coefficient; CI, confidence interval; SEM, standard error of measurement; MDC, minimal detectable change

rules could be considered but with detriment to precision. In this study, we applied a stopping rule with a smaller SE of 2.2 compared to the standard stopping rule (SE: 3.0)[13] to investigate optimal performance of the PROMIS CATs. We expect that application of the standard stopping rule will result in 36-43 items in total (5-6 items per CAT), with a minimum of 28 items since the standard stopping rule requires 4 items per CAT, and less than 45 items due to the higher SE in comparison to this study. Other alternative stopping rules to consider might be a lower maximum number of items (a maximum of 8 instead of 12 items per domain is currently being considered for the standard PROMIS CAT algorithms), stopping when the SE does not change much anymore (e.g. <0.1), or stopping when the score range is above or below a certain cut-off point on the scale (e.g. when the functionality or symptom burden is at such a level that it is probably not perceived as burdensome). The latter may be particularly beneficial for domains such as Pain Interference and Sleep Disturbance, to keep the number of items low for patients with no pain or sleep problems. Further research is needed to explore feasibility and the most optimal use of PROMIS CATs in routine nephrology care, in close collaboration with patients and healthcare professionals.

A limitation of PROMIS CATs is that they can only be completed digitally. Participants thus have to have access to an electronic device and be digitally skilled. In the Netherlands, approximately 80% of the population aged 55+ is sufficiently digitally skilled[44], but in many countries – also within Europe – citizens are less digitally skilled.[45, 46] Consequently, it may be challenging to reach the total advanced CKD population. In our study, we therefore enabled participation by telephone. For routine care, also other methods could be considered, for instance offering help or making tablets available on site.

An advantage of PROMIS CATs is that the PROM adapts to the patient, resulting in items that are more likely considered relevant by the patient. As a result, the PROM might be perceived as less burdensome. On the other hand, items may vary over time, meaning that progression of individual items cannot be easily monitored over time, which is in contrast to how the SF-12 (and DSI) is also being used in routine nephrology care.[3] In addition, the varying items and 'black box algorithm' (i.e. not a simple sum of scores) may also lead to patients and professionals finding it more difficult to interpret the scores. Qualitative research is needed to investigate patients' and professionals' preferences for use in routine nephrology care.

Furthermore, it may be important to mention that the SF-12 was selected by patients for use in routine nephrology care, partly because of the low number of items. Besides, the SF-12 was considered a good fit with the DSI to provide insight into both generic HRQOL and disease-specific symptom burden.[3, 10] Differences

in characteristics of the PROMIS CATs and SF-12, and how they complement other PRO(M)s, should thus be taken into account when considering which PROM fits routine nephrology care best to measure HRQOL.

An important strength of this study is that the PROMIS CATs were compared to the PROM that is currently being used in routine nephrology care to assess generic HRQOL (i.e. the SF-12). The findings from this study are therefore of clinical relevance and can contribute to considerations regarding which PROMs best fit routine practice to measure HRQOL. A disadvantage is that the SF-12 may not be the best comparator (i.e. 'golden standard') for the PROMIS CATs, for instance because of the low number of items per domain and the fact that in practice, both in research and in healthcare, less focus is often being paid towards individual SF-12 domains. To expand on current findings, future research could investigate the validity of PROMIS CATs in comparison to the SF-36.[11, 29]

Conclusion

All seven PROMIS CATs (assessing physical function, pain interference, fatigue, sleep disturbance, anxiety, depression, and ability to participate in social roles and activities) demonstrated evidence for sufficient construct validity and test-retest reliability in patients with advanced CKD in The Netherlands. PROMIS CATs required more items but showed better reliability than the SF-12. Future research is needed to investigate the optimal use of PROMIS CATs for routine nephrology care.

References

1. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

2. Voskamp PWM, van Diepen M, Evans M, et al. The impact of symptoms on health-related quality of life in elderly pre-dialysis patients: effect and importance in the EQUAL study. *Nephrol Dial Transplant* 2019;34(10):1707-15

3. van der Willik EM, Hemmelder MH, Bart HAJ, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. *Clin Kidney J* 2021;14(6):1535-44

4. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

5. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. *Clin J Am Soc Nephrol* 2014;9(10):1813-21

6. Urquhart-Secord R, Craig JC, Hemmelgarn B, et al. Patient and Caregiver Priorities for Outcomes in Hemodialysis: An International Nominal Group Technique Study. *Am J Kidney Dis* 2016;68(3):444-54

7. Weisbord SD, Fried LF, Mor MK, et al. Renal provider recognition of symptoms in patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2007;2(5):960-7

8. Wang Y, Snoep JD, Hemmelder MH, et al. Outcomes after kidney transplantation, let's focus on the patients' perspectives. *Clin Kidney J* 2021;14(6):1504-13

9. 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

10. van der Willik EM, Meuleman Y, Prantl K, et al. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease - a four-phase mixed methods study. *BMC Nephrol* 2019;20(1):344

11. Verberne WR, Das-Gupta Z, Allegretti AS, et al. Development of an International Standard Set of Value-Based Outcome Measures for Patients With Chronic Kidney Disease: A Report of the International Consortium for Health Outcomes Measurement (ICHOM) CKD Working Group. *American Journal of Kidney Diseases* 2019;73(3):372-84

12. Terwee CB, Vonkeman HE, Zuidgeest M: Het menu van generieke PROMs Advies, The Netherlands, Linnean initiatief, werkgroep PROMs en PREMs, 2019

13. HealthMeasures. Intro to PROMIS®, Northwestern University, 2021. https://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis

14. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology* 2010;63(11):1179-94

15. van Balen EC, Haverman L, Hassan S, et al. Validation of PROMIS Profile-29 in adults with hemophilia in the Netherlands. *J Thromb Haemost* 2021;19(11):2687-701

16. Rose AJ, Bayliss E, Huang W, et al. Evaluating the PROMIS-29 v2.0 for use among

older adults with multiple chronic conditions. *Quality of Life Research* 2018;27(11):2935-44

17. Katz P, Pedro S, Michaud K. Performance of the Patient-Reported Outcomes Measurement Information System 29-Item Profile in Rheumatoid Arthritis, Osteoarthritis, Fibromyalgia, and Systemic Lupus Erythematosus. *Arthritis Care & Research* 2017;69(9):1312-21

18. Tang E, Ekundayo O, Peipert JD, et al. Validation of the Patient-Reported Outcomes Measurement Information System (PROMIS)-57 and -29 item short forms among kidney transplant recipients. *Quality of Life Research* 2019;28(3):815-27

19. Selewski DT, Massengill SF, Troost JP, et al. Gaining the Patient Reported Outcomes Measurement Information System (PROMIS) perspective in chronic kidney disease: a Midwest Pediatric Nephrology Consortium study. *Pediatr Nephrol* 2014;29(12):2347-56

20. ERA-EDTA Registry Annual Report 2019, Amsterdam UMC, location AMC, Department of Medical Informatics, Amsterdam, the Netherlands, European Renal Association, 2021

21. Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *J Pain Symptom Manage* 2004;27(3):226-40

22. 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

23. Loosman WL, Hoekstra T, van Dijk S, et al. Short-Form 12 or Short-Form 36 to measure quality-of-life changes in dialysis patients? *Nephrol Dial Transplant* 2015;30(7):1170-6

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

25. Terwee CB, Roorda LD, de Vet HC, et al. Dutch-Flemish translation of 17 item banks from the patient-reported outcomes measurement information system (PROMIS). *Qual Life Res* 2014;23(6):1733-41

26. Terwee CB, Peipert JD, Chapman R, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. *Quality of Life Research* 2021;30(10):2729-54

27. Dutch-Flemisch PROMIS. http://www.dutchflemishpromis.nl/

28. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-33

29. Ware JE, Jr. SF-36 health survey update. *Spine (Phila Pa 1976)* 2000;25(24):3130-9

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

31. HealthMeasures. PROMIS® Score Cut Points, Northwestern University, 2021. https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/promis-score-cut-points

32. van Muilekom MM, Luijten MAJ, van Oers HA, et al. From statistics to clinics: the visual feedback of PROMIS® CATs. *J Patient Rep Outcomes* 2021;5(1):55

33. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge university press: 2011;

34. Jabrayilov R, Emons WHM, Sijtsma K. Comparison of Classical Test Theory and Item Response Theory in Individual Change Assessment. *Appl Psychol Meas* 2016;40(8):559-72

35. Hays RD, Spritzer KL, Reise SP. Using Item Response Theory to Identify Responders to Treatment: Examples with the Patient-Reported Outcomes Measurement Information System (PROMIS®) Physical Function Scale and Emotional Distress Composite. *Psychometrika* 2021;86(3):781-92

36. van der Willik EM, Terwee CB, Bos WJW, et al. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. *Nephrology (Carlton)* 2021;26(5):391-9

37. Vanden Wyngaert K, Van Craenenbroeck AH, Eloot S, et al. Associations between the measures of physical function, risk of falls and the quality of life in haemodialysis patients: a cross-sectional study. *BMC Nephrol* 2020;21(1):7

38. Sturgill DA, Bal N, Nagavally S, Wolfgram DF. The Relationship between Dialysis Metrics and Patient-Reported Cognition, Fatigue, and Physical Function. *Kidney Dis (Basel)* 2020;6(5):364-70

39. 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

40. Erez G, Selman L, Murtagh FE. Measuring health-related quality of life in patients with conservatively managed stage 5 chronic kidney disease: limitations of the Medical Outcomes Study Short Form 36: SF-36. *Qual Life Res* 2016;25(11):2799-809

41. Terwee CB, Crins MHP, Roorda LD, et al. International application of PROMIS computerized adaptive tests: US versus country-specific item parameters can be consequential for individual patient scores. *J Clin Epidemiol* 2021;134:1-13

42. Schalet BD, Lim S, Cella D, Choi SW. Linking Scores with Patient-Reported Health Outcome Instruments: a Validation Study and Comparison of Three Linking Methods. *Psychometrika* 2021;86(3):717-46

43. PROsetta Stone® Linking Patient-Reported Outcome Measures. https://www.prosettastone.org/

44. van Deursen AJAM: Digitale ongelijkheid in Nederland: Internetgebruik van mensen van 55 jaar en ouder., Enschede, Nederland, Universiteit Twente, 2019

45. James J. Confronting the scarcity of digital skills among the poor in developing countries. *Development Policy Review* 2021;39(2):324-39

46. Eurostat. Individuals' level of digital skills (from 2021 onwards) 2022. https://ec.europa.eu/eurostat/databrowser/view/ISOC_SK_DSKL_I21__custom_2397093/bookmark/table?lang=en&bookmarkId=dc481686-c938-4e07-b03c-8e039f532857. Accessed June 10, 2022

Supplemental Material for Chapter 3

Supplement A – Visual illustration of Computerized Adaptive Testing (CAT)

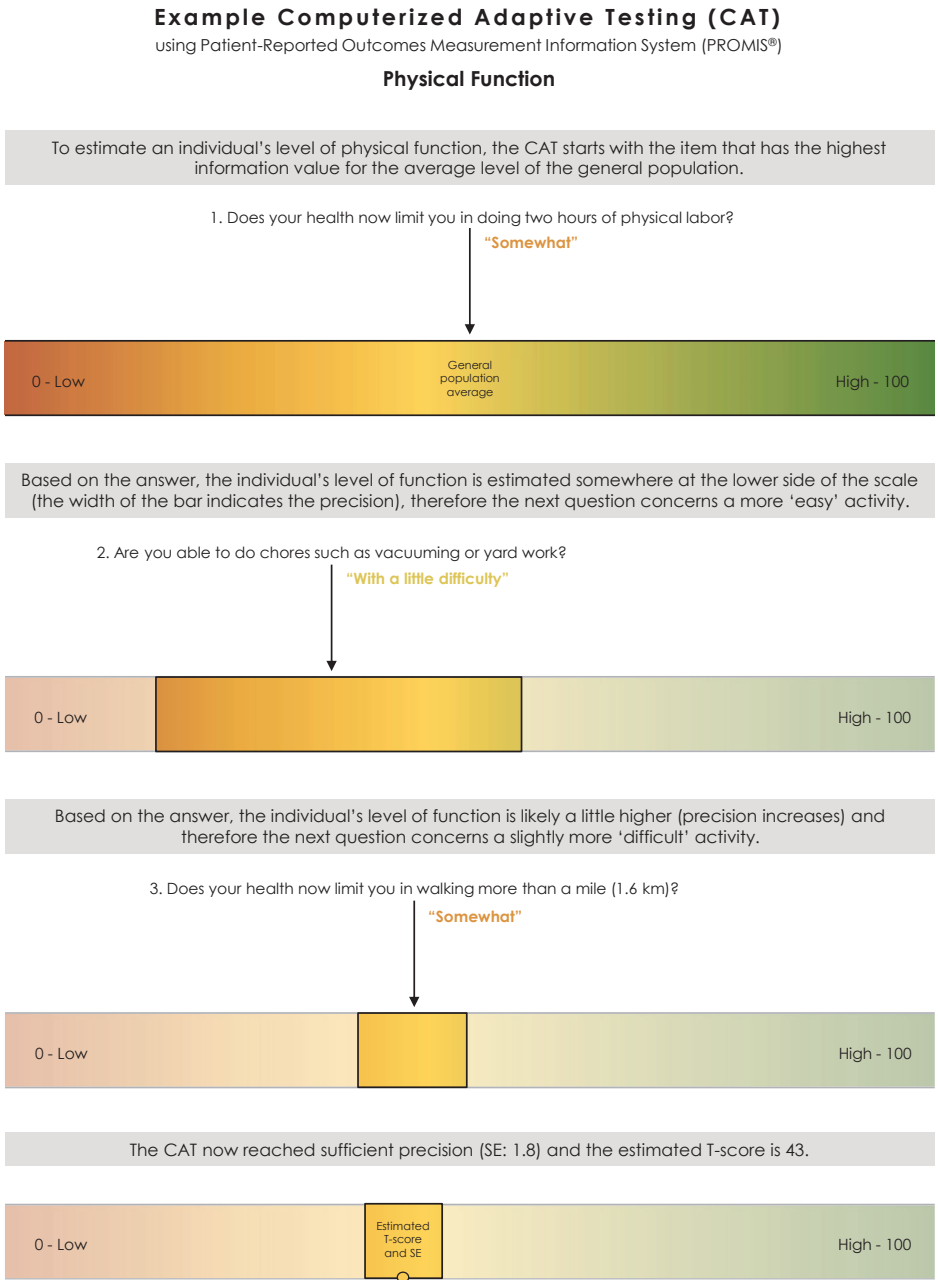


Figure S1. Visual illustration of Computerized Adaptive Testing (CAT) using Patient-Reported Outcomes Measurement Information System (PROMIS®) Physical Function

Supplement B – PROMIS CAT item characteristics

Table S1. Characteristics of items used in PROMIS CATs.

PROMIS CAT	Items used / total item bank	Top 3 items used	Items per patient
Physical Function	28 / 121	1. Does your health now limit you in doing two hours of physical labor? (n=205) 2. Are you able to do chores such as vacuuming or yard work? (n=122) 3. Does your health now limit you in walking more than a mile (1.6 km)? (n=119)	Median (IQR): 4 (3-6) Min-max: 3-12 12 items: n=8 (3.9%)
Pain Interference	24 / 40	1. How much did pain interfere with your day to day activities? (n=203) 2. How much did pain interfere with your ability to participate in social activities? (n=109) 3. How often was pain distressing to you? (n=95)	Median (IQR): 4 (2-12) Min-max: 2-12 12 items: n=75 (36.9%)
Fatigue	27 / 95	1. How often did you have to push yourself to get things done because of your fatigue? (n=203) 2. I have trouble starting things because I am tired. (n=180) 3. How exhausted were you on average? (n=33)	Median (IQR): 5 (4-6) Min-max: 4-12 12 items: n=6 (3.0%)
Sleep Disturbance	22 / 27	1. My sleep quality was… (n=203) 2. I had trouble sleeping. (n=203) 3. I had a problem with my sleep. (n=201)	Median (IQR): 10 (8-12) Min-max: 6-12 12 items: n=71 (35.0%)
Anxiety	21 / 29	1. I felt uneasy... (n=203) 2. I felt tense… (n=157) 3. I felt anxious and worried… (n=142)	Median (IQR): 7 (6-10) Min-max: 5-12 12 items: n=44 (21.7%)
Depression	21 / 28	1. I felt depressed. (n=204) 2. I felt unhappy. (n=144) 3. I felt discouraged about the future. (n=130)	Median (IQR): 8 (5-12) Min-max: 3-12 12 items: n=55 (27.0%)
Ability to Participate in Social Roles and Activities	23 / 35	1. I have trouble doing all of my regular leisure activities with others … (n=203) 2. I have trouble doing all of the activities with friends that I want to do... (n=132) 3. I have trouble doing all of the family activities that I want to do… (n=114)	Median (IQR): 5 (4-6) Min-max: 3-12 12 items: n=13 (6.4%)


4

Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures (PROMs)

Esmee M. van der Willik, Marc H. Hemmelder, Hans A. J. Bart, Frans J. van Ittersum, Judith M. Hoogendijk - van den Akker, Willem Jan W. Bos, Friedo W. Dekker, Yvette Meuleman

Clinical Kidney Journal. 2020;14(6):1535-1544.

Abstract

Background: The use of patient-reported outcome measures (PROMs) is becoming increasingly important in healthcare. However, incorporation of PROMs into routine nephrological care is challenging. This study describes the first experience with PROMs in Dutch routine dialysis care.

Methods: A pilot study was conducted in dialysis patients in 16 centres. Patients were invited to complete PROMs at baseline, 3 and 6 months. PROMs consisted of the 12-item Short-Form (SF-12) and Dialysis Symptom Index (DSI) to assess health-related quality of life (HRQOL) and symptom burden. Response rates, HRQOL and symptom burden scores were analysed. Qualitative research methods were used to gain insight into patients' view on using PROMs in clinical practice.

Results: In total, 512 patients (36%) completed 908 PROMs (24%) across three time-points. Response rates varied from 6-70% among centres. Mean (SD) scores for physical and mental HRQOL were 35.6 (10.2) and 47.7 (10.6), respectively. Patients experienced on average 10.8 (6.1) symptoms with a symptom burden score of 30.7 (22.0). Only 1-3% of the variation in PROM-scores can be explained by differences between centres. Patients perceived discussing their HRQOL and symptom scores as insightful and valuable. Individual feedback on results was considered crucial.

Conclusions: First results show low average response rates with high variability among centres. Dialysis patients experienced a high symptom burden and poor HRQOL. Using PROMs at individual patient level is suitable and may improve patient-professional communication and shared decision-making. Further research is needed to investigate how collection and use of PROMs can be successfully integrated into routine care in order to improve healthcare quality and outcomes.

Introduction

Patients with advanced chronic kidney disease (CKD) experience numerous physical and emotional disease-related symptoms and a poor health-related quality of life (HRQOL).[1-3] In daily healthcare, these patient-reported outcomes (PROs) are frequently underrecognized and underestimated [2, 4], and consequently, may remain unattended.[5] The under-identification may be partly explained by patients not sharing their symptoms and needs easily [6, 7], and by difficulties for clinicians to identify the full spectrum and severity of patients' symptoms and needs.[4, 7, 8]

The use of patient-reported outcome measures (PROMs) may facilitate communication about symptoms and needs, and may provide insight into PROs both at individual and at centre or national level.[9-13] Although the importance of PROs is recognized [14-16] and the use of PROMs in routine care is highly supported [9, 16], PROMs are often not yet part of standard nephrological care.[9, 13, 17] In Europe, few renal registries have initiated routine collection of PROMs [18-20]. The Scottish Renal Registry recently described their first experience with collecting PROMs and encountered challenges including a low response rate, selective response, organizational struggles and low commitment from centres.[18] Literature also corroborates that it is challenging to incorporate PROMs into routine care.[9-11, 17, 20-22] A major challenge is to incorporate PROMs in such a way that it can be used for different purposes at different levels; to evaluate healthcare quality at aggregated level, and, perhaps even more important for patients, to support patient-professional communication and decision-making at individual patient level.[9-11, 13] Using PROMs for different purposes requires engagement at all levels, high response rates and feedback on outcomes tailored to the context and the purpose.[9-13]

Currently, PROMs are being implemented into Dutch nephrological care to provide insight into PROs of individual patients and at centre and national level. PROMs will be collected in Renine, the Dutch Renal Registry (www.nefrovisie.nl/renine) in which all patients on renal replacement therapy (RRT) are registered. This study describes the first experience with PROMs in Dutch routine dialysis care. We aim to evaluate the introduction of the national registry of PROMs by answering the following research questions:

1. What is the response rate and how does the response rate vary among centres? Which differences in characteristics are observed between responders and non-responders?

2. What is the HRQOL and symptom burden of patients receiving dialysis, which variation in scores is observed among centres, and to which extent can variation in scores be explained by differences in patient population?

3. What are patients' experiences and views on the use of PROMs in clinical practice?

Materials and Methods

Study design and patients

The registry of PROMs was introduced in routine nephrological care through a pilot study among adult patients on dialysis in 16 Dutch centres from September 2016 to April 2017. These centres treat 26% of all Dutch patients receiving dialysis. Patients undergoing any type of dialysis were included. Clinicians invited their patients to complete the online PROMs at 3 time-points: at baseline, 3 and 6 months after study start. This frequency was considered suitable by patients [23] and is expected to be sufficient for centres to become familiar with PROMs. Aiming at optimal incorporation of PROMs in routine care, centres were free to develop the process of inviting and motivating patients that fits their workflow.[24, 25] Clinicians could decide not to invite a patient, for example because of the patient's holiday or medical condition. At 6 months, PROMs were available to complete in the following languages: Dutch, English, Turkish and Arabic.

PROMs: HRQOL and symptom burden questionnaires

The PROMs consist of two questionnaires: the 12-item Short-Form (SF-12) health survey to assess HRQOL and the Dialysis Symptom Index (DSI) to assess symptom burden. These questionnaires were carefully selected in close collaboration with patients, professionals, the Dutch Kidney Patients Association (NVN) and the quality institution Nefrovisie.[23] Literature also recommends the SF-12 as appropriate questionnaire to assess HRQOL in routine care, but no recommendation is provided for the assessment of symptom burden.[20] Therefore, a four-phase mixed methods study was conducted to select the best suitable symptom questionnaire, in collaboration with patients and experts, and by using existing symptom questionnaires and literature. In this study, the DSI was found the most relevant, complete and comprehensible symptom questionnaire for routine assessment in patients with advanced CKD. The details of this selection process have been described elsewhere.[23]

The SF-12 is a generic questionnaire consisting of 12 questions regarding physical and mental HRQOL, especially developed for large-scale monitoring.[26] Within the dialysis population, the SF-12 is frequently used and has shown to be a preferred and valid questionnaire.[20, 27] Norm-based scoring algorithms were used to calculate physical (PCS) and mental (MCS) component scores. Component scores range from 0 to 100, with higher scores indicating better HRQOL. PCS and MCS scales are standardized to the U.S. population with a mean of 50 and a standard deviation (SD) of 10.[26, 28]

The DSI is a 30-item disease-specific symptom questionnaire to assess physical and emotional symptom burden.[29] To ensure comprehensiveness for individual

patients, an open-ended question was added to report three additional symptoms.[23] Patients indicate for each symptom if it was present (yes/no) during the past week and, if so, how much it bothered (5-point scale ranging from 1 'not at all' to 5 'very much'). The number (0-30 symptoms) and burden (score ranging from 0 [no symptoms] to 150 [all 30 symptoms are present and are very burdensome]) of symptoms were calculated, with higher scores indicating higher symptom burden.[30] Scores were calculated for patients that filled in ≥28 questions, whereby missing symptoms were assumed absent (burden score 0).[30]

Potentially explanatory factors

From Renine we obtained patient, disease and treatment characteristics describing the study population: age, sex, primary kidney disease (according to European Renal Association – European Dialysis and Transplantation Association codes [31]), social economic status (SES; using zip code [32]), dialysis modality and time on RRT (using date of RRT initiation).

Statistical analysis

Statistical analyses were performed using SPSS version 23.0. *P*-values<0.05 were considered statistically significant. Variables are shown as mean (SD) or percentages. Non-normally distributed variables were log-transformed and presented as geometric mean (SD). Missing values for patient, disease and treatment characteristics (Table 1) were assumed 'missing at random' and estimated using multiple imputation.[33, 34] Ten imputed datasets were created.[34] The imputation model included all patient, disease and treatment characteristics (see potentially explanatory factors), centre, response, if patients received support completing PROMs, death during follow-up, cause of death and all outcomes (PCS, MCS, symptom number and burden score).[34, 35]

Patients who died or for which the centre indicated that they did not invite the patient were excluded from analyses for relevant time-points (Figure 1). Patients were considered a responder if they participated at least once. Student's t-test and Chi-squared tests were used to compare characteristics of responders and non-responders. To compare response rates between time-points and centre-volume (number of dialysis patients), Chi-squared test and linear regression analysis were performed, respectively. For patients who participated at multiple time-points, their first measurement was used in HRQOL and symptoms analyses (Figure 1).

MCS, PCS, symptom number and burden scores were calculated for responders who completed the full questionnaire (Figure 1). To explore variation among centres, MCS, PCS and symptom burden scores were assessed per centre and

compared to the overall study population through indirect standardization. To that end, the following steps were taken: first, an expected score was calculated per patient for each outcome separately by using a multivariable linear regression model including patient, disease and treatment characteristics as predictors. Second, mean expected and observed scores were calculated for each centre. Third, adjusted PROM-scores per centre were calculated as follows: *(Observed centre mean – Expected centre mean) + overall mean*, hereby creating centre scores comparable to the overall study population. Crude and adjusted PROM-scores are shown in funnel plots.[36] Funnel plots were created in R version 3.4.2.

To examine to what extent variation among centres can be explained by differences in patient population, intraclass correlation (ICC) – also referred to as 'rankability' – was calculated using multilevel regression analysis (MLRA).[37-39] ICC is the proportion of variance in MCS, PCS or symptom burden scores that occurs at centre level. This variance may be attributable to centre-factors or to the patient population (of responders) of centres.[37, 38] Patient, disease and treatment characteristics were included as fixed effects and centres as random effect in the MLRA model. Comparison of ICC before and after including characteristics in the model (i.e. comparing crude and adjusted ICC) shows to which extent centre variation can be explained by differences in characteristics of centre populations.[38]

Patients' experiences and preferences

As PROMs are intended to become part of regular care, we wanted to know more about patients' experiences with and preferences for discussing PROM-scores with their healthcare professional. At 3 months, all patients were asked if they would like to share and discuss their PROM-scores with their healthcare professional. Hereafter, in each centre, professionals received an individual digital report from 5 randomly selected patients who gave consent. Professionals were invited to discuss the report with these patients at their next consultation visit. This report contained patient's responses and PROM-scores with a comparison to all responders and – for MCS and PCS – the general Dutch population. At 6 months, patients and professionals were asked how they experienced the conversation about PROM-scores. Patients also reported which professional discussed the PROM-scores with them and how satisfied they were with the conversation (5-point scale: 1-5, poor-excellent).

Additionally, the use of PROMs was evaluated in a focus group with patients receiving dialysis. Patients were recruited by NVN via e-mail and social media. During the focus group, patients' views and preferences concerning the use of PROMs in clinical practice were discussed. The focus group lasted 2.5 hours and was chaired

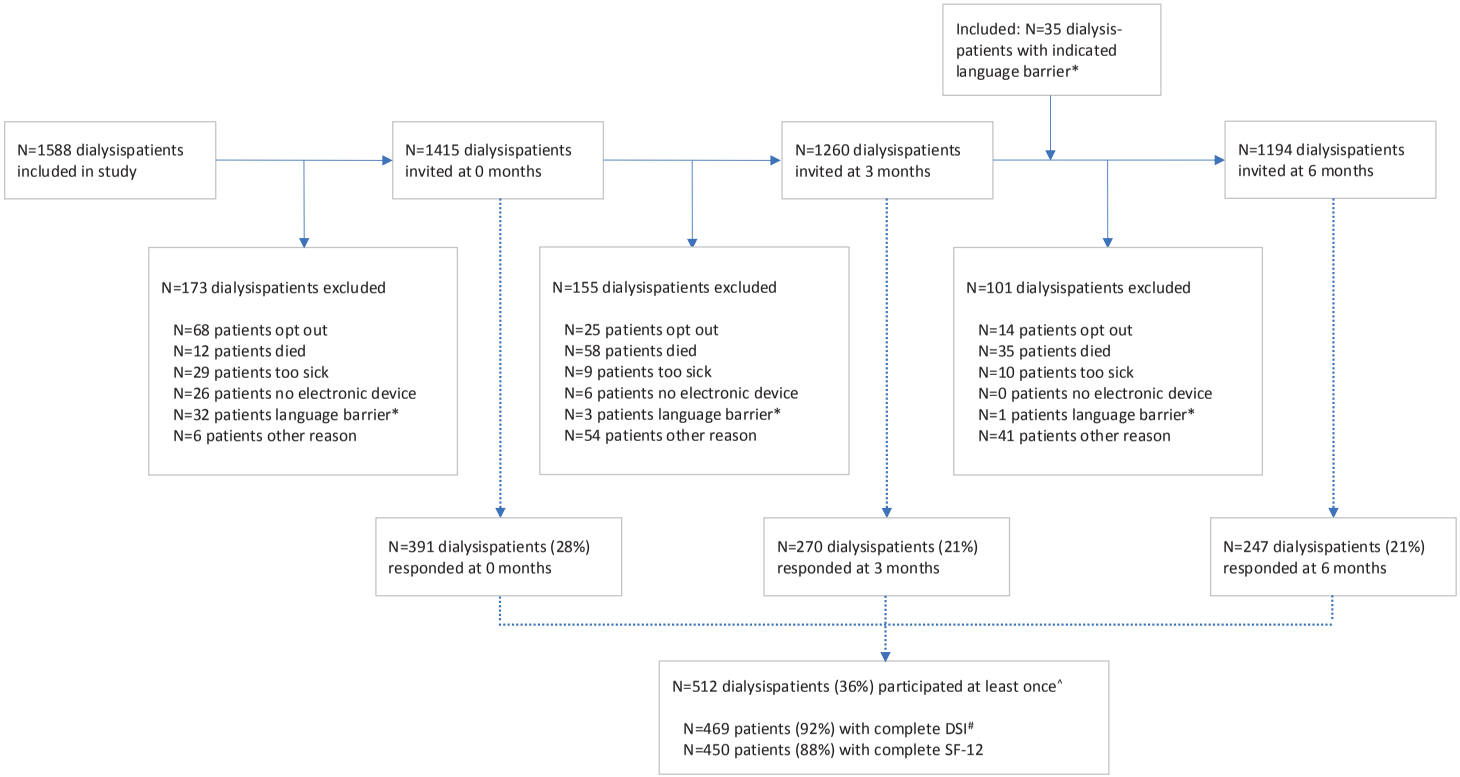


Figure 1. Flow chart for the number of patients included, invited and participating at each time-point.

* *Patients that were excluded because of a language barrier at 0 or 3 months were again included at 6 months: PROMs were also available in English, Turkish and Arabic at 6 months.*

^ *In total, 1440 patients were invited at least at one time-point.*

\# *The DSI was considered complete if ≥28 questions had been answered.*

Abbreviations: DSI, Dialysis Symptom Index; SF-12, 12-item Short-Form; PROMs, patient-reported outcome measures.

by a health educator (KP). Patients' discussion was recorded in detail by handwritten field notes and, when possible, verbatim by the chair and two observers (HB and FD). All written information was analysed using Atlas.ti. Statements were analysed by a researcher trained in qualitative research (EvdW) and discussed with an experienced qualitative researcher (YM).

Results

Response rate

Figure 1 shows the number of patients that were invited and responded across the time-points. In total, 1440 patients were invited at least once. The main reasons not to invite a patient were the medical condition of patients and that patients indicated that they did not want to be invited. In total, 512 patients (36%) responded at least once and altogether completed 908 PROMs (24%) across the three time-points.

The response rate was higher at baseline with 28% compared to 21% at 3 and 6 months ($p<0.001$). Figure 2 presents response rates per centre at all time-points. A large variation among centres was found with response rates ranging from 6% to 70%.

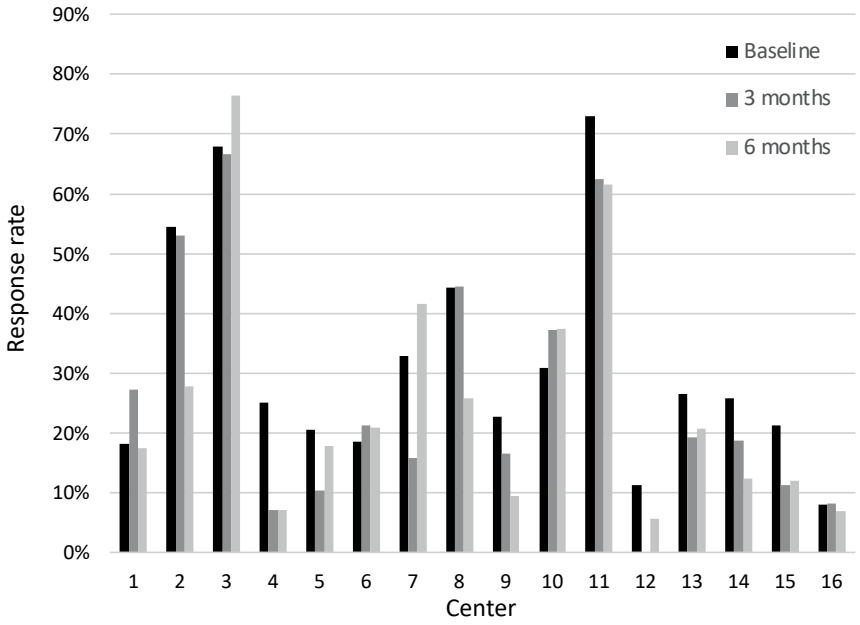


Figure 2. Response rates per time-point in 16 pilot centres.

Centres are ranked (low to high) according to the number of patients on dialysis included at baseline. Larger centres (i.e. higher number of patients included at baseline) had a slightly lower response rate compared to smaller centres: the response rate decreases with 2% for each additional 10 patients (p<0.001).

Table 1. Characteristics of responders and non-responders (N=1440)

	Responder* N=512	Non-responder* N=928	*p*-value
Sex, male [a]	342 (67.9)	484 (57.1)	<0.001
Age, mean (years) [b]	66.6 (13.8)	64.7 (16.0)	0.022
SES [c]			<0.001
Low	119 (24.1)	305 (36.5)	
Middle	309 (62.6)	430 (51.5)	
High	66 (13.4)	100 (12.0)	
Primary kidney disease [d]			0.005
Glomerulonephritis/sclerosis	55 (12.5)	98 (12.8)	
Pyelonephritis	23 (5.2)	40 (5.2)	
Polycystic kidney disease	42 (9.5)	40 (5.2)	
Hypertension	72 (16.4)	158 (20.6)	
Renal vascular disease	67 (15.2)	96 (12.5)	
Diabetes mellitus	84 (19.1)	194 (25.3)	
Miscellaneous	97 (22.0)	142 (18.5)	
Dialysis modality [e]			0.121
HD centre	407 (82.6)	695 (82.3)	
HD home	18 (3.7)	50 (5.9)	
PD	68 (13.8)	99 (11.7)	
Time on RRT, mean (years) [f #]	2.5 (3.8)	3.1 (3.4)	0.005

Values are shown in N (%) or mean (SD)
**Patients are considered responder if they participated at least once. Non-responders are invited at least once, but never participated.*
[a] Sex is available for 504 (98.4%) responders and 847 (91.3%) non-responders, [b] Age is available for 504 (98.4%) responders and 846 (91.2%) non-responders, [c] SES is available for 494 (96.5%) responders and 835 (90.0%) non-responders, [d] Primary kidney disease is available for 440 (85.9%) responders and 768 (82.8%) non-responders, [e] Dialysis modality is available for 493 (96.3%) responders and 844 (90.9%) non-responders, [f] Time on RRT is available for 497 (97.1%) responders and 847 (91.3%) non-responders.
[#] Time on RRT is shown as geometric mean (SD).
Abbreviations: SES, social economic status; HD, haemodialysis; PD, peritoneal dialysis; RRT, renal replacement therapy.

Responders

Table 1 shows the characteristics of responders (N=512) compared to non-responders (N=928). Responders were more frequently male, older, had a higher SES and started RRT more recently. Responders' primary kidney disease was more frequently polycystic kidney disease and less frequently hypertension or diabetes. Responders needed on average 12.2 (SD: 6.1) minutes to complete the PROMs. In total, 211 out of 512 patients (41%) received some support to complete the PROMs, ranging from 7% to 65% across centres. When support was provided, the support mainly consisted of: reading questions aloud (81%), filling in patients' answers (79%), translating of questions (6%), and completing the questionnaire on their behalf (e.g. their partner; 8%). Eleven patients (5%) indicated that other support was provided, such as assistance in using an electronic device or discussing questions with relatives to remember their experiences. Furthermore, some centres with high response rates indicated that they provided tablets, so that patients could complete the PROMs while receiving dialysis treatment. The non-Dutch questionnaires that were available at 6 months, were used twice: once in English and once in Arabic.

Table 2. Top 10 most frequent and most burdensome symptoms*

	Symptom frequency	N (%)	Symptom burden^	Mean (SD)
1	Feeling tired or lack of energy	366 (76.4)	Difficulty becoming sexually aroused	3.42 (1.4)
2	Dry skin	283 (58.7)	Trouble falling asleep	3.26 (1.1)
3	Trouble staying asleep	260 (54.3)	Decreased interest in sex	3.25 (1.5)
4	Muscle cramps	246 (51.0)	Feeling tired of lack of energy	3.24 (1.0)
5	Itching	240 (50.0)	Bone or joint pain	3.23 (1.1)
6	Bone or joint pain	225 (47.0)	Trouble staying asleep	3.18 (1.1)
7	Dry mouth	223 (46.8)	Dry skin	3.04 (1.2)
8	Trouble falling asleep	206 (43.2)	Numbness or tingling in feet	2.99 (1.0)
9	Shortness of breath	207 (43.1)	Restless legs or difficulty keeping legs still	2.94 (1.0)
10	Decreased interest in sex	193 (41.8)	Itching	2.88 (1.0)

** Symptom frequency and burden reported using the Dialysis Symptom Index (DSI): top 10 out of 30 symptoms. Symptoms were available for 459 to 484 patients (90% to 95%).*
^ Average burden score (range: 1-5) reported when symptom was present.

Patient-reported outcomes

Patients experienced on average 10.8 (SD: 6.1) out of 30 symptoms, ranging between 8.0-14.8 symptoms across centres. The overall mean symptom burden score was 30.7 (SD: 22.0) on a scale ranging from 0 (no symptoms) to 150 (all 30 symptoms bother 'very much'). Table 2 presents the 10 most frequently experienced symptoms and 10 most burdensome symptoms. The most common symptom was fatigue, which was experienced by 76% of the patients. 'Difficulty becoming sexually aroused' was – if present – reported as the most bothersome symptom, with a mean score of 3.4 on the 5-point scale. Figure 3 presents the variation among centres in symptom burden score in comparison with the overall mean score.
The mean (SD) scores for physical and mental HRQOL were 35.6 (10.2) and 47.7 (10.6), respectively. Figure 4 and 5 present the variation among centres in PCS and MCS in comparison to the overall mean scores.

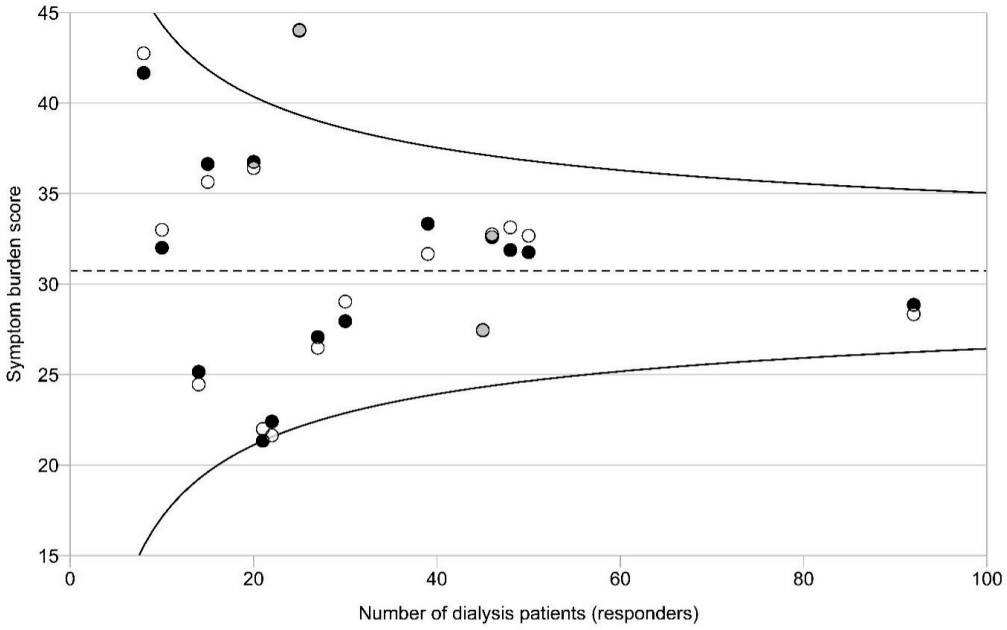


Figure 3. Observed and adjusted mean symptom burden score in 16 pilot centres.
Circles represent the mean observed (white circles) and adjusted (black circles) symptom burden score for each centre. Overlapping part of circles is depicted grey. The overall mean (dotted line) is used as reference in the comparison with each centre. The 95%-CI (curved lines) is provided around the overall mean. The mean score of one centre is outside the 95%-CI, indicating a statistically significant higher symptom burden score compared to the overall mean.*
**Adjusted for sex, age, SES, primary kidney disease, dialysis modality and time on RRT.*
Abbreviations: 95%-CI, 95%-confidence interval; SES, social economic status; RRT, renal replacement therapy.

Variance at centre level

The part of the observed variance in symptom burden, PCS and MCS scores explained by differences among centres was 2.6% (p=0.34), 1.0% (p=0.64) and 1.5% (p=0.45), respectively. The adjusted ICC was 3.1% (p=0.32), 0.6% (p=0.80) and 2.0% (p=0.41), respectively.

Patient experiences and preferences

At 3 months, 214 patients (79%) indicated that they wanted to share and discuss their results on HRQOL and symptom burden with their clinician. In total, 71 individual reports were sent to professionals: 5 patients in each centre, unless fewer patients gave their consent at 3 months (e.g. one centre had no responders at 3

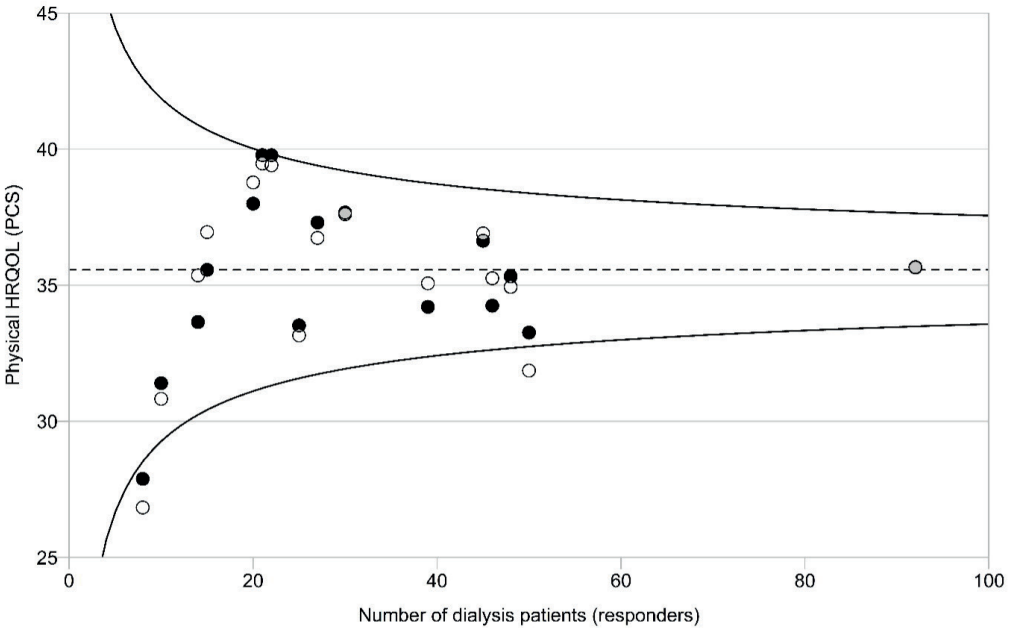


Figure 4. Observed and adjusted mean physical HRQOL (PCS) in 16 pilot centres.
Circles represent the mean observed (white circles) and adjusted (black circles) score for physical HRQOL per centre. Overlapping part of circles is depicted grey. The overall mean PCS (dotted line) is used as reference in the comparison with each centre. The 95%-CI (curved lines) is provided around the overall mean PCS. The adjusted mean score of one centre is outside the 95%-CI, indicating a statistically significant lower PCS compared to the overall mean PCS.*
**Adjusted for sex, age, SES, primary kidney disease, dialysis modality and time on RRT.*
Abbreviations: 95%-CI, 95%-confidence interval; HRQOL, health-related quality of life; PCS, physical component score; SES, social economic status; RRT, renal replacement therapy.

months). At 6 months, 16 patients from 10 different centres indicated that they had discussed the PROM-scores and gave feedback on how they experienced the conversation. Patients discussed the results with a nephrologist (N=11), a nurse (N=8) and/or a social worker (N=2). Patients rated the way in which results were discussed with a mean score of 3.8 (SD: 0.8, score range: 1-5; poor-excellent). Professionals also appreciated discussing patients' PROM-scores and experienced it as insightful. Additionally, professionals indicated that their involvement is important for implementing PROMs into routine care. Moreover, response rates were highest in centres where professionals indicated that they had put a lot of effort into informing and inviting patients.

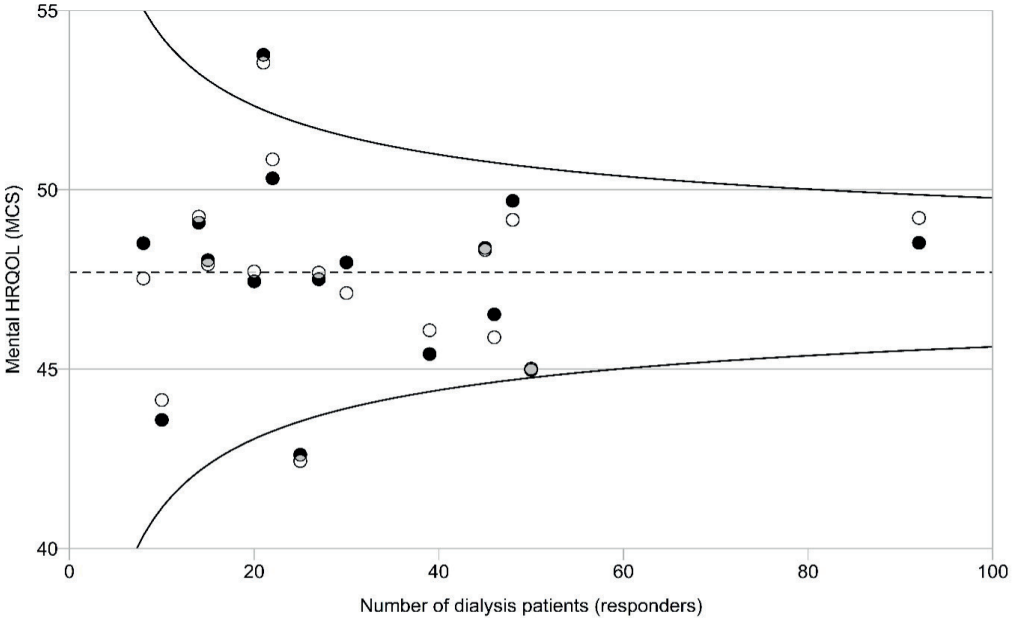


Figure 5. Observed and adjusted mean mental HRQOL (MCS) in 16 pilot centres.
Circles represent the mean observed (white circles) and adjusted (black circles) score for mental HRQOL per centre. Overlapping part of circles is depicted grey. The overall mean MCS (dotted line) is used as reference in the comparison with each centre. The 95%-CI (curved lines) is provided around the overall mean MCS. The mean scores of two centres are outside the 95%-CI: one above and one below the funnel, indicating a statistically significant higher and lower MCS compared to the overall mean MCS, respectively.*
**Adjusted for sex, age, SES, primary kidney disease, dialysis modality and time on RRT.*
Abbreviations: 95%-CI, 95%-confidence interval; HRQOL, health-related quality of life; MCS, mental component score; SES, social economic status; RRT, renal replacement therapy.

Eight patients participated in the focus group: 7 patients were male, aged 33-78 years old, 6 patients received haemodialysis and 2 patients peritoneal dialysis. Five themes were discussed: 'online tool', 'communication about content and purpose', 'benefits of using PROMs', 'feedback is crucial', and 'interpreting PROM-scores'. Examples of corresponding quotations by patients are presented in Table 3. Overall, patients were satisfied with the content, length and structure of the PROMs and the online completion was mentioned as an advantage. Communication about the content and the purpose of PROMs was not always clear for patients. Additional information is needed when receiving the invitation and when completing PROMs. Furthermore, patients indicated that the use of PROMs can contribute to their treatment by providing insight into patient's experienced health for both the clinician and the patient. Additionally, it may enhance patient-clinician communication, as it offers guidance during and in preparation of the conversation. Patients indicated that provision of individual feedback, written and oral, is crucial

Table 3. Examples of corresponding quotations by 8 patients receiving dialysis for the identified themes

Themes	Illustrative quotations
Online tool	"When filling it [the questionnaire] in online, you can also save and keep track of changes [in PROM-scores over time] yourself. This can be an advantage."
Communication about content and purpose	"Titles like PROMs, DSI and SF-12 make no sense. Use clear terms that appeal to the patients, such as 'symptom questionnaire' or 'quality of life questionnaire'."
Benefits of using PROMs	"The questionnaires can be used as a kind of checklist. To help you remember things. (…) The questionnaires help to come up with ideas." "Questionnaires help patients in initiating conversations. Some subjects are difficult to discuss." "You can adjust your treatment goal and plan according to these changes [in PROM-scores] over time, and this can be discussed with your healthcare professional."
Feedback is crucial	"Getting feedback on the results [PROM-scores] should be the basis of each PROMs measurement. After all, it is about your treatment." "Healthcare professionals have the important task to conduct the conversation well. Not every patient is out-spoken and active enough [to express needs and experiences]."
Interpreting PROM-scores	"It is nice to know what other kidney patients score, this gives some context…You want to know if a score of 46 is high, low or average." "I am not very interested in the average [PROM-] score in my hospital… Hospital scores should be available for patients (...) and local patients advocates to address quality improvement."

and that clinicians play an important role in this. Patients mentioned that individual feedback should be presented in a relevant context. They stressed the need for a reference score (e.g. average score of similar patients) to interpret their own results, not to compare their results.

Discussion

This study describes the first experience with PROMs in Dutch routine dialysis care. Overall, response rates were low with high variability among centres. Patients receiving dialysis experienced a high symptom burden and a decreased HRQOL. With regard to these PROM-scores, no centre effect could be observed. Patients believed that discussing HRQOL and symptom burden scores with their healthcare professional was highly insightful and valuable. Individual feedback on PROM-scores was considered crucial.

This is the first study presenting results on HRQOL and symptom burden in Dutch routine dialysis care setting. Patients receiving dialysis experienced a decreased physical HRQOL with an average score of 36 compared to 51 in the general Dutch population (aged 60-69).[40] Mental HRQOL was comparable to the general Dutch population.[40] The substantial symptom burden found is comparable to literature as well.[3, 23, 30] In line with a recent study [41], this study shows that the most common symptoms are not necessarily the ones that bother patients the most. The importance of certain symptoms may be different for each patient. Therefore, it is important to monitor and discuss the presence and burden of symptoms in order to understand what is most important to each patient. Further research is needed on how individual PROM-scores can be best used to address their needs.

Patients and professionals were very positive about the use of PROMs, in which they considered provision of individual feedback to be crucial. These first results are promising and imply that PROMs are suitable for use at individual level. The number of patients (n=16, 23%) that indicated to have discussed their PROM-scores seems low, but is proportionally similar compared to the response rate at 6 months. Moreover, the real number of patients that discussed their PROM-scores is probably higher, for instance because they discussed their report after the third time-point. Since all patients and professionals who discussed the PROM-scores highly appreciated the conversation, we decided to send the individual reports of the remaining patients (who gave consent) to their professionals and to include individual PROM reports into the electronic registration system.

Results from the focus group suggest that PROMs can provide insight into experienced health and needs, improve patient-professional communication and in-

crease shared decision-making. Similar potential benefits of PROMs are described in literature [9-13], however, there is a paucity of evidence on whether and how the use of PROMs actually leads to improvements in patients' outcomes. There are some studies suggesting that using PROMs will lead to better outcomes, for example: a randomized controlled trial in routine cancer care showed that web-based symptom monitoring resulted in improved HRQOL after 6 months, less hospital admissions and better 1-year survival, even though no specific guidance was provided to professionals on how to respond to reported symptoms.[42] Scholars also argue that patients receiving dialysis expect improvements when using PROMs, for instance: improved symptom experience as a consequence of improved patient-professional communication about symptoms.[41] However, further research is needed to investigate whether and how the use of PROMs leads to long term improvements in healthcare quality and outcomes in patients receiving dialysis.

The low response rate in this study is similar to the response rate (31%) of the Scottish Renal Registry when first introducing PROMs, confirming that it is challenging to incorporate PROMs into routine care.[18] Several factors may explain our results. First, professionals play an important role in informing and motivating patients. Highest response rates were observed in centres where professionals were highly engaged in the process. Therefore, interventions to increase professionals' engagement may be beneficial. Previous studies show that training and guidance on why and how to use PROMs and how to act in response to individual PROM-scores may facilitate the uptake of PROMs by professionals.[24, 43, 44] In the Scottish registry, interventions to improve patient information letters and staff awareness indeed resulted in an increase of their response rate to 48%.[18] Second, the process of inviting patients was regulated by each centre independently to promote incorporation into their workflow. A drawback of this approach may be that not every centre organized this in a structured way or had the desired facilities (e.g. availability of resources such as a process coordinator, printer and internet access) and consequently, some patients may not have been invited. Moreover, differences across centres existed with regard to the type and amount of support that patients received when filling in PROMs (e.g. availability of electronic devices in centres), by whom support is provided (e.g. medical staff or partner) and at which location (home or medical centre). It is possible that the centres' support and the possibility to complete PROMs on site, contributed to higher response rates. On the other hand, the availability or lack of support in centres could also have influenced patients' responses. However, we did not observe differences among centres with regard to PROs. Third, some patients are more likely to participate than others. In line with literature, we found that older patients with a higher SES [18] and male

[45] patients were more likely to participate. Further research focussing on non-responders is needed to gain more insight into barriers and potential facilitators for participation in order to implement recruitment strategies tailored to these more difficult-to-reach patients.[46, 47]

Higher response rates are needed for optimal use of PROMs at patient level (e.g. individualized prognosis) and aggregated level (e.g. evaluation of healthcare quality).[9] Based on this study and literature, we provide the following recommendations to increase the response rate. First, provide additional training and support to increase engagement of healthcare professionals and to reinforce the professionals' feeling of being comfortable and able to handle PROM-scores.[48, 49] Second, recruitment strategies should be improved and, given that dialysis patients regularly encounter healthcare professionals, recruitment strategies should particularly focus on tailored communication (e.g. on personal relevance and confidentiality) and support (e.g. completing PROMs online).[25] Third, communication between stakeholders should be improved, for instance by supportive resources such as provision of material to inform patients, individualized reports on PROM-scores, and updates on centres' experiences (best practices), response rates and outcomes.[25, 48] Fourth, logistics should be further developed to improve response rates and to support professionals and patients in using PROMs in clinical practice, for example: provide individual reports directly after PROMs completion, incorporate PROMs into the electronic health record, and send automated invitations (e.g. prior to patient's upcoming consultation visit [24]) and reminders to complete PROMs.[50] Finally, we propose to assess and discuss PROM-scores twice per year, as we believe this provides insight into patient's outcomes over time with minimal burden to patients and professionals. Some centres suggested using PROMs during a more extended consultation, such as an annual check-up, to discuss PROs progression over time, patients' needs and treatment goals.

The low response rate and selective response are important results, but also limitations in this study. For instance, our results suggest that there is no relevant centre-effect on patients' HRQOL and symptom burden, however, possibly real centre-effects could not be detected due to low and selective responses. Furthermore, responders are likely to be more health conscious and involved in healthcare compared to non-responders (i.e. healthy responder bias). For example, the patients who shared their experience about discussing PROM-scores may be more involved and may have a more positive attitude towards using PROMs in clinical practice, which should be taken into account when interpreting the results. Additionally, the selective response may have led to effect underestimation of patients' outcomes: symptom burden is likely to be higher and HRQOL lower in the

total dialysis population. However, information about non-responders was also presented and can therefore be taken into account when interpreting the results. Although current data may be insufficient to evaluate healthcare quality, the electronic registration of PROMs as part of Renine is designed in such a way that future data may be used for this purpose.[9] We believe that it is a major strength that PROMs can be used both at individual level in clinical practice and at aggregated level to evaluate healthcare quality. Possibly, this combination is crucial, as the use of PROMs for individual patient's treatment may be the most important factor in reaching sufficient response rates to enable evaluation of healthcare quality.

Another strength is the multicentre study design and methods used in this study. With 16 participating centres, a substantial sample of all Dutch dialysis patients was included. Additionally, by leaving centres free to incorporate PROMs into their workflow, a broad variation of in-centre processes was included, which may provide valuable information (e.g. insight into best practices for using PROMs in clinical settings) and may eventually promote adaptation and implementation of PROMs into clinical practice (e.g. due to limited workflow disruptions and research processes that are in line with priorities of patients and professionals).[13, 24, 25] Moreover, all relevant stakeholders (e.g. patients, healthcare professionals and researchers) were involved from the start, resulting in widely supported PROMs that fit clinical practice and research.[13, 24, 25] Besides, during the developmental phase much attention has been paid to the electronic registration system and selection of valid questionnaires.[23, 24] The pilot study confirms that the questionnaires were suitable and feasible, with only minor suggestions for improvement. Finally, by making use of both quantitative and qualitative methods, we obtained a broad picture of perceived benefits and barriers for implementing PROMs into nephrological care and possibilities for improvement.

In conclusion, first results from the Dutch registry of PROMs in patients receiving dialysis showed low response rates with a high centre variability. Achieving higher response rates is challenging and requires extra encouragement of patients and professionals. Patients experienced a high symptom burden and a decreased physical HRQOL. Discussing symptom and HRQOL results was greatly appreciated and considered crucial for the use of PROMs in routine care. Further research is needed to investigate how collection and use of PROMs can be successfully integrated into routine care in order to improve healthcare quality and outcomes.

References

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

2. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

3. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

4. Weisbord SD, Fried LF, Mor MK, et al. Renal provider recognition of symptoms in patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2007;2(5):960-7

5. Claxton RN, Blackhall L, Weisbord SD, Holley JL. Undertreatment of symptoms in patients on maintenance hemodialysis. *J Pain Symptom Manage* 2010;39(2):211-8

6. Pugh-Clarke K, Read SC, Sim J. Symptom experience in non-dialysis-dependent chronic kidney disease: A qualitative descriptive study. *J Ren Care* 2017;43(4):197-208

7. Feldman R, Berman N, Reid MC, et al. Improving symptom management in hemodialysis patients: identifying barriers and future directions. *J Palliat Med* 2013;16(12):1528-33

8. Stephens RJ, Hopwood P, Girling DJ, Machin D. Randomized trials with quality of life endpoints: are doctors' ratings of patients' physical symptoms interchangeable with patients' self-ratings? *Qual Life Res* 1997;6(3):225-36

9. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q* 2014;92(4):754-75

10. Bingham CO, 3rd, Noonan VK, Auger C, Feldman DE, Ahmed S, Bartlett SJ. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 4: patient-reported outcomes can inform clinical decision making in chronic care. *J Clin Epidemiol* 2017;89:136-41

11. Catarinella FS, Peters DTJM, de Roos K-P, et al. Praktische handvatten voor het gebruik van Patient-Reported Outcome Measures (PROMs). *Nederlands Tijdschrift voor Heelkunde* 2017;26(4):29-33

12. Calvert M, Kyte D, Price G, Valderas JM, Hjollund NH. Maximising the impact of patient reported outcome assessment for patients and society. *Bmj* 2019;364:k5267

13. Anderson NE, Calvert M, Cockwell P, Dutton M, Kyte D. The Use of Patient-Reported Outcomes in Patients Treated With Maintenance Hemodialysis: A Perspective. *Am J Kidney Dis* 2019

14. Urquhart-Secord R, Craig JC, Hemmelgarn B, et al. Patient and Caregiver Priorities for Outcomes in Hemodialysis: An International Nominal Group Technique Study. *Am J Kidney Dis* 2016;68(3):444-54

15. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. *Clin J Am Soc Nephrol* 2014;9(10):1813-21

16. Verberne WR, Das-Gupta Z, Allegretti AS, et al. Development of an International Standard Set of Value-Based Outcome Measures for Patients With Chronic Kidney Disease: A Report of the International Consortium for Health Outcomes Measurement (ICHOM) CKD Working Group. *Am J Kidney Dis* 2019;73(3):372-84

17. van der Veer SN, Aresi G, Gair R. Incorporating patient-reported symptom assessments into routine care for people with chronic kidney disease. *Clin Kidney J* 2017;10(6):783-7

18. Nimmo A, Bell S, Brunton C, et al. Collection and determinants of patient reported outcome measures in haemodialysis patients in Scotland. *Qjm* 2018;111(1):15-21

19. Evans K, Pyart R, Steenkamp R, et al. UK Renal Registry 20th Annual Report: Introduction. *Nephron* 2018;139 Suppl 1:1-12

20. Breckenridge K, Bekker HL, Gibbons E, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. *Nephrol Dial Transplant* 2015;30(10):1605-14

21. Biber J, Ose D, Reese J, et al. Patient reported outcomes - experiences with implementation in a University Health Care setting. *J Patient Rep Outcomes* 2017;2:34

22. Noonan VK, Lyddiatt A, Ware P, et al. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 3: patient-reported outcomes can facilitate shared decision-making and guide self-management. *J Clin Epidemiol* 2017;89:125-35

23. van der Willik EM, Meuleman Y, Prantl K, et al. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease – a four-phase mixed methods study. *BMC Nephrology* 2019;20(1):344

24. Foster A, Croot L, Brazier J, Harris J, O'Cathain A. 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

25. Flythe JE, Narendra JH, Hilliard T, et al. Cultivating a Research-Ready Dialysis Community. *J Am Soc Nephrol* 2019

26. Ware JE, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-33

27. Loosman WL, Hoekstra T, van Dijk S, et al. Short-Form 12 or Short-Form 36 to measure quality-of-life changes in dialysis patients? *Nephrol Dial Transplant* 2015;30(7):1170-6

28. Ware JE, Kosinski M, Keller SD: SF-12: How to score the SF-12 physical and mental health summary scales, Boston, MA, The Health Institute, New England Medical Center, 1995

29. Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *J Pain Symptom Manage* 2004;27(3):226-40

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

31. Registry E-E. ERA-EDTA Registry Annual Report 2016. 2018.

32. Statusscores 2016 [Internet]. 2018 [cited 10 December 2018]. Available from: https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores/statusscores_2016.

33. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30(4):377-99

34. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epide-

miological and clinical research: potential and pitfalls. *Bmj* 2009;338:b2393

35. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006;59(10):1092-101

36. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;24(8):1185-202

37. Merlo J, Chaix B, Yang M, Lynch J, Rastam L. A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon. *J Epidemiol Community Health* 2005;59(6):443-9

38. Merlo J, Yang M, Chaix B, Lynch J, Rastam L. A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people. *J Epidemiol Community Health* 2005;59(9):729-36

39. van Dishoeck AM, Lingsma HF, Mackenbach JP, Steyerberg EW. Random variation and rankability of hospitals using outcome indicators. *BMJ Qual Saf* 2011;20(10):869-74

40. Mols F, Pelle AJ, Kupper N. Normative data of the SF-12 health survey with validation using postmyocardial infarction patients in the Dutch population. *Qual Life Res* 2009;18(4):403-14

41. Flythe JE, Hilliard T, Castillo G, et al. Symptom Prioritization among Adults Receiving In-Center Hemodialysis: A Mixed Methods Study. *Clin J Am Soc Nephrol* 2018;13(5):735-45

42. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* 2016;34(6):557-65

43. Bausewein C, Simon ST, Benalia H, et al. Implementing patient reported outcome measures (PROMs) in palliative care--users' cry for help. *Health Qual Life Outcomes* 2011;9:27

44. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23(6):508-18

45. Stengel B, Metzger M, Combe C, et al. Risk profile, quality of life and care of patients with moderate and advanced CKD : The French CKD-REIN Cohort Study. *Nephrol Dial Transplant* 2018

46. Thurman WA, Harrison TC. Reaching the "Hard-to-Reach": Recruitment of Rural-Dwelling Adults With Disabilities. *J Transcult Nurs* 2019:1043659619856667

47. van der Waerden JE, Hoefnagels C, Jansen MW, Hosman CM. Exploring recruitment, willingness to participate, and retention of low-SES women in stress and depression prevention. *BMC Public Health* 2010;10:588

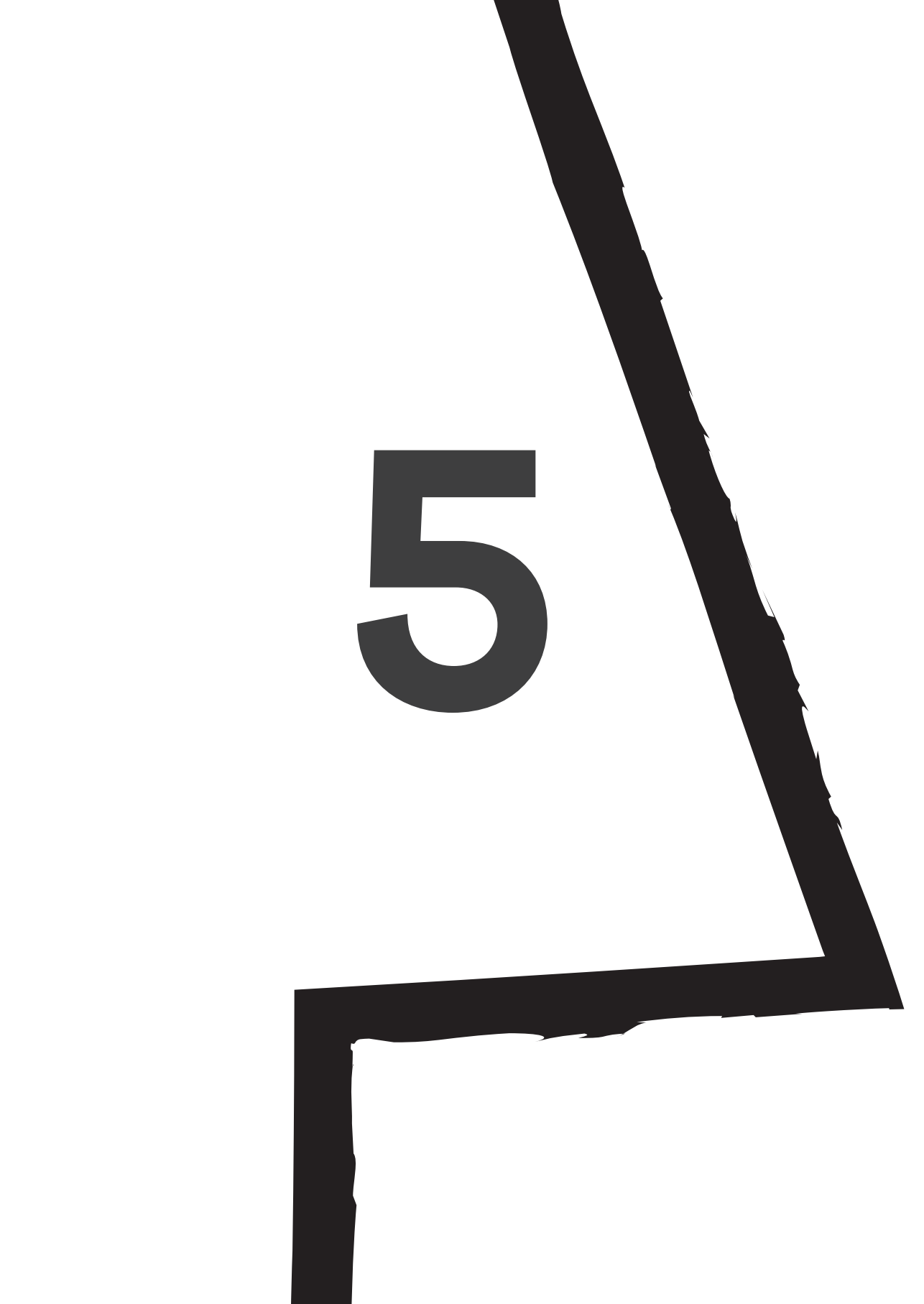
48. Zhang R, Burgess ER, Reddy MC, et al. Provider perspectives on the integration of patient-reported outcomes in an electronic health record. *JAMIA Open* 2019;2(1):73-80

49. Hardee JT, Rehring TF, Cassara JE, Weiss K, Perrine N. Effect and Durability of an In-depth Training Course on Physician Communication Skills. *Perm J* 2019;23

50. Nakash RA, Hutton JL, Jorstad-Stein EC, Gates S, Lamb SE. Maximising response to postal questionnaires--a systematic review of randomised trials in health research. *BMC Med Res Methodol* 2006;6:5

Part 2

Use of PROMs at population level and in individual patients in routine nephrology care


5

Funnel plots of patient-reported outcomes (PROs) to evaluate healthcare quality: basic principles, pitfalls and considerations

Esmee M. van der Willik, Erik W. van Zwet, Tiny Hoekstra, Frans J. van Ittersum, Marc H. Hemmelder, Carmine Zoccali, Kitty J. Jager, Friedo W. Dekker, Yvette Meuleman

Nephrology (Carlton). 2021;26(2):95-104.

Abstract

A funnel plot is a graphical method to evaluate healthcare quality by comparing hospital performances on certain outcomes. So far, in nephrology, this method has been applied to clinical outcomes like mortality and complications. However, patient-reported outcomes (PROs; e.g. health-related quality of life [HRQOL]) are becoming increasingly important and should be incorporated into this quality assessment. Using funnel plots has several advantages, including: clearly visualized precision, detection of volume-effects, discouragement of ranking hospitals and easy interpretation of results. However, without sufficient knowledge of underlying methods, it is easy to stumble into pitfalls, such as: overinterpretation of standardized scores, incorrect direct comparisons of hospitals and assuming a hospital to be in-control (i.e. to perform as expected) based on underpowered comparisons. Furthermore, application of funnel plots to PROs is accompanied by additional challenges related to the multidimensional nature of PROs and difficulties with measuring PROs. Before using funnel plots for PROs, high and consistent response rates, adequate case mix correction and high-quality PRO measures are required. In this article, we aim to provide insight into the use and interpretation of funnel plots by presenting an overview of the basic principles, pitfalls and considerations when applied to PROs, using examples from Dutch routine dialysis care.

Introduction

In the last decade, healthcare has shifted towards a more patient-centred and value-based approach, resulting in a stronger focus on healthcare outcomes.[1, 2] Reasons for measuring outcomes are to gain insight into hospital performance and encourage healthcare quality improvement.[2-4] Quality can be improved, for instance, because hospitals can learn from each other (i.e. adopt best practice) and initiate improvement strategies.[3, 4] Patients can also make better informed decisions, for example in which hospital to start dialysis treatment.[3-5] Additionally, strategies by insurance companies (e.g. value-based payment) and government (e.g. regulations on quality) can also reward and stimulate higher quality of care.[3, 4] Insight into hospital performance can be obtained through outcome comparison using funnel plots.[6] This graphical method is common in meta-analysis to gain insight into potential publication bias. For hospital comparison, funnel plots have been applied to clinical outcomes, for example: the standardized mortality ratio in which the observed and expected number of deaths are compared.[7] Figure 1 depicts such an example from Dutch dialysis care[8]: the standardized mortality rate in each dialysis centre (circles) is being compared to the national mortality rate in

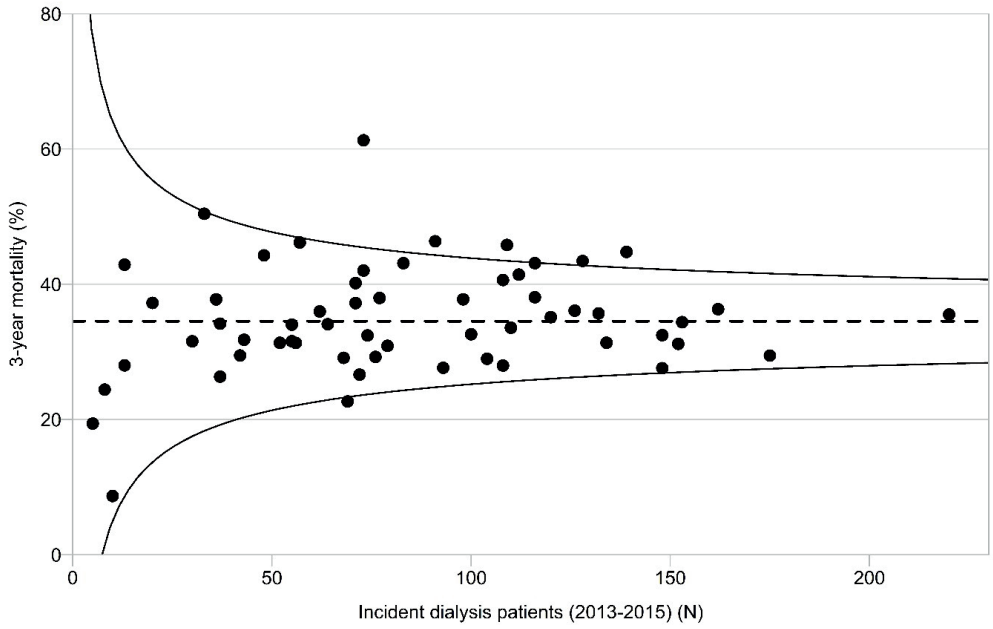


Figure 1. Funnel plot on 3-year mortality in incident dialysis patients.
Inclusion period 2013-2015. Circles represent the standardized mortality rates of 58 Dutch dialysis centres. The overall mortality rate in all incident dialysis patients is used as reference standard. *Case mix factors include: age, sex, social economic status and primary kidney disease categories. (Figure obtained from Renine annual report 2018[8]).*

dialysis patients (dashed line). Some variation in outcome can be observed across the centres and a few centres exceed the funnel-shaped control limits, which may indicate either excellent performance or underperformance. In such cases, further investigation and initiatives may be necessary to improve healthcare quality. Although funnel plots are regularly regarded as being intuitive and easy to interpret[6, 9], some knowledge about the method is needed for correct interpretation. For example: the hospital rates depicted in Figure 1 may, intuitively, be interpreted as observed mortality rates, while actually relative measures are presented for comparison with the national mortality rate in dialysis patients. This example underlines the necessity for understanding the underlying methods to prevent incorrect interpretation.

Furthermore, various outcomes can provide insight into healthcare quality and should be taken into account when evaluating hospital performances. Nowadays, patient-reported outcomes (PROs; e.g. health-related quality of life [HRQOL] and symptom burden) are considered important healthcare outcomes and PRO measures (PROMs) are increasingly being implemented into routine care, including nephrological care.[10-13] Therefore, the logical next step is to include PROs – in addition to clinical outcomes – in the process of healthcare quality evaluation. However, incorporation of PROs and using funnel plots for PROs is accompanied with additional challenges. For example: low and selective response rates are common for PROs and may lead to generalisability problems and incorrect conclusions. Therefore, in this paper we will provide insight into the use and interpretation of funnel plots for PROs by presenting an overview of the basic principles, common pitfalls and considerations, using examples from Dutch routine dialysis care.

Basic principles of funnel plots

Funnel plots are considered a suitable graphical method to present information on hospital performance in comparison to a reference standard and by taking random variation into account.[6, 9] A funnel plot consists of 4 components (Figure 2): 1. an indicator, which is the measure of performance on a certain outcome; 2. a benchmark, which is the reference standard to compare hospitals with; 3. a measure of precision that is related to the certainty of the comparison; and 4. control limits to identify statistical differences for a certain p-value. Hospitals exceeding these control limits may be considered as either underperforming or overperforming. The statistical details of these different components have been described elsewhere.[6] Below, we will elaborate on the underlying methods of funnel plot components, using examples from Dutch routine dialysis care. Data on

PROs (HRQOL and symptom burden), sociodemographic and clinical characteristics of patients receiving dialysis treatment were obtained from Renine, the Dutch renal registry (www.nefrovisie.nl/renine). For more information about the Dutch PROMs registry, see Van der Willik et al. (2019, 2020).[10, 14]

Indicator of performance

In a funnel plot, hospital comparisons are made for a certain outcome using an *indicator* or *performance-indicator*. To be considered a valuable indicator, an outcome has to meet certain criteria, for example: it must be relevant, measurable, changeable and related to healthcare quality, and there must be variation across hospitals. The indicator is presented on the y-axis of the funnel plot and can be either the outcome as observed (i.e. crude analysis) or an indicator wherein differences in hospital populations are taken into account (i.e. adjusted analysis). The latter indicator includes the comparison between the observed outcome and the outcome that would be expected in that specific hospital (see heading 'Adjustment for differences in hospital populations').

Benchmark: reference standard

Benchmarking is the process of measuring and evaluating the hospital's own performance by comparing it to a reference standard (i.e. the benchmark) with the purpose of improving the hospital's own performance and quality of care. Often the total population of interest (e.g. national average) or a certain norm is chosen as reference standard for comparison. In a funnel plot, the *reference standard* or *target outcome* is presented as a horizontal line at the corresponding value for the indicator on the y-axis. For example: the national 1-year mortality rate (Figure 1) or the average physical HRQOL score (Figure 2) of Dutch dialysis patients (i.e. the reference population) can serve as a reference standard.

Selecting a suitable reference standard can be challenging since the reference standard must be a fair and feasible comparator for all hospitals. Some background knowledge on the outcome in the specific population of interest is needed to assess what can be expected or considered relevant. Additionally, high-quality data on the reference population must be available. The latter could be a concern when using PROs, since response rates rarely reach 100% in routine care (Figure 3) and some people are more likely to participate than others, resulting in a reference standard that may not fully represent the population of interest.[10-12] Box 1 describes how this selective response may cause generalisability problems or even selection bias.

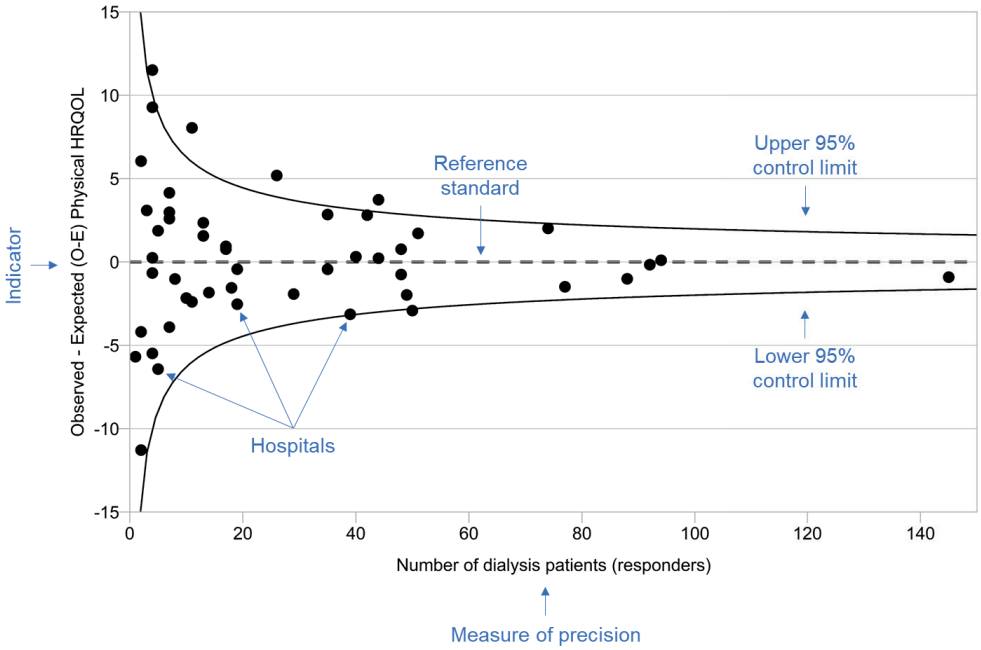


Figure 2. Components of a funnel plot for hospital comparison.
An example is shown of a funnel plot on physical HRQOL in 48 Dutch dialysis centres that participated in the Dutch registry of PROMs in 2019. The indicator shows the comparisons between the centres' observed and expected scores on physical HRQOL. The total study population of Dutch dialysis patients is used as a reference standard. The 95% control limits are provided around the reference standard.*
**Expected scores were based on the following case mix factors: sex, age, social economic status, primary kidney disease, dialysis modality and time on renal replacement therapy.*

Measure of precision

The x-axis of a funnel plot presents a *measure of precision*, which is a variable that determines the precision of the indicator. Usually, the sample size or the number of (expected) cases is used as measure of precision, since a larger sample size is accompanied with more precision. By choosing such an easily interpretable measure, both the random variation (through 'control limits'; see heading below) and potential volume-effects (see 'relationship with volume') are clearly visualized.

Control limits

Control limits corresponding to a certain *p*-value are plotted around the reference standard. As control limits include a measure of precision, the width of the limits changes with the x-axis, resulting in funnel-shaped limits around the reference standard. Often the 95% control limits (corresponding to $p = 0.05$) are presented,

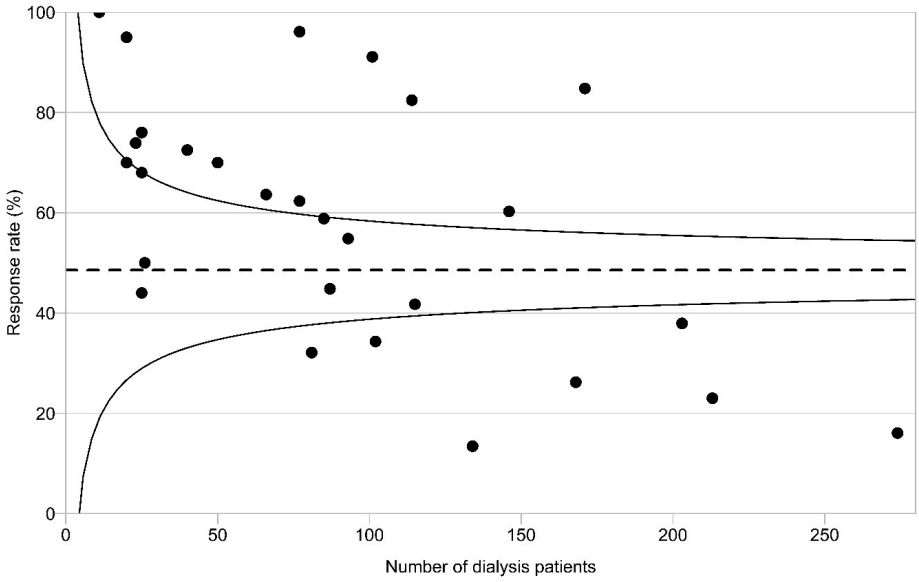


Figure 3. Funnel plot of response rates on PROMs in 28 Dutch dialysis centres.
Circles represent the response rates in Dutch dialysis centres that participated in the Dutch registry of PROMs in 2019. The total number of dialysis patients that was invited to complete the PROMs is presented on the x-axis. The figure shows large variation in response rates across dialysis centres. The response rate seems lower in centres that invited more patients, which may indicate a volume-effect.*
**The total number of dialysis patients was based on the number of patients for which an invitation to complete the PROMs was downloaded from the electronic registry environment. Twenty centres (42%) did not use the registry invitations and their data only included patients that participated through the DOMESTICO study.[15] For these centres the number of invited patients is unknown in the registry, and therefore these centres were excluded from this funnel plot.*

whereby a 5% chance of a type I error is accepted. In other words, hospitals that perform similar to the reference population have a 5% chance to exceed the limits: 2.5% at the upper limit and 2.5% at the lower limit.

Adjustment for differences in hospital populations

Case mix

To enable fair hospital comparisons, differences in characteristics of the hospital population or 'case mix' must be taken into account to ensure that differences in hospitals' performance are investigated rather than differences in population. Hence, adjusting for case mix is identical to adjusting for confounding. For example: differences across dialysis centres with regard to patients' age or sex should

Box 1. Response rates – why are high and consistent rates needed?

In contrast to clinical outcomes, PROs can only be observed and reported by the patient himself, which inherently leads to concerns about response rates. Especially in routine chronic and advanced care, response rates that reach 100% are very rarely achieved.[10-12] Obviously, lower response rates result in lower sample sizes and thus, less precision (as clearly visualized by the funnel shaped control limits that narrow with larger sample sizes). Low response rates may be reasons for concern, especially for low-volume hospitals who already deal with power issues.[16] However, the main problem of low response rates is the selective response: some people are more likely to participate than others[10, 11], which may result in generalisability problems and selection bias. (See also Figure 3 and Supplementary Table S1)

Generalisability

The reference standard is based on people that completed PROMs, which could make the selection of a suitable reference standard challenging. Selective response in the reference population, results in a reference standard that may not fully reflect the population of interest. The same issue exists on a hospital level: the group responders may not be generalizable to the total hospital population, making it difficult to draw conclusions about performance in patients treated in that hospital. Insight into characteristics of (non-) responders can be helpful when interpreting the results. Additionally, recruitment strategies should be aimed at reaching all (types of) patients.

Selection bias

Several factors may determine whether patients complete PROMs or not. For example: participation may be influenced by the hospital's facilities and engagement of the medical team, and by the patient's characteristics or health state (e.g. fatigue). If this factor is also associated with the outcome, selection bias may occur. By including only responders in the analysis, an association is created between the hospital and the outcome that may not actually exist. To account for this, insight into these mechanisms and data on factors influencing response from both responders and non-responders are needed. Furthermore, it is important to use similar recruitment strategies and to strive for high but also comparable response rates across hospitals.

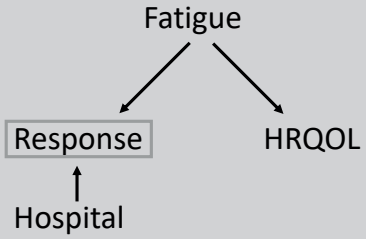


Figure B1. Example selection bias

Box 2. Identifying case mix factors for PROs – what makes it so difficult?

Hospital comparison research usually aims to explore whether there is an association between the treating hospital and the patients' outcome. Herein, factors that affect both the outcome and the hospital in which the patient is treated should be taken into account, i.e. confounding factors. To this end, the term case mix is used: the composition of patient- and disease characteristics (that affect the outcome) in the hospitals' populations, for which you want to correct. For each outcome, different case mix variables may be important to correct for. Therefore, case mix adjustment models are very likely to differ across outcomes (e.g. clinical outcomes and PROs will most likely have different underlying mechanisms).[19] The difficulty lies in selecting the right case mix factors to correct for. For example: symptom burden is associated with the outcome HRQOL [20] and may vary across hospitals [10]. If we assume symptom burden to be a disease characteristic reflecting a certain health state or the severity of disease, we may want to adjust for this. However, scholars also argue that symptom burden can be influenced by healthcare and can therefore be considered a consequence of healthcare quality as well, for which we do not want to correct. Thus, the selection of case mix factors is dependent on the assumptions made, which is often based on literature. Given the multidimensional and complex nature of PROs such as HRQOL, it may be challenging to achieve sufficient case mix correction. More research on which factors and through which mechanisms PROs are influenced may contribute to the selection of an adequate set of covariates to correct for.

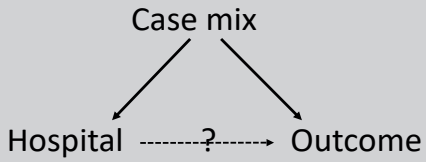


Figure B2. Example confounding

be taken into account (see also Supplementary Table S1). The difficulty is selecting a sufficient set of true case mix factors (e.g. no mediators) to correct for[17], which may be even more difficult for PROs, given the multidimensional nature of outcomes such as HRQOL (see Box 2 for further explanation).[3, 18] Moreover, for both clinical outcomes and PROs, some residual confounding is inevitable.

Indirect standardization

In funnel plots, case mix differences are taken into account by performing indirect standardization.[21] This method is suitable for the evaluation of a hospital's performance as it demonstrates how the outcomes observed in the hospital relate to what can be expected based on the reference standard and given the hospital's case mix. When using indirect standardization, the performance of the reference standard is applied to the hospital population (by strata of case mix characteristics). For each patient, based on his characteristics, the outcome (e.g. HRQOL score) is calculated that he would have had, if he had been treated in a hospital that performs similarly to the reference standard. The calculation of these individual predicted scores is usually performed using regression analysis. The mean of all individual predicted scores is equal to the expected (E) score of the hospital and this expected score is then compared to the observed (O) score of the hospital.[21]
The comparison between O and E (i.e. the indicator) is presented on the y-axis either as a ratio (O/E), a difference (O-E) or a standardized score (multiplicative: O/E*reference score or additive: O-E+reference score). Depending on whether the indicator is presented as ratio or as difference, the target outcome is 1 or 0 respectively, because E equals O within the reference population (O/E=1 or O-E=0). The multiplicative and additive standardized scores differ only in 'starting point' on the scale from the ratio and difference, respectively, and thus, result in the same picture for hospital comparison. For example: Figure 4a (O-E) and Figure 4b (O-E+reference score) present the same data, both on an additive scale (see also Box 3).
Irrespective of how the results are presented, the hospital's score should be interpreted in comparison to the reference standard. Individual hospitals are, even after standardization, not directly comparable, because each hospital's own population is used to calculate the expected scores. The indicator thus shows how well a hospital performs within its own population, in comparison to the performance of the reference standard. Box 3 elaborates on how results can and cannot be interpreted.

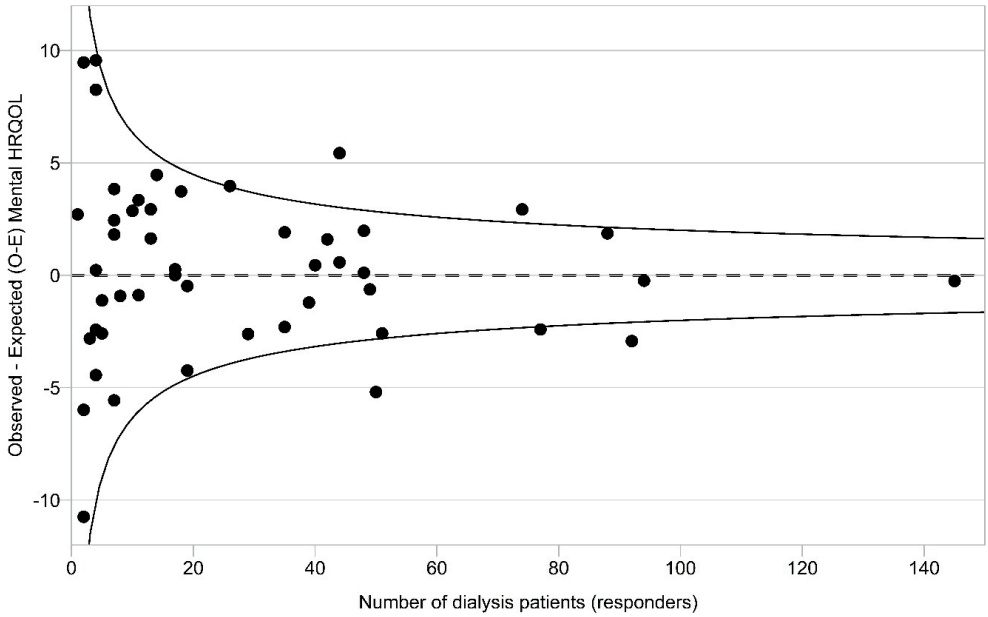


Figure 4a. Funnel plot of comparison between observed and expected scores on mental HRQOL in 48 Dutch dialysis centres.

Circles represent the difference between the centres' observed and expected scores on mental HRQOL of 48 centres that participated in the Dutch registry of PROMs in 2019. The total study population of Dutch dialysis patients is used as a reference standard (dashed line) to compare centres with. The 95% control limits (curved lines) are provided around the reference standard. Four centres exceed the 95% control limits, indicating statistically significant lower (two centres) or higher (two centres) scores on mental HRQOL compared to the reference standard. *Case mix factors included: sex, age, social economic status, primary kidney disease, dialysis modality and time on renal replacement therapy.*

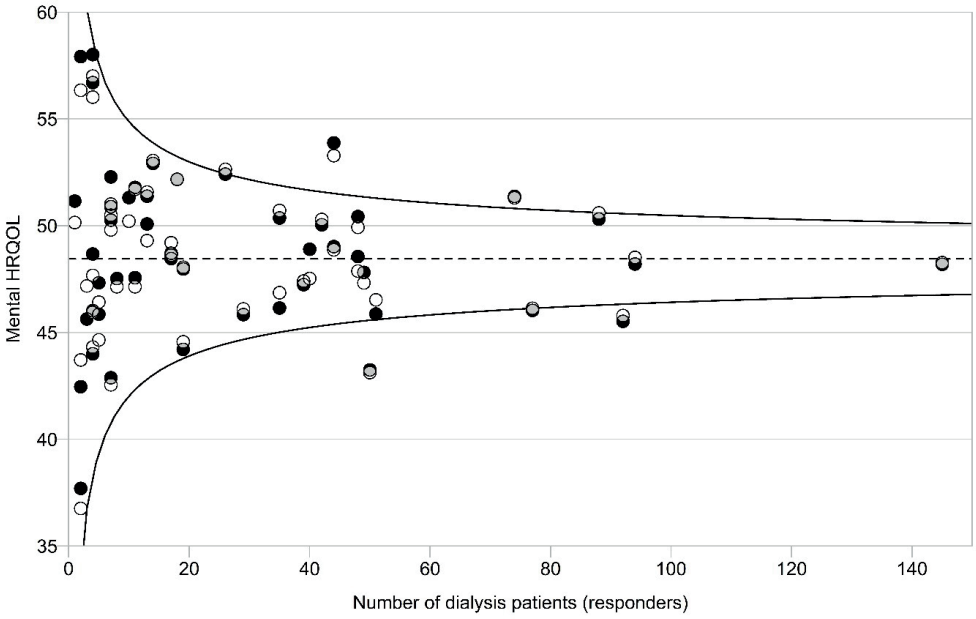


Figure 4b. Funnel plot of observed and standardized scores on mental HRQOL in 48 Dutch dialysis centres.

Circles represent the mean observed (white circles) and standardized (black circles) scores on mental HRQOL of 48 centres that participated in the Dutch registry of PROMs in 2019. Overlapping part of circles is depicted grey. The overall mean score on mental HRQOL of all Dutch dialysis patients (dashed line) is used as reference standard to compare centres with. The 95% control limits (curved lines) are provided around the reference standard. The standardized scores of four centres exceed the 95% control limits, indicating statistically significant lower (two centres) or higher (two centres) scores on mental HRQOL compared to the reference standard. *Standardized score = observed score – expected score + reference score. The following case mix factors were included to calculate the expected scores: sex, age, social economic status, primary kidney disease, dialysis modality and time on renal replacement therapy.*

Box 3. Indirect standardization – what do results say, and what not?

In indirect standardization, the observed outcome in each hospital is compared to the expected outcome, which is the outcome that would be observed if the hospital's performance is equal to the reference standard. To illustrate this, we will use an example: Hospital A and B are compared to the total Dutch dialysis population (i.e. the reference standard). Hospital A has an older and more fragile dialysis population, and Hospital B has a younger and less fragile dialysis population. The total Dutch dialysis population contains a heterogeneous group of patients, from which the outcomes in the populations of Hospital A and B can be predicted. Example scores on mental HRQOL are shown below (Table B3).

Table B3. Example observed, expected and standardized scores on mental HRQOL.

	Older and more fragile patients (Hospital A)	All dialysis patients (Reference standard)	Younger and less fragile patients (Hospital B)
Observed score (O)	45	48	50
Expected score (E)	40	48	58
O – E	+ 5	0	- 8
O – E + reference score (standardized score)	53	48	40

Table B3 clearly shows that Hospital A is performing better (+ 5 points) and Hospital B is performing worse (- 8 points) within their population compared to the reference standard (i.e. all dialysis patients). This example also illustrates why Hospital A and Hospital B cannot be compared: both have a different population, and thus a different expected score. We do not know how Hospital A will perform in younger and less fragile patients, and we also do not know how Hospital B will score in older and more fragile patients. Of course, in practice, there is some overlap in population characteristics, but as long as the composition differs, you cannot make direct comparisons. If you want to compare Hospital A to Hospital B, one or the other must be used as a reference standard or direct standardization methods should be applied.

The comparison between observed and expected scores can be presented as either a difference, a ratio or a standardized score. Preference may be given to presenting the difference or ratio, since these measures clearly describe the comparison. The standardized score seems attractive, since the original scale of the outcome can be used and therefore also observed scores can be presented using the same funnel plot (Figure 4b), but can easily be overinterpreted. The standardized score is also meant to be interpreted in comparison to the reference score and the standardized score itself has no clear interpretation. For example: Hospital A's standardized score of 53 is not the mental HRQOL-score that you would expect from the population of Hospital A, neither the predicted score if Hospital A had treated all Dutch dialysis patients or any other population. It is only a representation of the 5 points difference with the reference standard. This comparison is illustrated below in Figure B3.

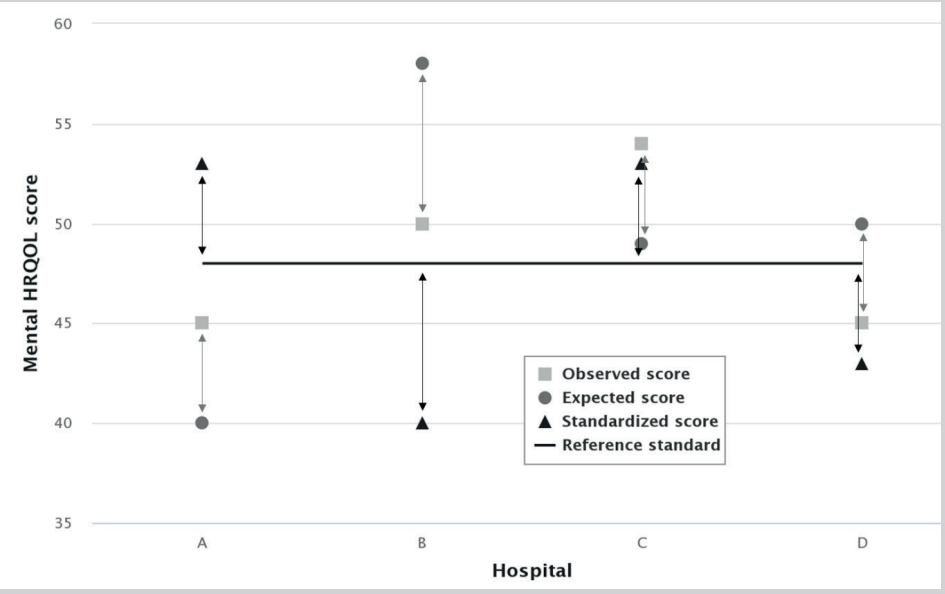


Figure B3. *Illustration of observed, expected and standardized score in Hospital A-D based on fictive data on mental HRQOL. Hospital A and B are also presented in Table B3. Note that the distance between observed and expected score is equal to the distance between standardized score and reference standard.*

Interpretation of funnel plots

General interpretation

In the first place, funnel plots provide a general overview of the variability between hospitals and present information for benchmarking purposes: it provides hospitals with insight into their performance within their own population in comparison to the reference standard. Hospitals' scores that exceed the lower or upper control limit indicate a statistically significant lower or higher score, i.e. over- or underperformance, compared to the reference score. For example: after looking at Figure 4, it becomes clear that little variation exists between the hospitals (i.e. almost all hospitals are within the 95% control limits), but that two centres may be considered as excellent performers and two centres as under-performers. A difficulty here is the 5% chance of a type I error: for each 20 hospitals, 1 hospital is expected to be outside the 95% control limits (i.e. a false-positive) if in fact the level of quality at all hospitals is according to the benchmark. On the other hand, hospitals inside the control limits may wrongly be assumed to be in-control. Due to the often low patient numbers in funnel plots, the power can be low, meaning that there is a small chance of detecting existing differences in performance.[16] Assuming that hospitals are in-control based on under-powered comparisons is a common misconception (conform the well-known expression "absence of evidence is not evidence of absence"). Therefore, risks of unfairly criticising hospitals or missing under-performers must be weighed and results should be interpreted with caution.[2, 16] More conservative methods such as 99.8% control limits can also be used, hereby yielding fewer false-positives but also less power. Besides this, it may be advisable to monitor the hospital performances over a longer period of time or to pool data over similar groups of patients to explore whether differences in outcomes persist.

An advantage of presenting hospital comparisons in funnel plots is that funnel plots do not involve ordering or ranking of hospitals.[6] In a funnel plot, the hospitals' outcomes (i.e. positions in the funnel plot) remain independent from each other – in contrast to a ranking list or league table, a change in outcome in one hospital does not influence the position of another hospital in a funnel plot.[6] Furthermore, with a funnel plot, one is less inclined to make direct comparisons between hospitals. This is important, because outcomes of individual hospitals are unsuitable for between-hospital comparisons due to the underlying method of indirect standardisation using populations unique to each hospital (see also Box 3).[6]

Relationship with volume

Funnel plots clearly visualize the relation between sample size and precision: the control limits and the distribution of hospital outcomes become smaller with higher volume (i.e. number of patients).[6, 9] The presentation of volume on the x-axis also provides the opportunity to observe an association between volume and outcome (see Figure 3), which is particularly interesting when the outcome is expected to be partly dependent on hospital-volume, for instance when volume is a proxy for experience with certain treatment that may lead to better outcomes.[6, 22]

High and consistent response rates are also necessary to investigate volume effects: if response rates vary highly across hospitals, the sample size (i.e. number of responders presented on the x-axis) is not a good representation of volume (see also Box 1 for other consequences). However, if a fixed number of patients is invited and included in the analysis (e.g. 100 consecutive patients per hospital), the number of responders is equal to the response rate and thus, can be used to explore the association between response rate and outcome. A relationship between response rates and outcomes could be informative, for example when response rates are considered a proxy for certain structures or processes of care organization that may influence the outcome (assuming adequate adjustment for case mix). For example, digitization in hospitals can ease recruitment and may also improve outcomes.[23]

PROs to evaluate quality of care

When using funnel plots for PROs, the following aspects related to the selection, measurement and analysis of PROs should be taken into account.

First, the purpose of healthcare quality evaluation must be taken into account when selecting PROs. It is possible that a PRO is very important for use at the individual level (e.g. during consultations), but that it is not suitable for comparing healthcare quality. To evaluate healthcare quality, PROs should be selected for which an association with healthcare quality is plausible or established. To make relevant comparisons, there must also be room for improvement (i.e. variation across hospitals) and actionable care plans must exist. Umeukeje et al. (2020) provide an example where pain is considered not to be included as performance-indicator in dialysis patients because pain management strategies are lacking and there is too little room for improvement (90% of dialysis centres had the highest score possible).[3] Hence, although pain is a relevant PRO for routine care, in this example, pain seems unsuitable for healthcare quality evaluation.

Second, PRO measurement can be more challenging compared to clinical out-

comes. PROs can only be observed and registered by the patients themselves, making it more difficult to obtain complete data at fixed time-points. Hospital recruitment strategies can also vary and influence patient participation, resulting in selective response and differences in response rates across hospitals (see Box 1). In nephrology, deciding on the right timing to collect PRO outcomes may also be challenging since there is often no clear starting point in chronic care (e.g. prevalent dialysis patients) and because outcomes are likely to vary over time (in contrast to dichotomous outcomes such mortality). Furthermore, the usability of PRO-data is partly determined by the selected PROM (i.e. the questionnaire used to measure the PRO): the psychometric properties of the PROM determines the suitability of the PRO for quality purposes. The PROM must be valid and reliable within the context of the field, and must be responsive to change in such way that differences in healthcare quality can be detected over time or between similar patients receiving different quality of care.[18] Additionally, all hospitals should use the same PROM to measure the same PRO, as different instruments often cannot be easily compared due to differences between questionnaires (e.g. different scales, items or domains).

Third, adequate case mix correction is required to enable fair comparisons and to draw conclusions about differences in performance. Identifying a sufficient set of case mix factors may be more challenging for PROs compared to clinical outcomes, given the complexity of the constructs (e.g. the multidimensional character of PROs: HRQOL includes various domains; see Box 2).[4, 18] Furthermore, for meaningful comparisons, PRO-data of large numbers of patients is needed to have sufficient power and the data should be representative of the total population of interest. Thus, recruitment strategies that yield high and consistent response rates are needed before valid conclusions can be drawn from funnel plots of PROs. Although the validity of the data strongly depends on the randomness of the (non-) response (i.e. representativeness of the study sample), thresholds of 60-80% have been proposed in the literature as adequate response rates.[24-26] Despite the fact that there are still steps to be taken, there are already some examples in the literature showing that PROs can be of added value in healthcare quality evaluation.[27-30] Although beyond the scope of this review, it is important to note that PROs are also being used in routine care at the individual patient level to provide insight into patients' outcomes, enhance patient-professional communication and shared decision-making, identify patients in need for additional support, and consequently, improve patient outcomes and healthcare quality.[2, 4] Patients and professionals particularly consider the individual use of PROs of great added value and an important reason to complete PROMs.[10] Individual use may therefore be the primary

purpose of collecting PROs in routine care. That being said, we should keep in mind that individual and aggregated use often go together and may strengthen each other, for example: aggregated information is valuable when considering treatment choices and may contribute to shared decision-making (e.g. prognoses on outcomes after treatments).[31] Furthermore, the use at individual level is expected to improve response rates, which in turn results in better quality of aggregated information. Finally, the ultimate aim of collecting PROs is to improve patient outcomes and quality of care, and in order to evaluate whether the use of PROs at individual level indeed results in quality improvements, data on an aggregated level is required[32], for instance by using funnel plots.

Conclusion

PROs are becoming increasingly important in healthcare and should be included in healthcare quality evaluation. A funnel plot is a feasible graphical method for this purpose, as it is easily interpretable and precision is clearly visualized. However, some challenges need to be addressed before using funnel plots for PROs, namely: high and consistent response rates, adequate case mix correction and high-quality PRO measures.

References

1. Porter ME. What is value in health care? *The New England journal of medicine* 2010;363(26):2477-81

2. Black N. Patient reported outcome measures could help transform healthcare. *BMJ (Clinical research ed)* 2013;346:f167-f

3. Umeukeje EM, Nair D, Fissell RB, Cavanaugh KL. Incorporating patient-reported outcomes (PROs) into dialysis policy: Current initiatives, challenges, and opportunities. 2020;33(1):18-25

4. Van Der Wees PJ, Nijhuis-Van Der Sanden MWG, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *The Milbank quarterly* 2014;92(4):754-75

5. Gutacker N, Siciliani L, Moscelli G, Gravelle H. Choice of hospital: Which type of quality matters? *Journal of health economics* 2016;50:230-46

6. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;24(8):1185-202

7. Taylor P. Standardized mortality ratios. *International journal of epidemiology* 2013;42(6):1882-90

8. Hoekstra T, Dekker FW, Cransberg K, Bos WJW, van Buren M, Hemmelder MH: RENINE annual report 2018, Utrecht, Nefrovisie, 2019

9. van Dishoeck AM, Looman CW, van der Wilden-van Lier EC, Mackenbach JP, Steyerberg EW. Displaying random variation in comparing hospital performance. *BMJ Qual Saf* 2011;20(8):651-7

10. van der Willik EM, Hemmelder MH, Bart HAJ, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. *Clinical Kidney Journal* 2021;14(6):1535–44

11. Nimmo A, Bell S, Brunton C, et al. Collection and determinants of patient reported outcome measures in haemodialysis patients in Scotland. *QJM : monthly journal of the Association of Physicians* 2018;111(1):15-21

12. Pagels AA, Stendahl M, Evans M. Patient-reported outcome measures as a new application in the Swedish Renal Registry: health-related quality of life through RAND-36. *Clinical Kidney Journal* 2019

13. Morton RL, Lioufas N, Dansie K, et al. Use of patient-reported outcome measures and patient-reported experience measures in renal units in Australia and New Zealand: A cross-sectional survey study. *Nephrology (Carlton)* 2019

14. van der Willik EM, Meuleman Y, Prantl K, et al. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease - a four-phase mixed methods study. *BMC nephrology* 2019;20(1):344-

15. van Eck van der Sluijs A, Bonenkamp AA, Dekker FW, Abrahams AC, van Jaarsveld BC. Dutch nOcturnal and hoME dialysis Study To Improve Clinical Outcomes (DOMESTICO): rationale and design. *BMC Nephrol* 2019;20(1):361

16. Seaton SE, Barker L, Lingsma HF, Steyerberg EW, Manktelow BN. What is the

probability of detecting poorly performing hospitals using funnel plots? *BMJ quality & safety* 2013;22(10):870-6

17. VanderWeele TJ. Principles of confounder selection. *European journal of epidemiology* 2019;34(3):211-9

18. Dennison CR. The role of patient-reported outcomes in evaluating the quality of oncology care. *Am J Manag Care* 2002;8(18 Suppl):S580-6

19. Oemrawsingh A, van Leeuwen N, Venema E, et al. Value-based healthcare in ischemic stroke care: case-mix adjustment models for clinical and patient-reported outcomes. *BMC medical research methodology* 2019;19(1):229-

20. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton, Vic)* 2017;22(3):228-33

21. Naing NN. Easy way to learn standardization : direct and indirect methods. *The Malaysian journal of medical sciences : MJMS* 2000;7(1):10-5

22. van Groningen JT, Marang-van de Mheen PJ, Henneman D, Beets GL, Wouters MWJM. Surgeon perceived most important factors to achieve the best hospital performance on colorectal cancer surgery: a Dutch modified Delphi method. *BMJ open* 2019;9(9):e025304-e

23. van Poelgeest R, van Groningen JT, Daniels JH, et al. Level of Digitization in Dutch Hospitals and the Lengths of Stay of Patients with Colorectal Cancer. *Journal of medical systems* 2017;41(5):84-

24. Rolfson O, Bohm E, Franklin P, et al. Patient-reported outcome measures in arthroplasty registries Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. *Acta orthopaedica* 2016;87 Suppl 1(Suppl 1):9-23

25. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care* 1998;10(4):311-7

26. Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. *American journal of pharmaceutical education* 2008;72(2):43-

27. Bronserud MM, Iachina M, Green A, Groenvold M, Jakobsen E. Patient reported outcome data as performance indicators in surgically treated lung cancer patients. *Lung Cancer* 2019;130:143-8

28. Varagunam M, Hutchings A, Black N. Do patient-reported outcomes offer a more sensitive method for comparing the outcomes of consultants than mortality? A multilevel analysis of routine data. *BMJ quality & safety* 2015;24(3):195-202

29. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* 2016;34(6):557-65

30. Gibbons E, Black N, Fallowfield L, Newhouse R, Fitzpatrick R. Patient-reported outcome measures and the evaluation of services. In Raine R, Fitzpatrick R, Barratt H, Bevan G, Black N, Boaden R, et al. Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. *Health Serv Deliv Res* 2016;4(16):55-68

31. Damman OC, Jani A, de Jong BA, et al. The use of PROMs and shared deci-

sion-making in medical encounters with patients: An opportunity to deliver value-based health care to patients. *J Eval Clin Pract* 2019;n/a(n/a):1-17

32. Varagunam M, Hutchings A, Neuburger J, Black N. Impact on hospital performance of introducing routine patient reported outcome measures in surgery. *Journal of health services research & policy* 2014;19(2):77-84

Supplemental Material for Chapter 5

Table S1. Characteristics of patients receiving dialysis in Dutch dialysis centres, stratified by participation to PRO measurements.

	Total dialysis population§		Responders^	Non-responders^
Characteristics	n = 2711	Range across dialysis centres#	n = 1388 (51.2%)	n = 1323 (48.8%)
Sex (male), n (%)	1601 (59.1)	40.0 - 85.7	838 (60.4)	763 (57.7)
Age (years), mean (SD)	66.8 (14.6)	54.8 - 73.2	67.3 (14.0)	66.2 (15.1)
SES, n (%)				
Low	1380 (51.4)	20.0 - 91.2	656 (47.6)	724 (55.3)
Middle	756 (28.1)	0.0 - 59.4	422 (30.6)	334 (25.5)
High	551 (20.5)	0.0 - 57.1	299 (21.7)	252 (19.2)
Primary kidney disease, n (%)				
Glomerulonephritis/ sclerosis	295 (10.9)	0.0 - 42.9	154 (11.1)	141 (10.7)
Pyelonephritis	131 (4.8)	0.0 - 28.6	62 (4.5)	69 (5.2)
Polycystic kidney disease	134 (4.9)	0.0 - 14.3	82 (5.9)	52 (3.9)
Hypertension	424 (15.6)	0.0 - 71.4	198 (14.3)	226 (17.1)
Renal vascular disease	290 (10.7)	0.0 - 71.4	174 (12.5)	116 (8.8)
Diabetes mellitus	575 (21.2)	0.0 - 57.1	261 (18.8)	314 (23.7)
Miscellaneous	490 (18.1)	0.0 - 56.0	270 (19.5)	220 (16.6)
Unknown	372 (13.7)	0.0 - 39.2	187 (13.5)	185 (14.0)
Dialysis modality, n (%)				
HD	2354 (86.8)	60.0 - 100.0	1242 (89.5)	1112 (84.1)
PD	357 (13.2)	0.0 - 40.0	146 (10.5)	211 (15.9)
Time on RRT (years), geometric mean (SD)	2.2 (4.1)	0.2 - 4.0	2.0 (4.3)	2.4 (3.8)

§ *Total dialysis population includes all patients receiving dialysis in Dutch dialysis centres that participated in the Dutch registry of PROMs in 2019, i.e. all dialysis centres for which at least one patient completed the PROMs.*

^ *Patients are considered responders if they participated to a PRO measurement at least once. Non-responders were invited at least once, but never participated to a PRO measurement.*

Range in percentage or mean of characteristic across n=38 dialysis centres. N=10 dialysis centres included < 5 patients and were excluded from the calculation of the range.

Abbreviations: SES, social economic status; HD, haemodialysis; PD, peritoneal dialysis; RRT, renal replacement therapy

6

Itching in dialysis patients: impact on health-related quality of life and interactions with sleep problems and psychological symptoms – results from the RENINE/PROMs registry

Esmee M. van der Willik, Robin Lengton, Marc H. Hemmelder, Ellen K. Hoogeveen, Hans A. J. Bart, Frans J. van Ittersum, Marc A.G.J. ten Dam, Willem Jan W. Bos, Friedo W. Dekker, Yvette Meuleman

Nephrology Dialysis Transplantation. 2022;37(9):1731-1741.

Abstract

Background: Itching (pruritus) is common in dialysis patients, but little is known about its impact on health-related quality of life (HRQOL), sleep problems and psychological symptoms. This study investigates the impact of itching in dialysis patients by looking into the persistence of itching, the effect of itching on the course of HRQOL, and the combined effect of itching with sleep problems and with psychological symptoms on HRQOL.

Methods: Data were obtained from the RENINE/PROMs registry and included 2978 dialysis patients who completed patient-reported outcome measures between 2018-2020. Itching, sleep problems and psychological symptoms were assessed with the DSI, and HRQOL with the SF-12. Effects of itching on HRQOL and interactions with sleep problems and psychological symptoms were investigated cross-sectionally and longitudinally, using linear regression and linear mixed models.

Results: Half of the patients experienced itching and in 70% of them, itching was persistent. Itching was associated with a lower physical and mental HRQOL (-3.35 [95%CI: -4.12;-2.59] and -3.79 [95%CI: -4.56;-3.03]). HRQOL remained stable during two years and trajectories did not differ between patients with or without itching. Sleep problems (70% vs 52%) and psychological symptoms (36% vs 19%) were more common in patients with itching. These symptoms had an additional negative effect on HRQOL, but did not interact with itching.

Conclusions: The persistence of itching, its impact on HRQOL over time, and the additional effect on HRQOL of sleep problems and psychological symptoms, emphasize the need for recognition and effective treatment of itching to reduce symptom burden and improve HRQOL.

Introduction

Patients with end-stage kidney disease (ESKD) experience numerous physical and emotional disease-related symptoms, such as fatigue, muscle cramps, itching, sleep problems and depressive symptoms.[1, 2] The heavy symptom burden has a disruptive impact on individuals' lives and has been shown to be associated with the impaired health-related quality of life (HRQOL) in this population.[3, 4]

A common and highly distressing symptom is chronic kidney disease-associated pruritus, better known as itching. Itching is experienced by both hemodialysis (HD) and peritoneal dialysis (PD) patients with a prevalence of approximately 50%.[1, 2, 5, 6] Itching was found to be one of the ten most burdensome symptoms experienced by dialysis patients[2] and is considered a main research priority by patients with ESKD, their caregivers and healthcare professionals.[7] The pathogenesis of itching in dialysis patients is not yet fully understood, but several factors seem to influence the occurrence or burden of itching, including abnormal calcium, phosphate and parathyroid hormone levels, opioid imbalance, peripheral neuropathy, dialysis efficiency and a dry skin.[5, 6] Furthermore, itching has been associated with adverse clinical outcomes, such as hospitalization and mortality, and poor patient-reported outcomes, such as a decreased HRQOL, psychological symptoms (e.g. depressive symptoms) and sleep problems.[5, 6, 8, 9] Large cohort studies have found that HRQOL, depressive symptoms and sleep quality were worse with more severe itching in dialysis patients [8, 9]; these associations suggest a causal effect of itching on HRQOL. However, information about the impact of itching on the course of HRQOL over time is lacking. Moreover, although itching is often accompanied with sleep problems and psychological symptoms, no literature is currently available about the extent to which the combinations of these symptoms affect patients' physical and mental HRQOL.

Insight into the impact of itching on HRQOL and into the combined effect of itching with sleep problems and with psychological symptoms in the association with HRQOL, could help to better understand patients' outcomes and ultimately to reduce symptom burden and increase HRQOL. Therefore, the aim of this study is to investigate the impact of itching in patients receiving dialysis treatment by looking into the persistence of itching over time, the relationship between itching and HRQOL, and the combined effect of itching with sleep problems and psychological symptoms on HRQOL. These associations will be examined both cross-sectionally and longitudinally, using data from routine Dutch dialysis care.

Methods

Study design and population

Data were obtained from RENINE (Registratie Nierfunctievervanging Nederland: www.nefrovisie.nl/renine), the nationwide Dutch renal registry of patients receiving kidney replacement therapy. The registry collects information on demographics and clinical characteristics which are registered every 3 months. In addition, patient-reported outcome measures (PROMs) were introduced into routine dialysis care in September 2016 as part of a pilot study in 16 Dutch dialysis centres and have now been implemented nationally since November 2018.[2] The PROMs were selected in collaboration with patients and experts[10], and include the 12-item Short Form Health Survey (SF-12)[11] to assess HRQOL and the Dialysis Symptom Index (DSI)[12] to assess symptom burden. PROMs-invitations are distributed 1-2 times per year in all patients receiving dialysis treatment (i.e. total prevalent dialysis population). To ensure inclusion of all hospitals and a consistent follow-up period, data from 2018-2020 were used for this study. All dialysis patients that completed the PROMs at least once in this period were included in the analysis.

All patients included in RENINE gave consent to collect and use their data for scientific research purposes. Additionally, the current study protocol was reviewed and approved by the scientific committees of Nefrovisie and of the clinical epidemiology department at Leiden University Medical Centre (LUMC). The study is reported according to the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines[13] with the extension of the REporting of studies Conducted using Observational Routinely collected health Data (RECORD) statement[14].

Itching, sleep problems and psychological symptoms

The DSI is a 30-item kidney disease specific questionnaire to assess physical and emotional symptom burden. Patients were asked to report the presence of 30 symptoms (yes/no) during the past week and, if present, the burden of each symptom on a 5-point Likert scale ranging from 1 'not at all' to 5 'very much' bothersome.[12] Two overall scores were calculated: 1) the total number of symptoms present (0-30 symptoms), and 2) the total symptom burden score (score range 0-150), which is the sum of burden on individual symptoms whereby missing items were assumed absent (i.e. burden score: 0).[2, 15]

The symptoms of interest in this study – itching, sleep problems and psychological symptoms – were assessed by means of the DSI. Itching was reported using a single item assessing whether itching was experienced in the past week and, if present, how bothersome this was. For the main analysis, patients were stratified

based on the presence of itching (yes/no) at baseline (i.e. the patient's first PROM measurement). The burden score of itching ranges from 0 to 5, with higher scores indicating a higher burden.
Sleep problems were assessed using two symptoms, namely 'trouble falling asleep' and 'trouble staying asleep'. Sleep problems were defined as at least one of these two symptoms being present. The burden score of sleep problems ranges from 0 to 10, with higher scores indicating a higher burden.
The psychological cluster includes the following five symptoms: 'worrying', 'feeling nervous', 'feeling irritable', 'feeling sad' and 'feeling anxious'. Psychological symptoms were considered present when at least three out of these five symptoms were experienced by the patient. The total burden score of psychological symptoms ranges from 0 to 25, with higher scores indicating a higher burden.

Health-related quality of life

The SF-12 is a generic health questionnaire consisting of 12 questions assessing the following 8 domains of HRQOL: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional and mental health. These domains contribute (in different proportions) to the scoring of a physical component summary (hereafter referred to as 'physical HRQOL') and a mental component summary (hereafter referred to as 'mental HRQOL'). HRQOL scores range from 0 to 100, with higher scores indicating a better physical and mental HRQOL.[11]

Population characteristics

Demographics and clinical characteristics were: age, sex, primary kidney disease (PKD) according to European Renal Association codes[16], socio-economic status (SES, classified as low, middle and high using standard deviation [SD] scores based on zip code), dialysis modality (haemodialysis [HD] or peritoneal dialysis [PD]), number of dialysis sessions and hours per week in HD, time since dialysis initiation, kidney transplantation in the past (yes/no), residual glomerular filtration rate (rGFR, ml/min/1.73m^2), single pool Kt/V per dialysis session in HD patients, total Kt/V per week in PD patients, haemoglobin (mmol/L), ferritin (µg/L), transferrin saturation (%), albumin-adjusted calcium (mmol/L), phosphate (mmol/L), parathyroid hormone (pmol/L).

Statistical analysis

Baseline was defined as the patient's first PROMs-measurement. Baseline characteristics of the dialysis population, stratified for the experience of itching (yes/

no) are presented as frequencies with percentages for categorical variables, as mean with SD for normally distributed continuous variables and as median with interquartile range (IQR) for skewed continuous variables. The prevalence and persistence of itching over time are shown graphically based on calendar time (prevalence of itching over time) and on patients' follow-up time stratified for itching at baseline (persistence of itching over time).

The main analyses were performed both cross-sectionally and longitudinally, so that all patients and all PROMs-measurements could be included in the analyses, and to expand on existing (mainly cross-sectional) literature.[5] The cross-sectional analysis was performed at baseline and includes all patients in the study population (n=2978 patients). The longitudinal analysis includes all PROMs-measurements (n=5042), with 1218 (40.9%) patients having multiple PROMs-measurements. Of the individuals that had only one PROMs-measurement (n=1760), 1032 (58.6%) patients started with PROMs in 2020 and 322 (18.3%) patients died, which prevented follow-up data being available. The main analyses were performed crude and adjusted for the following potential baseline confounders: age, sex, PKD, SES, dialysis modality, time since dialysis initiation and kidney transplantation in past. The symptoms, HRQOL and potential confounding variables included in the main analyses had no or less than two percent missing values.

Cross-sectionally, the association between the presence of itching and physical and mental HRQOL was investigated using linear regression analysis. Furthermore, in two separate linear regression models, a cross-product interaction term for itching (yes/no) and sleep problems (yes/no) and for itching and psychological symptoms (yes/no) was included to assess the interaction effects of these symptoms in the association with HRQOL.

The associations described above were also investigated longitudinally using linear mixed models. By using this statistical method, all measurements from all individuals could be included, as the model takes account of a varying number of follow-up measurements across individuals and even single measurements can be included in the estimation of the trajectory over time at population level.[17] The presence of itching at baseline was included in the model as fixed independent variable, time as random variable, and the continuous physical and mental HRQOL over time as dependent variable. The interaction between time and itching was included, indicating the annual change in HRQOL for individuals with itching compared to individuals without itching.

Sensitivity analyses were conducted to assess the robustness of our main results. Both the cross-sectional and the longitudinal analyses were repeated using the continuous burden score of itching, sleep problems and psychological symptoms.

The analyses were also performed with the symptoms classified based on low or high burden, e.g. no or mild itching (burden score: 0-2) versus moderate to severe itching (burden score: 3-5), and similar categories for sleep problems and psychological symptoms. Furthermore, the analyses were repeated comparing persistent itching (i.e. presence of itching reported both at baseline and at the first follow-up measurement) with no or non-persistent itching. Last, analyses were performed using 2019-2020 data, to only include measurements from the official start of the PROMs registry at November 2018.
All statistical analyses were performed using SPSS version 25.0 (IBM, Armonk, NY, USA).

Results

Population characteristics

Table 1 presents the characteristics of all patients (n=2978) that completed the PROMs at least once in 2018-2020, stratified for the presence of itching at baseline. Itching was present in approximately half of the patients and was more common in individuals receiving PD (59.4%) compared to HD (48.7%). (See also Supplement A for the population characteristics stratified by dialysis modality). Patients with itching were more often male, had a higher SES and more often diabetes as primary kidney disease, compared to patients without itching. No differences were observed in calcium, phosphate and parathyroid hormone levels. The total symptom burden was higher in patients who experienced itching, with on average 14 symptoms with a median (IQR) total burden score of 35 (23-51), compared to 8 symptoms with a median (IQR) total burden score of 19 (10-32) in patients who did not experience itching. Patients with itching had more often a dry skin compared to patients without itching (73% versus 43%, resp.). Sleep problems were experienced by 70% of the patients with itching and by 52% of the patients without itching. Psychological symptoms occurred in 36% of the patients with itching compared to 19% in patients without itching.

Table 1. Characteristics of dialysis patients, stratified by presence of itching (yes/no) at baseline

	Total dialysis population (n=2978)	Patients with itching (n=1493, 50.1%)	Patients without itching (n=1485, 49.9%)
Age (years)	67.3 (14.1)	67.4 (14.0)	67.3 (14.2)
Sex (male)	1827 (61.4)	927 (62.1)	900 (60.7)
SES			
Low	1430 (48.4)	711 (48.0)	719 (48.8)
Middle	907 (30.7)	435 (29.4)	472 (32.1)
High	617 (20.9)	336 (22.7)	281 (19.1)
Primary kidney disease			
Glomerulonephritis/sclerosis	333 (11.2)	160 (10.7)	173 (11.7)
Pyelonephritis	140 (4.7)	68 (4.6)	72 (4.9)
Polycystic kidney disease	171 (5.7)	86 (5.8)	85 (5.7)
Hypertension/renal vascular disease	809 (27.2)	411 (27.5)	398 (26.8)
Diabetes mellitus type 1/2	601 (20.2)	320 (21.4)	281 (18.9)
Miscellaneous	535 (18.0)	257 (17.2)	278 (18.7)
Unknown	387 (13.0)	191 (12.8)	196 (13.2)
Dialysis modality			
HD	2583 (87.9)	1258 (85.6)	1325 (90.1)
PD	357 (12.1)	212 (14.4)	145 (9.9)
Dialysis sessions per week (HD)			
< 3	266 (12.8)	125 (12.3)	141 (13.2)
3	1672 (80.3)	817 (80.6)	855 (80.1)
> 3	144 (6.9)	72 (7.1)	72 (6.7)
Dialysis hours per week (HD)	11.2 (4.3)	11.1 (3.8)	11.4 (4.8)
Time since dialysis initiation (months)	15 (3-43)	14 (3-41)	17 (3-46)
Kidney transplantation in past (yes)	327 (11.2)	162 (11.1)	165 (11.3)
rGFR (mL/min/1.73m^2)	4.7 (2.0-7.6)	5.0 (2.0-8.0)	4.5 (2.1-7.1)
Single pool Kt/V in HD	1.47 (0.53)	1.46 (0.54)	1.48 (0.52)
Total Kt/V in PD	2.63 (1.06)	2.70 (1.10)	2.52 (0.99)
Haemoglobin (mmol/L)	6.8 (0.9)	6.8 (0.9)	6.8 (0.9)

Ferritin (µg/L)	318 (168-534)	300 (153-517)	330 (187-547)
Transferrin saturation (%)	22.3 (10.5)	21.8 (10.4)	22.8 (10.6)
Calcium (mmol/L)[#]	2.31 (0.19)	2.31 (0.19)	2.31 (0.18)
Phosphate (mmol/L)	1.59 (0.48)	1.62 (0.49)	1.57 (0.47)
Parathyroid hormone (pmol/L)	30 (17-51)	30 (17-52)	30 (17-50)
Symptom burden			
Total number of symptoms (0-30)	11.0 (6.4)	13.7 (6.2)	8.3 (5.3)
Total symptom burden score (0-150)	27 (14-42)	35 (23-51)	19 (10-32)
Dry skin (yes)	1726 (58.0)	1091 (73.1)	635 (42.8)
Sleep problems			
Sleep problems (yes)[^]	1816 (61.0)	1044 (69.9)	772 (52.0)
Trouble falling asleep (yes)	1312 (44.1)	798 (53.4)	514 (34.6)
Trouble staying asleep (yes)	1549 (52.0)	907 (60.8)	642 (43.2)
Psychological symptoms			
Psychological symptoms (yes)[$]	820 (27.5)	534 (35.8)	286 (19.3)
Worrying (yes)	1195 (40.1)	708 (47.4)	487 (32.8)
Feeling nervous (yes)	816 (27.4)	517 (34.6)	299 (20.1)
Feeling irritable (yes)	843 (28.3)	558 (37.4)	285 (19.2)
Feeling sad (yes)	1088 (36.5)	669 (44.8)	419 (28.2)
Feeling anxious (yes)	647 (21.7)	425 (28.5)	222 (14.9)

Values are shown in n (%), mean (SD) or median (IQR).
Missing values for age: n=3 (0.10%), sex: n=2 (0.07%), SES: n=24 (0.81%), primary kidney disease: n=2 (0.07%), dialysis modality: n=38 (1.29%), dialysis sessions per week (HD): n=501 (19.4%), dialysis hours per week (HD): n=505 (19.6%), time since dialysis initiation: n=58 (1.99%), kidney transplantation in past: n=58 (1.99%), residual GFR: n=1977 (66.4%), single pool Kt/V in HD: n=647 (25.0%), Total Kt/V in PD: n=221 (61.9%), haemoglobin: n=286 (9.60%), ferritin: n=416 (14.0%), transferrin saturation: n=993 (33.3%) calcium: n=369 (12.4%), phosphate: n=280 (9.40%), parathyroid hormone: n=522 (17.5%).
[#] Albumin-adjusted calcium.
[^] Sleep problems are considered present if at least one of the two symptoms are experienced by the patient.
[$] Psychological symptoms are considered present if three out of the five symptoms are experienced by the patient.
Abbreviations: SES, social economic status; HD, haemodialysis; PD, peritoneal dialysis; rGFR, residual glomerular filtration rate.

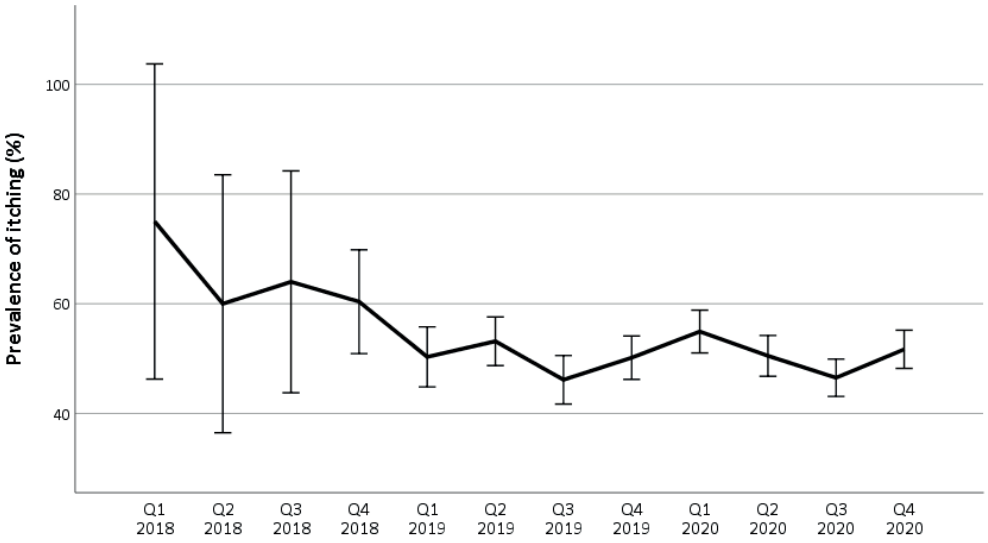


Figure 1a. Prevalence of itching (yes/no) over study period (calendar time).
Percentage of dialysis patients who experience itching (solid line) with 95% confidence intervals (bars) at each quartile in 2018-2020. Note that Q1-Q3 2018 includes a small number of patients (n=12, n=20, n=25, resp.), as the PROMs registry officially started from November 2018. Some patients already participated in Q1-Q3 2018 for scientific research purposes.

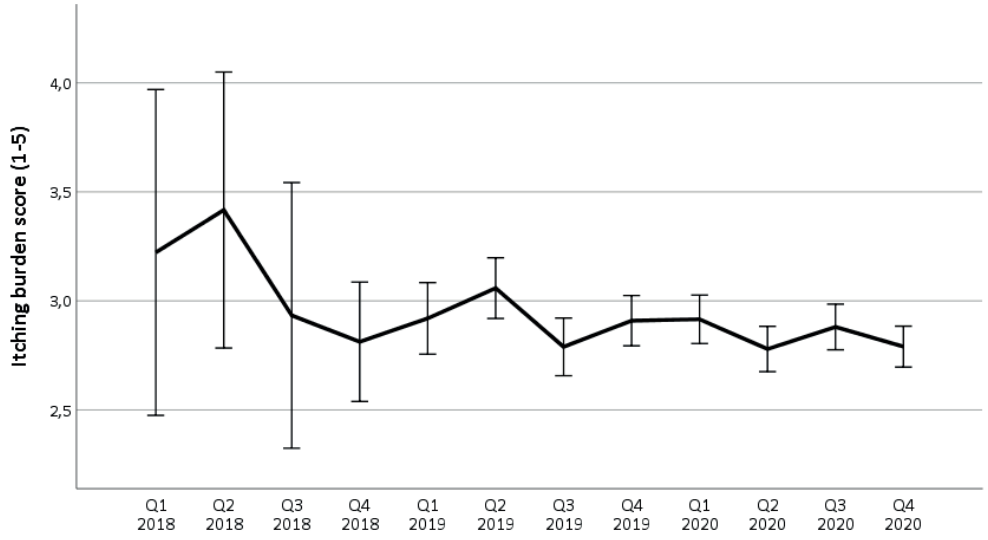


Figure 1b. Burden of itching (scale: 1-5) over study period (calendar time) in patients who experienced itching.
Average itching burden score (solid line) with 95% confidence intervals (bars) in dialysis patients who experienced itching at each quartile in 2018-2020. Note that Q1-Q3 2018 includes a small number of patients (n=12, n=20, n=25, resp.), as the PROMs registry officially started from November 2018. Some patients already participated in Q1-Q3 2018 for scientific research purposes.

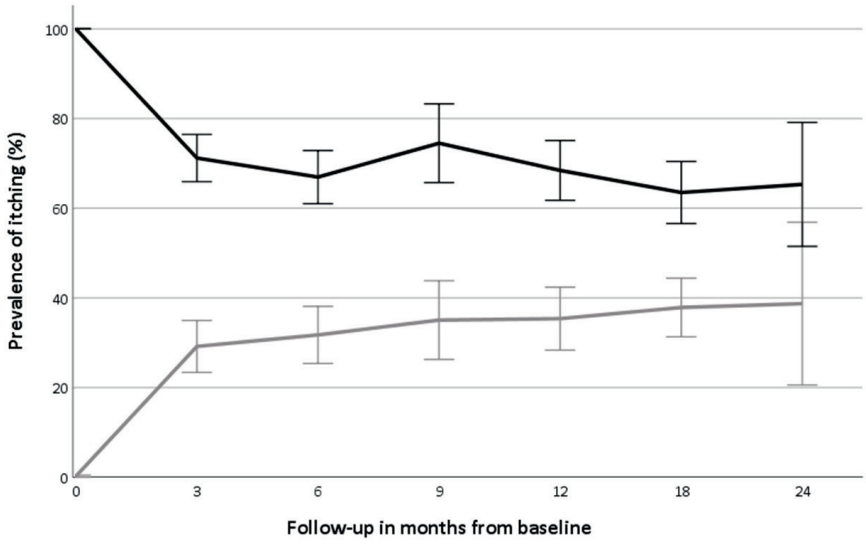


Figure 2. Persistence of itching during follow-up in patients with itching (black) and patients without itching (grey) at baseline.

Black solid line (black bars) shows the percentage (95% Confidence Interval [CI]) of dialysis patients in which itching is persistent during follow-up since baseline. Grey solid line (grey bars) shows the percentage (95% CI) of dialysis patients in which itching was newly developed during follow-up since baseline. Note that the average time between follow-up measurements was 6.7 months, meaning that the number of patients that contribute data fluctuates across the time-points in the graph.

Prevalence and persistence of itching over time

In total, 1218 patients have multiple PROMs-measurements (median: 2, IQR: 2-3 measurements), with on average 6.7 months (SD: 5.0) between baseline and the second PROMs-measurement. Throughout the whole study period, the prevalence of itching is around 50% with a moderate burden (mean burden scores between 2.8 and 3.4 on 1-5 scale) (Figure 1a/b). No clear differences in prevalence or burden of itching could be detected between the yearly quartiles (i.e. no seasonal effects). Figure 2 shows that itching persisted over time in approximately 70% of the patients that experienced itching at baseline. Of the patients without itching at baseline, 30-40% developed itching during follow-up. Sleep problems and psychological symptoms also persisted over time in the majority of the patients (see Supplement B).

Table 2. Cross-sectional effects of the presence of itching (yes/no), combined with sleep problems and psychological symptoms, on physical and mental HRQOL

	Physical HRQOL		Mental HRQOL	
	Coef. (95%CI)	p-value	Coef. (95%CI)	p-value
Itching				
Model 1, unadjusted	-3.36 (-4.13; -2.59)	<0.001	-3.82 (-4.58; -3.06)	<0.001
Model 2, adjusted^	-3.35 (-4.12; -2.59)	<0.001	-3.79 (-4.56; -3.03)	<0.001
Itching and sleep problems (Model 3$)				
Itching	-3.38 (-4.62; -2.13)	<0.001	-2.38 (-3.61; -1.15)	<0.001
Sleep problems	-3.85 (-4.92; -2.78)	<0.001	-3.37 (-4.42; -2.31)	<0.001
Itching * sleep problems	1.00 (-0.58; 2.58)	0.214#	-1.18 (-2.74; 0.38)	0.139#
Itching and psychological symptoms (Model 4$)				
Itching	-2.81 (-3.70; -1.92)	<0.001	-2.35 (-3.15; -1.56)	<0.001
Psychological symptoms	-3.51 (-4.87; -2.15)	<0.001	-11.34 (-12.56; -10.13)	<0.001
Itching * psychological symptoms	0.03 (-1.72; 1.79)	0.971#	1.01 (-0.56; 2.58)	0.208#

^ *Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.*
$ *Model 3 and 4 build on model 2 and include the interaction with sleep problems and psychological symptoms, respectively.*
# *P-value for interaction.*

Association between itching and HRQOL at baseline

Mean (SD) physical and mental HRQOL scores in the total dialysis population were 35.8 (10.4) and 48.1 (10.4), respectively. Table 2 shows the cross-sectional effects of itching, combined with sleep problems and psychological symptoms, on physical and mental HRQOL. Patients with itching experienced a lower physical (-3.35 [95% Confidence Interval (CI): -4.12 to -2.59; p<0.001]) and mental HRQOL (-3.79 [95% CI: -4.56 to -3.03; p<0.001]), compared to patients without itching. Sleep problems and psychological symptoms had an additional negative effect on HRQOL. No interaction was observed between itching and sleep problems or psychological symptoms in the association with HRQOL. Table 3a shows the average physical and mental HRQOL in patients with itching, sleep problems, or a combination of both. Table 3b shows the average physical and mental HRQOL in patients with itching, psychological symptoms, or a combination of both.

Table 3a. Physical and mental HRQOL in patients with itching, sleep problems, or a combination of both.

		Physical HRQOL^		Mental HRQOL^	
Itching	**Sleep problems**	Mean	SD	Mean	SD
no	no	39.44	2.23	51.89	1.79
yes	no	36.01	2.03	49.68	1.69
no	yes	35.58	2.21	48.47	1.83
yes	yes	33.25	2.22	44.89	1.81

^ Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.

Table 3b. Physical and mental HRQOL in patients with itching, psychological symptoms, or a combination of both.

		Physical HRQOL^		Mental HRQOL^	
Itching	**Psychological symptoms**	Mean	SD	Mean	SD
no	no	38.12	2.15	52.34	1.46
yes	no	35.33	2.03	50.02	1.40
no	yes	34.51	2.17	40.80	1.54
yes	yes	31.75	2.19	39.46	1.47

^ Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.

Association between itching and HRQOL over time

Figures 3a and 3b show the trajectories of physical and mental HRQOL during follow-up, stratified by itching at baseline. Findings from the longitudinal analyses using linear mixed models were similar to the cross-sectional analyses, showing that patients with itching experienced a lower physical and mental HRQOL compared to patients without itching (-3.12 [95% CI: -3.86 to -2.38; p<0.001] and -3.62 [95% CI: -4.35 to -2.88; p<0.001], resp.). No changes in physical and mental HRQOL over time were observed in the total population throughout follow-up (annual change: 0.01 [95% CI: -0.68 to 0.70; p=0.97] and -0.04 [95% CI: -0.75 to 0.67; p=0.91], resp.). No differences in physical and mental HRQOL trajectories were observed between patients with and without itching (extra annual change in patients with itching: 0.10 [95% CI: -0.83 to 1.03; p=0.83] and -0.07 [95% CI: -1.04 to 0.90; p=0.88], resp.). Also longitudinally, in the association with physical and mental HRQOL, there was no significant interaction between itching and sleep problems (p=0.52 and p=0.22, resp.) or itching and psychological symptoms (p=0.66 and p=0.29, resp.).

A post hoc subgroup analysis showed an increase in physical and mental HRQOL when itching disappeared (+0.56; p=0.49 and +1.78; p=0.02, resp.) and a decrease when itching newly occurred (-0.44; p=0.61 and -0.68; p=0.38, resp.) between the patients' first and second PROMs-measurement (see Supplement C).

Sensitivity analyses

All sensitivity analyses yielded results comparable to the main analyses, both cross-sectionally and longitudinally (see Supplement D). Analyses using the continuous burden scores for symptoms showed that physical and mental HRQOL were -1.26 (95% CI: -1.50 to -1.02; p<0.001) and -1.42 (95% CI: -1.65 to -1.18; p<0.001) points lower, respectively, for each point increase in burden of itching. Using the continuous burden scores, the interaction between itching and psychological symptoms in the association with physical and mental HRQOL became statistically significant, though with a similarly small effect. Moderate to severe itching (prevalence: 26.1%) compared to no or mild itching showed a larger decrease in physical and mental HRQOL (-4.20 [95% CI: -5.07 to -3.33; p<0.001] and -4.90 [95% CI: -5.76 to -4.03; p<0.001], resp.). Comparing persistent itching to no or non-persistent itching showed comparable results. Restriction to 2019-2020 data yielded similar results.

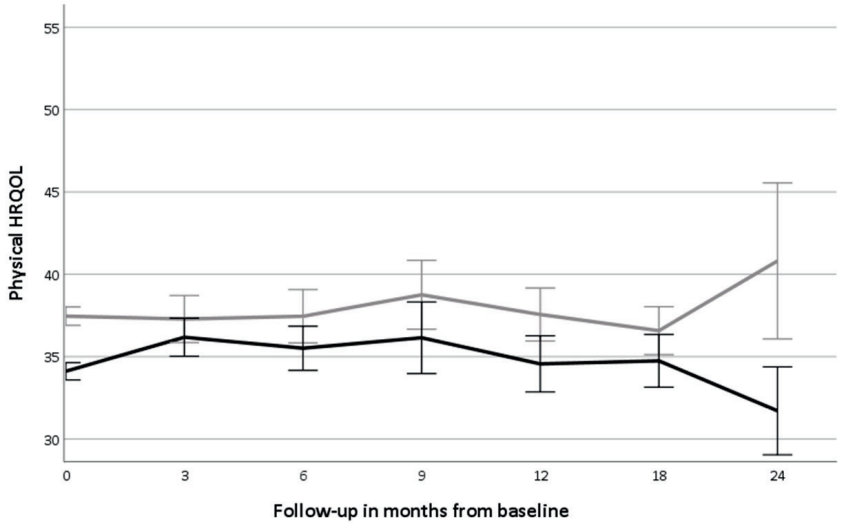


Figure 3a. Trajectory of physical HRQOL during follow-up in patients with itching (black) and patients without itching (grey) at baseline.

Black solid line (black bars) shows the mean physical HRQOL (95% Confidence Interval [CI]) over time in dialysis patients with itching at baseline. Grey solid line (grey bars) shows the mean physical HRQOL (95% CI) over time in dialysis patients without itching at baseline. Note that the average time between follow-up measurements was 6.7 months, meaning that the number of patients that contribute data fluctuates across the timepoints in the graph.

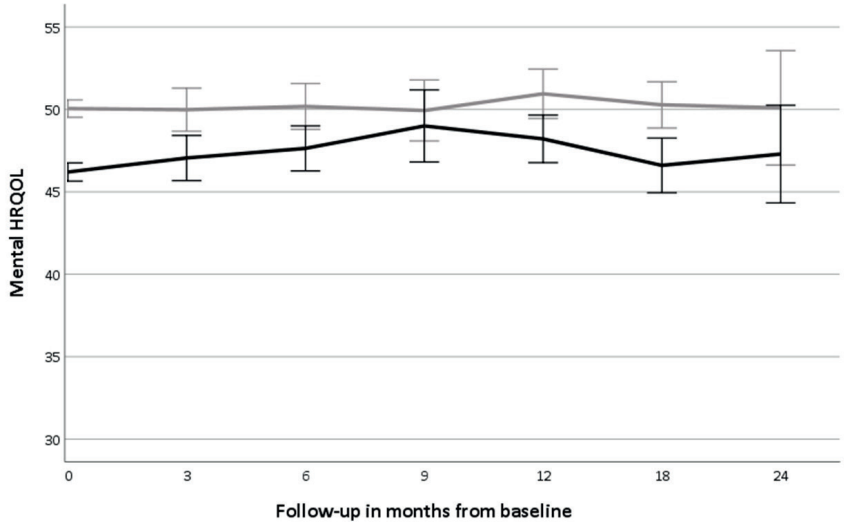


Figure 3b. Trajectory of mental HRQOL during follow-up in patients with itching (black) and patients without itching (grey) at baseline.

Black solid line (black bars) shows the mean mental HRQOL (95% Confidence Interval [CI]) over time in dialysis patients with itching at baseline. Grey solid line (grey bars) shows the mean mental HRQOL (95% CI) over time in dialysis patients without itching at baseline. Note that the average time between follow-up measurements was 6.7 months, meaning that the number of patients that contribute data fluctuates across the timepoints in the graph.

Discussion

This nationwide Dutch study investigated the impact of itching (pruritus) on HRQOL and interactions with sleep problems and psychological symptoms in patients receiving dialysis treatment. Half of the dialysis patients experienced itching and in 70% of them, itching was persistent over time. Individuals with itching experienced a lower physical and mental HRQOL. This is the first study showing that HRQOL remained stable during the two years of follow-up and HRQOL trajectories did not differ between patients with or without itching. Furthermore, we found that sleep problems and psychological symptoms were more common in individuals who also experienced itching. These symptoms had an additional negative effect on physical and mental HRQOL, but did not interact with itching (i.e. the combination of both symptoms did not result in a significantly lower or higher HRQOL than the sum of individual effects).

The high prevalence of itching and its persistence over time demonstrate that itching is a major problem in patients receiving dialysis treatment. Although the estimated prevalence varies between 20-90% across studies[5], for instance due to differences in populations and in definitions of itching, it is clear that itching affects many dialysis patients' lives, especially given that itching appeared to be persistent in many patients. In line with our results, two other studies found that itching was persistent for >1 year in 50-69% of the dialysis patients.[18, 19] The reason why itching was persistent in some patients and not in others, remained unclear. According to one of the studies, differences could not be explained by whether or not the patients received treatment for their itching, as itching was often underestimated and left untreated.[18] Therefore, more research is needed to identify patients with persistent itching in order to treat them in a timely manner.

The impact of itching on patients' lives is clearly reflected in the decreased HRQOL scores. In line with existing literature, physical and mental HRQOL were around three to four points lower in patients who experienced itching compared to those without itching, and HRQOL scores decreased further with more severe itching.[8, 9, 20-22] Information regarding the relevance of this difference according to dialysis patients is lacking[23], but comparable differences in HRQOL have been considered important in other populations.[24-26] In addition to the existing literature, this study also investigated the impact of itching on the course of HRQOL and showed that the difference in HRQOL between individuals with and without itching persisted over time. A possible explanation for this result is that itching also persisted over time in the majority of the patients. However, in contrast to what might be expected based on previous research about the effect of itching on clinical outcomes over time (e.g. mortality and hospitalizations)[5, 8, 27, 28], this study showed no faster

deterioration of HRQOL in patients with (persistent) itching during two years of follow-up. Future research should investigate these relationships using a longer follow-up period.

Findings from our study confirm that sleep problems and psychological symptoms often co-occur with itching.[6, 8, 29-31] Results from previous studies suggest that sleep problems and psychological symptoms may partly explain the effect of itching on HRQOL.[5, 20, 30] Our study does not contradict this suggestion, but it does show that sleep problems and psychological symptoms also independently affect HRQOL, in addition to the effect that itching has on HRQOL. Since these symptoms often co-occur in dialysis patients, many individuals have to deal with a substantially decreased HRQOL.

Our findings emphasize the importance of an effective treatment for itching in dialysis patients. Although unadjusted for potential confounding, findings from two observational studies suggest that HRQOL improves when itching disappears.[18, 19] A post hoc analysis in our data also showed an improved physical HRQOL ($p<0.05$) and mental HRQOL (p=ns) when itching disappeared. Additionally, several treatment trials showed that a reduced itching intensity may already result in improved HRQOL scores and sleep quality.[5] Furthermore, literature suggests that a better management of itching and HRQOL might even result in improved clinical outcomes, such as mortality and hospitalizations.[5, 8, 27, 28] The need for and the potential benefits of a treatment for itching are thus evident. However, effective treatment of itching in dialysis patients appeared challenging: some treatment options are available (e.g. prevention of hyperphosphatemia, adequate dialysis dose, ultraviolet-B light therapy, gabapentin and several emollient creams), but seem to have limited efficacy or side effects.[5, 6, 32]

With this study, we aim to provide insight in and awareness of the high prevalence and impact of itching, as this may still be underestimated.[33, 34] We believe that our research can contribute to existing knowledge in particular due to the longitudinal design using national data from routine dialysis care. As PROMs are part of and used for routine care, patients are more likely to participate (compared to research purposes only), which enhances the generalizability of our results. On the downside, due to this design the study mainly includes prevalent dialysis patients, which means that patients have been followed from different points in their trajectory (e.g. differences in time since start of dialysis). However, we do not believe this has affected the relationship between itching and HRQOL. Another limitation of this study is that no information was available on treatments that may have induced or reduced itching. It is therefore unclear how treatment may have affected the results (e.g. is the prevalence and persistence of itching this high despite treat-

ment of itching?) or to what extent available treatment options may decrease the burden of itching in dialysis patients. Furthermore, additional knowledge about factors that may influence itching is needed, and may be informative for treatment choices, for example to tailor the dialysis schedule or nutritional advice. Taken together, current findings show that itching is a major problem in dialysis patients and call for further research to effectively identify and treat (persistent) itching to reduce symptom burden and improve HRQOL.

Of course, to reduce the burden of itching in dialysis patients, attention must be paid to itching on the individual patient level. Literature suggests however that itching remains underreported and therefore undertreated due to a lack of knowledge and assessment during consultations.[34] We believe that the use of PROMs in routine dialysis care improves the reporting, and prompts discussion of patients' experiences and treatment options.[2] Current findings can be used to better inform patients and may enhance shared decision making.

In conclusion, the high prevalence and persistence of itching, its impact on HRQOL over time, and the additional effect on HRQOL of the often co-occurring sleep problems and psychological symptoms, emphasize the need for recognition and effective treatment of itching to reduce symptom burden and improve HRQOL in dialysis patients. No individual prognoses can be derived from our study, but the findings may be used in shared decision making. We hope that this study provided insight into and awareness of the major impact that itching can have, to enable early recognition and treatment of itching.

References

1. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

2. van der Willik EM, Hemmelder MH, Bart HAJ, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. *Clinical Kidney Journal* 2021;14(6):1535-44

3. Voskamp PWM, van Diepen M, Evans M, et al. The impact of symptoms on health-related quality of life in elderly pre-dialysis patients: effect and importance in the EQUAL study. *Nephrol Dial Transplant* 2019;34(10):1707-15

4. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

5. Shirazian S, Aina O, Park Y, et al. Chronic kidney disease-associated pruritus: impact on quality of life and current management challenges. *Int J Nephrol Renovasc Dis* 2017;10:11-26

6. Verduzco HA, Shirazian S. CKD-Associated Pruritus: New Insights Into Diagnosis, Pathogenesis, and Management. *Kidney Int Rep* 2020;5(9):1387-402

7. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. *Clin J Am Soc Nephrol* 2014;9(10):1813-21

8. Sukul N, Karaboyas A, Csomor PA, et al. Self-reported Pruritus and Clinical, Dialysis-Related, and Patient-Reported Outcomes in Hemodialysis Patients. *Kidney Med* 2021;3(1):42-53.e1

9. Ramakrishnan K, Bond TC, Claxton A, et al. Clinical characteristics and outcomes of end-stage renal disease patients with self-reported pruritus symptoms. *Int J Nephrol Renovasc Dis* 2013;7:1-12

10. van der Willik EM, Meuleman Y, Prantl K, et al. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease - a four-phase mixed methods study. *BMC Nephrol* 2019;20(1):344

11. Ware Jr JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care* 1996:220-33

12. Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *J Pain Symptom Manage* 2004;27(3):226-40

13. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61(4):344-9

14. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12(10):e1001885

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

16. ERA-EDTA Registry Annual Report 2019, Amsterdam UMC, location AMC, Department of Medical Informatics, Amsterdam, the Netherlands, ERA-EDTA Registry, 2021

17. Janmaat CJ, van Diepen M, Tsonaka R, Jager KJ, Zoccali C, Dekker FW. Pitfalls of linear regression for estimating slopes over time and how to avoid them by using linear mixed-effects models. *Nephrol Dial Transplant* 2019;34(4):561-6

18. Plewig N, Ofenloch R, Mettang T, Weisshaar E. The course of chronic itch in hemodialysis patients: results of a 4-year follow-up study of GEHIS (German Epidemiological Hemodialysis Itch Study). *J Eur Acad Dermatol Venereol* 2019;33(7):1429-35

19. Mathur VS, Lindberg J, Germain M, et al. A longitudinal study of uremic pruritus in hemodialysis patients. *Clin J Am Soc Nephrol* 2010;5(8):1410-9

20. Pisoni RL, Wikström B, Elder SJ, et al. Pruritus in haemodialysis patients: International results from the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Nephrol Dial Transplant* 2006;21(12):3495-505

21. Weiss M, Mettang T, Tschulena U, Weisshaar E. Health-related quality of life in haemodialysis patients suffering from chronic itch: results from GEHIS (German Epidemiology Haemodialysis Itch Study). *Qual Life Res* 2016;25(12):3097-106

22. Kimata N, Fuller DS, Saito A, et al. Pruritus in hemodialysis patients: Results from the Japanese Dialysis Outcomes and Practice Patterns Study (JDOPPS). *Hemodial Int* 2014;18(3):657-67

23. van der Willik EM, Terwee CB, Bos WJW, et al. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. *Nephrology (Carlton)* 2021;26(5):391-9

24. Clement ND, Weir D, Holland J, Gerrand C, Deehan DJ. Meaningful changes in the Short Form 12 physical and mental summary scores after total knee arthroplasty. *The Knee* 2019;26(4):861-8

25. von der Lippe N, Waldum B, Brekke FB, Amro AA, Reisæter AV, Os I. From dialysis to transplantation: a 5-year longitudinal study on self-reported quality of life. *BMC Nephrol* 2014;15:191

26. Ogura K, Yakoub MA, Christ AB, et al. What Are the Minimum Clinically Important Differences in SF-36 Scores in Patients with Orthopaedic Oncologic Conditions? *Clin Orthop Relat Res* 2020;478(9):2148-58

27. Ting SW, Fan PC, Lin YS, et al. Uremic pruritus and long-term morbidities in the dialysis population. *PLoS One* 2020;15(10):e0241088

28. Orasan OH, Muresan F, Mot A, et al. Hemodialysis Patients with Pruritus and Insomnia Have Increased Risk of Death. *Blood Purif* 2020;49(4):419-25

29. Rehman IU, Chohan TA, Bukhsh A, Khan TM. Impact of Pruritus on Sleep Quality of Hemodialysis Patients: A Systematic Review and Meta-Analysis. *Medicina (Kaunas)* 2019;55(10)

30. Li J, Guo Q, Lin J, Yi C, Yang X, Yu X. Prevalence and Associated Factors of Uraemic Pruritus in Continuous Ambulatory Peritoneal Dialysis Patients. *Intern Med* 2015;54(22):2827-33

31. Zucker I, Yosipovitch G, David M, Gafter U, Boner G. Prevalence and characterization of uremic pruritus in patients undergoing hemodialysis: uremic pruritus is still a major problem for patients with end-stage renal disease. *J Am Acad Dermatol* 2003;49(5):842-6

32. Simonsen E, Komenda P, Lerner B, et al. Treatment of Uremic Pruritus: A Systematic Review. *American Journal of Kidney Diseases* 2017;70(5):638-55

33. Rayner HC, Larkina M, Wang M, et al. International Comparisons of Prevalence, Awareness, and Treatment of Pruritus in People on Hemodialysis. *Clin J Am Soc Nephrol* 2017;12(12):2000-7

34. Aresi G, Rayner HC, Hassan L, et al. Reasons for Underreporting of Uremic Pruritus in People With Chronic Kidney Disease: A Qualitative Study. *Journal of Pain and Symptom Management* 2019;58(4):578-86.e2

Supplemental Material for Chapter 6

A. Population characteristics in patients receiving haemodialysis and peritoneal dialysis

Table S1 presents the characteristics of all haemodialysis and peritoneal dialysis patients that completed the PROMs at least once in 2018-2020, stratified for the presence of itching at baseline.

Table S1. Characteristics of haemodialysis and peritoneal dialysis patients, stratified by presence of itching (yes/no) at baseline

	Haemodialysis patients		Peritoneal dialysis patients	
	with itching (n=1258, 48.7%)	**without itching** (n=1325, 51.3%)	**with itching** (n=212, 59.4%)	**without itching** (n=145, 40.6%)
Age (years)	67.8 (14.2)	67.5 (14.1)	65.6 (12.8)	65.3 (15.4)
Sex (male)	775 (61.6)	804 (60.7)	135 (63.7)	85 (58.6)
SES				
Low	632 (50.5)	660 (50.2)	72 (34.6)	57 (39.3)
Middle	353 (28.2)	410 (31.2)	75 (36.1)	58 (40.0)
High	266 (21.3)	245 (18.6)	61 (29.3)	30 (20.7)
Primary kidney disease				
Glomerulonephritis/sclerosis	128 (10.2)	154 (11.6)	28 (13.2)	18 (12.4)
Pyelonephritis	60 (4.8)	63 (4.8)	6 (2.8)	9 (6.2)
Polycystic kidney disease	72 (5.7)	73 (5.5)	13 (6.1)	11 (7.6)
Hypertension/ renal vascular disease	336 (26.7)	351 (26.5)	70 (33.0)	46 (31.8)
Diabetes mellitus type 1/2	278 (22.1)	256 (19.4)	37 (17.4)	23 (15.9)
Miscellaneous	223 (17.7)	258 (19.5)	30 (14.2)	18 (12.4)
Unknown	161 (12.8)	170 (12.8)	28 (13.2)	20 (13.8)
Time since dialysis initiation (months)	17 (3-44)	19 (3-49)	4 (1-17)	3 (1-14)

Kidney transplantation in past (yes)	148 (11.8)	157 (11.9)	14 (6.7)	8 (5.6)
rGFR (mL/min/1.73m²)	5.0 (2.0-8.0)	4.6 (2.1-7.1)	5.2 (2.0-7.1)	4.0 (3.0-5.5)
Kt/V@	1.46 (0.54)	1.48 (0.52)	2.70 (1.10)	2.52 (0.99)
Haemoglobin (mmol/L)	6.8 (0.9)	6.8 (0.9)	6.9 (0.9)	6.9 (0.9)
Ferritin (µg/L)	316 (167-532)	340 (194-550)	178 (86-383)	226 (113-483)
Transferrin saturation (%)	21.8 (10.4)	22.6 (10.6)	22.0 (10.1)	25.3 (10.5)
Calcium (mmol/L)#	2.30 (0.19)	2.30 (0.18)	2.32 (0.18)	2.32 (0.20)
Phosphate (mmol/L)	1.63 (0.51)	1.57 (0.47)	1.55 (0.36)	1.53 (0.43)
Parathyroid hormone (pmol/L)	31 (18-54)	30 (17-51)	25 (15-40)	27 (18-40)
Symptom burden				
Total number of symptoms (0-30)	13.7 (6.2)	8.2 (5.4)	13.3 (6.1)	8.3 (4.5)
Total symptom burden score (0-150)	35 (23-52)	19 (9-32)	33 (23-46)	19 (11-31)
Dry skin (yes)	907 (72.2)	574 (44.2)	167 (78.8)	56 (40.3)
Sleep problems				
Sleep problems (yes)^	875 (69.6)	694 (52.4)	152 (71.7)	71 (49.0)
Trouble falling asleep (yes)	674 (53.6)	468 (35.3)	110 (51.9)	42 (29.0)
Trouble staying asleep (yes)	758 (60.3)	579 (43.7)	135 (63.7)	56 (38.6)
Psychological symptoms				
Psychological symptoms (yes)$	464 (36.9)	262 (19.8)	61 (28.8)	22 (15.2)
Worrying (yes)	611 (48.6)	439 (33.1)	85 (40.1)	44 (30.3)
Feeling nervous (yes)	454 (36.1)	270 (20.4)	55 (25.9)	26 (17.9)

Feeling irritable (yes)	473 (37.6)	254 (19.2)	75 (35.4)	28 (19.3)
Feeling sad (yes)	571 (45.4)	387 (29.2)	86 (40.6)	29 (20.2)
Feeling anxious (yes)	374 (29.7)	204 (15.4)	44 (20.8)	15 (10.3)

Values are shown in n (%), mean (SD) or median (IQR).
[@] Single pool Kt/V in haemodialysis and total Kt/V in peritoneal dialysis.
[#] Albumin-adjusted calcium.
[^] Sleep problems are considered present if at least one of the two symptoms are experienced by the patient.
[$] Psychological symptoms are considered present if three out of the five symptoms are experienced by the patient.
Abbreviations: SES, social economic status; rGFR, residual glomerular filtration rate.

B. Persistence of sleep problems and psychological symptoms

PERSISTENCE OF SLEEP PROBLEMS

The persistence of sleep problems is shown graphically based on patients' follow-up time, stratified for sleep problems at baseline (Figure S1a) and for itching at baseline (Figure S1b).

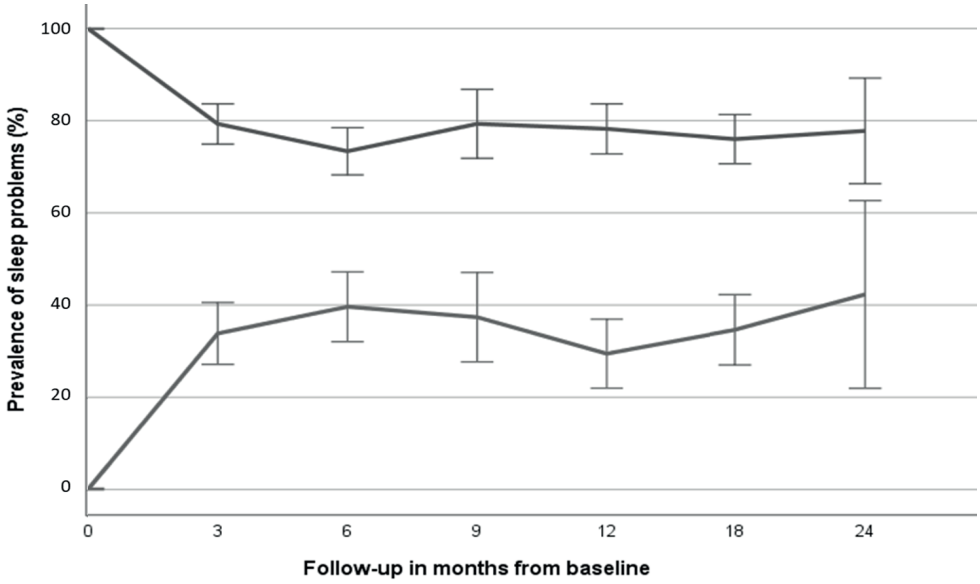


Figure S1a. Persistence of sleep problems during follow-up in patients with sleep problems (upper line) and patients without sleep problems (lower line) at baseline.
Upper solid line (upper bars) shows the percentage (95% Confidence Interval [CI]) of dialysis patients in which sleep problems are persistent during follow-up since baseline. Lower solid line (lower bars) shows the percentage (95% CI) of dialysis patients in which sleep problems were newly developed during follow-up since baseline.

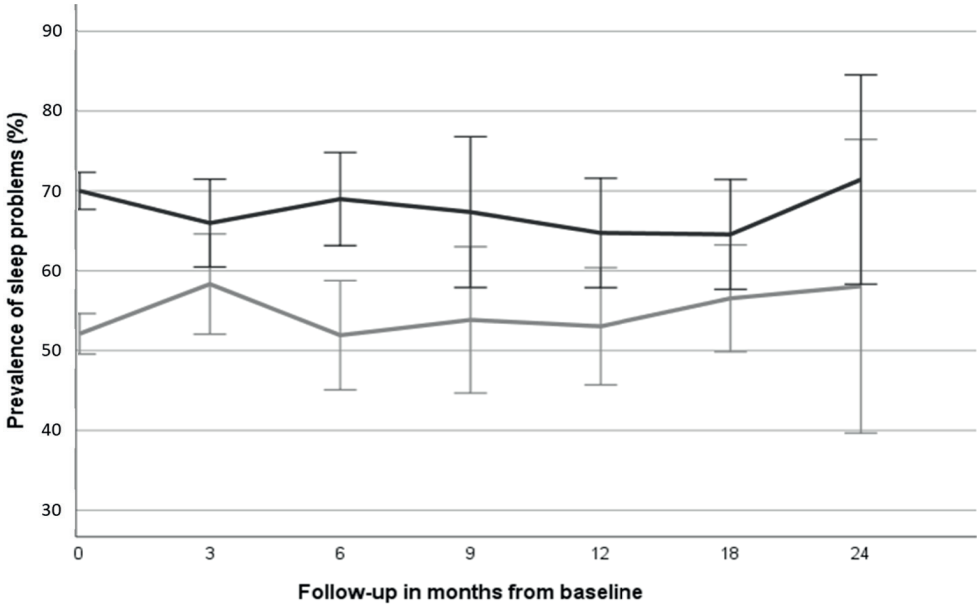


Figure S1b. Persistence of sleep problems during follow-up in patients with itching (upper line) and patients without itching (lower line) at baseline.

Solid lines (bars) show the percentages (95% Confidence Intervals [CI]) in which sleep problems are present during follow-up, stratified for patients with itching (upper line) and without itching (lower line) at baseline. Note that itching is persistent in approximately 70% of the dialysis patients.

PERSISTENCE OF PSYCHOLOGICAL SYMPTOMS

The persistence of psychological symptoms is shown graphically based on patients' follow-up time, stratified for psychological symptoms at baseline (Figure S2a) and itching at baseline (Figure S2b).

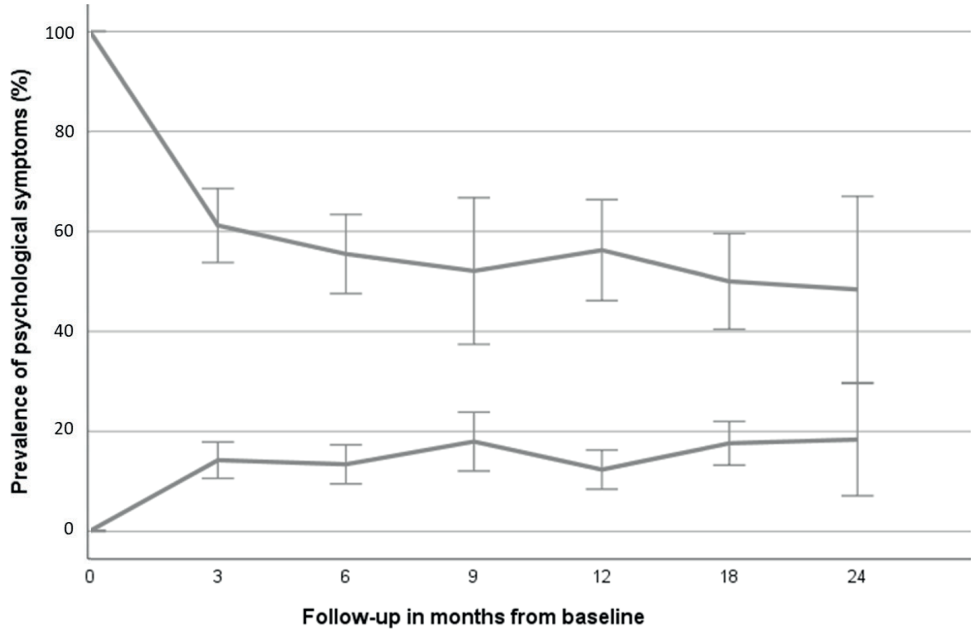


Figure S2a. Persistence of psychological symptoms during follow-up in patients with psychological symptoms (upper line) and patients without psychological symptoms (lower line) at baseline.

Upper solid line (upper bars) shows the percentage (95% Confidence Interval [CI]) of dialysis patients in which psychological symptoms are persistent during follow-up since baseline. Lower solid line (lower bars) shows the percentage (95% CI) of dialysis patients in which psychological symptoms were newly developed during follow-up since baseline.

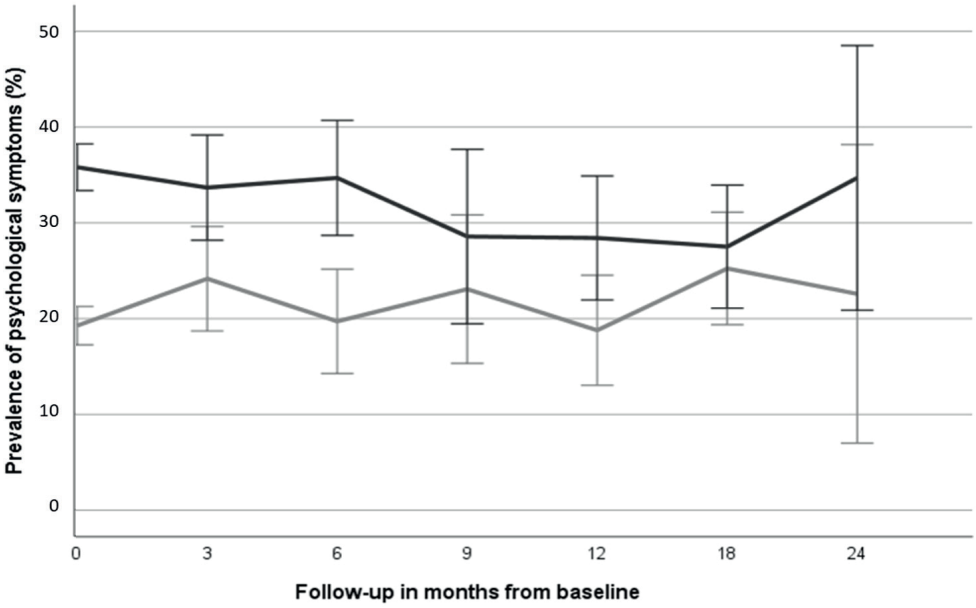


Figure S2b. Persistence of psychological symptoms during follow-up in patients with itching (upper line) and patients without itching (lower line) at baseline.
Solid lines (bars) show the percentages (95% Confidence Intervals [CI]) in which psychological symptoms are present during follow-up, stratified for patients with itching (upper line) and without itching (lower line) at baseline. Note that itching is persistent in approximately 70% of the dialysis patients.

C. Change in itching and HRQOL

A post hoc analysis was performed on the change in HRQOL in dialysis patients where itching disappeared (n=185; 15.2%) and where itching newly occurred (n=181; 14.9%) between their first and second PROMs measurement (see Table S2).

Table S2. Change in physical and mental HRQOL in dialysis patients where itching disappeared or newly occurred

	Δ Physical HRQOL	p-value*	Δ Mental HRQOL	p-value*
Itching disappeared	+ 0.56	0.489	+ 1.78	0.023
Itching newly occurred	- 0.44	0.613	- 0.68	0.380

** p-value based on paired samples t-test.*

D. Sensitivity analyses

BURDEN OF ITCHING AND HRQOL

The main analyses were also performed using the continuous burden scores for itching (range: 0-5), sleep problems (range: 0-10) and psychological symptoms (range: 0-25). A burden score of 0 refers to the symptom being absent and a higher score indicates the level of symptom burden that is experienced by the patient. Table S3 shows the cross-sectional effects of the burden of itching, combined with the burden of sleep problems and psychological symptoms, on physical and mental HRQOL.

Table S3. Cross-sectional effects of the burden of itching, combined with the burden of sleep problems and psychological symptoms, on physical and mental HRQOL

	Physical HRQOL		Mental HRQOL	
	Coef. (95%CI)	p-value	Coef. (95%CI)	p-value
Itching[§]				
Model 1, unadjusted	-1.29 (-1.52; -1.05)	<0.001	-1.45 (-1.69; -1.22)	<0.001
Model 2, adjusted[^]	-1.26 (-1.50; -1.02)	<0.001	-1.42 (-1.65; -1.18)	<0.001
Itching and sleep problems[§] (Model 3[$])				
Itching	-1.18 (-1.52; -0.83)	<0.001	-1.05 (-1.39; -0.71)	<0.001
Sleep problems	-0.72 (-0.90; -0.54)	<0.001	-0.87 (-1.05; -0.70)	<0.001
Itching * sleep problems	0.07 (-0.00; 0.14)	0.063[#]	0.02 (-0.05; 0.09)	0.557[#]
Itching and psychological symptoms[§] (Model 4[$])				
Itching	-1.21 (-1.52; -0.90)	<0.001	-0.77 (-1.02; -0.52)	<0.001
Psychological symptoms	-0.44 (-0.55; -0.33)	<0.001	-1.29 (-1.38; -1.20)	<0.001
Itching * psychological symptoms	0.05 (0.01; 0.09)	0.018[#]	0.06 (0.03; 0.09)	<0.001[#]

[§] Burden of itching on a 0-5 scale, burden of sleep problems on a 0-10 scale and burden of psychological symptoms on a 0-25 scale, with higher scores indicating a higher burden.

[^] Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.

[$] Model 3 and 4 build on model 2 and include the interaction with sleep problems and psychological symptoms, respectively.

[#] P-value for interaction.

Findings from the longitudinal analyses using linear mixed models were similar to the cross-sectional analyses, showing a monotonic association between burden of itching and HRQOL. Physical and mental HRQOL was -1.19 (95% CI: -1.41 to -0.96; p<0.001) and -1.37 (95% CI: -1.59 to -1.14; p<0.001) points lower, respectively, for each point increase in burden of itching. No significant changes in physical and mental HRQOL were observed in the total population during follow-up (annual change: -0.23 [95% CI: -0.88 to 0.41; p=0.48] and 0.02 [95% CI: -0.65 to 0.69; p=0.95], resp.). No differences in physical and mental HRQOL trajectories were observed for higher burden of itching (extra annual change for each point increase in burden of itching: 0.13 [95% CI: -0.16 to 0.41; p=0.40] and -0.01 [95% CI: -0.31 to 0.29; p=0.92], resp.). There was no significant interaction between the burden of itching and sleep problems in the association with physical and mental HRQOL (p=0.13 and p=0.89, resp.). The interaction between burden of itching and psychological symptoms in the association with physical and mental HRQOL became statistically significant (p=0.04 and p<0.001, resp.), though with a similarly small effect.

MODERATE TO SEVERE ITCHING AND HRQOL

The main analyses were also performed with the symptoms classified based on low or high burden: no or mild itching (burden score: 0-2) versus moderate to severe itching (burden score: 3-5), combined with no or mild sleep problems (burden score: 0-4) versus moderate to severe sleep problems (burden score: 5-10) and with no or mild psychological symptoms (burden score: 0-10) versus moderate to severe psychological symptoms (burden score: 10-25).

In total, 773 (26.1%) patients had moderate to severe itching, 814 (27.3%) patients had moderate to severe sleep problems and 380 (12.8%) patients had moderate to severe psychological symptoms.

Table S4 shows the cross-sectional effects of moderate to severe itching, combined with moderate to severe sleep problems and psychological symptoms, on physical and mental HRQOL.

Findings from the longitudinal analyses using linear mixed models were similar to the cross-sectional analyses, showing that patients with moderate to severe itching experienced a lower physical and mental HRQOL compared to patients with no or mild itching (-3.98 [95% CI: -4.82 to -3.14; p<0.001] and -4.66 [95% CI: -5.49 to -3.83; p<0.001], resp.). No significant changes in physical and mental HRQOL were observed in the total population during follow-up (annual change: 0.22 [95% CI: -0.73 to 1.16; p=0.66] and -0.19 [95% CI: -1.17 to 0.78; p=0.70], resp.). No differences in physical and mental HRQOL trajectories were observed between pa-

tients with moderate to severe itching and no or mild itching (extra annual change in patients with moderate to severe itching: 0.37 [95% CI: -0.71 to 1.44; p=0.50] and -0.26 [95% CI: -1.37 to 0.85; p=0.65], resp.). Also longitudinally, in the association with physical and mental HRQOL, there was no significant interaction between moderate to severe itching and sleep problems (p=0.30 and p=0.35, resp.) or moderate to severe itching and psychological symptoms (p=0.63 and p=0.71, resp.).

Table S4. Cross-sectional effects of moderate to severe itching, combined with moderate to severe sleep problems and psychological symptoms, on physical and mental HRQOL

	Physical HRQOL		Mental HRQOL	
	Coef. (95%CI)	p-value	Coef. (95%CI)	p-value
Itching				
Model 1, unadjusted	-4.33 (-5.20; -3.46)	<0.001	-4.98 (-5.85; -4.12)	<0.001
Model 2, adjusted^	-4.20 (-5.07; -3.33)	<0.001	-4.90 (-5.76; -4.03)	<0.001
Itching and sleep problems (Model 3$)				
Itching	-3.88 (-4.99; -2.77)	<0.001	-4.31 (-5.40; -3.22)	<0.001
Sleep problems	-3.55 (-4.63; -2.47)	<0.001	-5.07 (-6.13; -4.02)	<0.001
Itching * sleep problems	1.11 (-0.74; 2.95)	0.239#	1.30 (-0.51; 3.11)	0.158#
Itching and psychological symptoms (Model 4$)				
Itching	-3.90 (-4.87; -2.93)	<0.001	-3.04 (-3.92; -2.17)	<0.001
Psychological symptoms	-3.57 (-5.13; -2.02)	<0.001	-13.21 (-14.62; -11.80)	<0.001
Itching * psychological symptoms	0.94 (-1.40; 3.28)	0.432#	0.28 (-1.84; 2.40)	0.794#

^ *Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.*
$ *Model 3 and 4 build on model 2 and include the interaction with sleep problems and psychological symptoms, respectively.*
# *P-value for interaction.*

PERSISTENT ITCHING AND HRQOL

The main analyses were repeated in individuals with multiple PROMs measurements (n=1218) to compare persistent itching with no or non-persistent itching. Persistent itching was defined as the presence of itching at baseline and at the first follow-up measurement. In total, 430 (35.3%) patients had persistent itching.
Table S5 shows the cross-sectional effects of persistent itching, combined with sleep problems and psychological symptoms, on physical and mental HRQOL.

Table S5. Cross-sectional effects of persistent itching, combined with sleep problems and psychological symptoms, on physical and mental HRQOL

	Physical HRQOL		Mental HRQOL	
	Coef. (95%CI)	p-value	Coef. (95%CI)	p-value
Itching				
Model 1, unadjusted	-2.97 (-4.22; -1.72)	<0.001	-3.57 (-4.81; -2.34)	<0.001
Model 2, adjusted^	-3.27 (-4.55; -2.00)	<0.001	-3.29 (-4.53; -2.04)	<0.001
Itching and sleep problems (Model 3$)				
Itching	-4.07 (-6.20; -1.94)	<0.001	-2.72 (-4.80; -0.64)	0.010
Sleep problems	-3.26 (-4.77; -1.75)	<0.001	-3.01 (-4.48; -1.54)	<0.001
Itching * sleep problems	1.78 (-0.87; 4.42)	0.187#	-0.25 (-2.83; 2.33)	0.851#
Itching and psychological symptoms (Model 4$)				
Itching	-2.45 (-3.98; -0.92)	0.002	-1.74 (-3.10; -0.38)	0.012
Psychological symptoms	-3.56 (-5.41; -1.72)	<0.001	-10.58 (-12.21; -8.94)	<0.001
Itching * psychological symptoms	-0.52 (-3.27; 2.23)	0.709#	0.78 (-1.66; 3.21)	0.532#

^ *Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.*
$ *Model 3 and 4 build on model 2 and include the interaction with sleep problems and psychological symptoms, respectively.*
# *P-value for interaction.*

Findings from the longitudinal analyses using linear mixed models were similar to the cross-sectional analyses, showing that patients with persistent itching experienced a lower physical and mental HRQOL compared to patients with no or non-persistent itching (-2.99 [95% CI: -4.14 to -1.82; p<0.001] and -3.46 [95% CI: -4.59 to -2.32; p<0.001], resp.). No significant changes in physical and mental HRQOL were observed in the total population during follow-up (annual change: -0.73 [95% CI: -1.61 to 0.16; p=0.11] and -0.48 [95% CI: -1.42 to 0.46; p=0.31], resp.). No differences in physical and mental HRQOL trajectories were observed between patients with persistent itching and no or non-persistent itching (extra annual change in patients with persistent itching: -0.64 [95% CI: -1.66 to 0.38; p=0.22] and 0.12 [95% CI: -0.96 to 1.20; p=0.83], resp.). Also longitudinally, in the association with physical and mental HRQOL, there was no significant interaction between persistent itching and sleep problems (p=0.88 and p=0.86, resp.) or persistent itching and psychological symptoms (p=0.49 and p=1.00, resp.).

ITCHING AND HRQOL USING 2019-2020 DATA

The main analyses were repeated using data from 2019 (n=1416) and 2020 (n=1436), to only include measurements from the official start of the PROMs registry at November 2018.

Table S6 shows the cross-sectional effects of the presence of itching, combined with sleep problems and psychological symptoms, on physical and mental HRQOL. Findings from the longitudinal analyses using linear mixed models were similar to the cross-sectional analyses, showing that patients with itching experienced a lower physical and mental HRQOL compared to patients without itching (-3.12 [95% CI: -3.86 to -2.37; p<0.001] and -3.61 [95% CI: -4.35 to -2.87; p<0.001], resp.). No significant changes in physical and mental HRQOL were observed in the total population during follow-up (annual change: 0.29 [95% CI: -0.45 to 1.02; p=0.45] and 0.09 [95% CI: -0.66 to 0.83; p=0.82], resp.). No differences in physical and mental HRQOL trajectories were observed between patients with and without itching (extra annual change in patients with itching: 0.38 [95% CI: -0.61 to 1.37; p=0.45] and 0.05 [95% CI: -0.96 to 1.06; p=0.92], resp.). Also longitudinally, in the association with physical and mental HRQOL, there was no significant interaction between itching and sleep problems (p=0.49 and p=0.34, resp.) or itching and psychological symptoms (p=0.54 and p=0.18, resp.).

Table S6. Cross-sectional effects of the presence of itching, combined with sleep problems and psychological symptoms, on physical and mental HRQOL (2019-2020)

	Physical HRQOL		Mental HRQOL	
	Coef. (95%CI)	p-value	Coef. (95%CI)	p-value
Itching				
Model 1, unadjusted	-3.28 (-4.05; -2.50)	<0.001	-3.85 (-4.61; -3.08)	<0.001
Model 2, adjusted^	-3.32 (-4.10; -2.55)	<0.001	-3.84 (-4.61; -3.08)	<0.001
Itching and sleep problems (Model 3$)				
Itching	-3.28 (-4.54; -2.03)	<0.001	-2.58 (-3.81; -1.34)	<0.001
Sleep problems	-3.75 (-4.82; -2.68)	<0.001	-3.28 (-4.35; -2.22)	<0.001
Itching * sleep problems	0.90 (-0.68; 2.49)	0.264#	-0.97 (-2.54; 0.61)	0.228#
Itching and psychological symptoms (Model 4$)				
Itching	-2.73 (-3.63; -1.83)	<0.001	-2.40 (-3.20; -1.60)	<0.001
Psychological symptoms	-3.27 (-4.65; -1.90)	<0.001	-11.56 (-12.78; -10.33)	<0.001
Itching * psychological symptoms	-0.16 (-1.94; 1.61)	0.859#	1.30 (-0.27; 2.88)	0.105#

^ *Adjusted for age, sex, primary kidney disease, socio-economic status, dialysis modality, time since dialysis initiation and kidney transplantation in past.*
$ *Model 3 and 4 build on model 2 and include the interaction with sleep problems and psychological symptoms, respectively.*
# *P-value for interaction.*

7

Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time

Esmee M. van der Willik, Caroline B. Terwee, Willem Jan W. Bos, Marc H. Hemmelder, Kitty J. Jager, Carmine Zoccali, Friedo W. Dekker, Yvette Meuleman

Nephrology (Carlton). 2021;26(5):391-399.

Abstract

Patient-reported outcome measures (PROMs) are increasingly being used in nephrology care. However, in contrast to well-known clinical measures such as blood pressure, healthcare professionals are less familiar with PROMs and the interpretation of PROM scores is therefore perceived as challenging. In this paper, we provide insight into the interpretation of PROM scores by introducing the different types and characteristics of PROMs, and the most relevant concepts for the interpretation of PROM scores. Concepts such as minimal detectable change, minimal important change and response shift are explained and illustrated with examples from nephrology care.

Introduction

Over the last decades, a shift towards a more value-based and patient-centred healthcare has taken place, resulting in a stronger focus on patient-reported outcomes (PROs) such as health-related quality of life (HRQOL) and symptom burden.[1, 2] PRO measures (PROMs) are nowadays introduced in nephrology care and may be used at individual level for personalized care and at aggregated level to evaluate healthcare quality. The use of PROMs at individual level as part of personalized care has been considered of great added value, as it may provide insight into patients' perceived health and their needs, and enhance patient-professional communication and shared decision making.[3, 4] Ultimately, PROMs can be used to improve symptom management, HRQOL and other outcomes of healthcare.[5, 6]
To achieve such goals, knowledge about PROMs and the interpretation of PROM scores are needed. In contrast to well-known clinical outcomes such as blood pressure, healthcare professionals and researchers are not yet familiar with PROMs and the interpretation of PROM scores is therefore perceived as challenging. For example: What does a symptom burden score of 27 mean? Is a HRQOL-score of 36 normal for a certain patient or in a certain situation? Is a change in HROQL-score of 4 points clinically relevant? And why does the change in PROM score not always reflect the clinical change in health status?
In this paper, we provide insight into the interpretation of PROM scores by introducing the different types and characteristics of PROMs, and by presenting the most relevant concepts for the interpretation of PROM scores (i.e. minimal detectable change, minimal important change, and response shift), illustrated with examples from nephrology care.

Patient-reported outcome measures

PROs are *outcomes* on aspects of patients' perceived health, which includes a variety of concepts, for instance: HRQOL, functional status or symptom burden. PROs can be best measured by asking the patient himself and are reported by the patient himself (support may be offered when filling in PROMs, as long as responses reflect the patient's perspective). PROMs are *questionnaires* that assess these aspects of perceived health. PROMs do not include experiences with, or perceptions and evaluations of healthcare provision; for this purpose, other measures are used, namely patient-reported experience measures (PREMs). Table 1 provides an overview of the terms used in this article.

Table 1. Overview of terms used in this article

Patient reported outcome (PRO)	Outcomes on aspects of patients' perceived health, reported from the patient's perspective. E.g. health-related quality of life (HRQOL), functional status or symptom burden.
Patient reported outcome measure (PROM)	Questionnaire to measure one or multiple PROs (i.e. uni- or multidimensional PROM). PROMs are often classified as either a generic PROM or a specific PROM (i.e. for a certain disease or condition).
PROM score	Score for a PRO as measured by a PROM (i.e. the result from a PROM), which can be a score for one item or multiple items.
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."[7]
Minimal detectable change (MDC)	A parameter of reliability that is defined as the "smallest change in score that can be detected beyond measurement error."[8]
Minimal important change (MIC)	"The smallest change in score in the construct to be measured which patients perceive as important."[8]
Response shift	"A change in the meaning of one's self-evaluation, which can be a result of recalibration, reprioritization and/or reconceptualization of the PRO."[9]

Various types of PROMs exist and knowledge about certain characteristics of the PROMs is required to properly interpret PROM scores. Therefore, we will briefly introduce different types and characteristics of PROMs and elaborate on how they relate to the interpretation of PROM scores.

Generic and specific PROMs

PROMs can roughly be classified as either generic or specific for a certain disease, condition or treatment. Generic PROMs measure a wide variety of health aspects and usually include aspects of people's health that are widely relevant (e.g. functional status or HRQOL in its broadest sense). Generic PROMs can therefore be used in any population, hereby enabling comparisons across populations or treatments, and are very suitable for heterogeneous and multimorbid populations (e.g. the elderly patient with chronic kidney disease (CKD) who often suffers from multiple comorbid conditions such as diabetes mellitus and cardiovascular disease). A disadvantage of this broad applicability is that it often goes with less precise PRO estimates and that nuances or small differences in PROs between or within specific populations may remain undetected.

Specific PROMs are tailored to a certain disease, condition or treatment, and address issues that are relevant to a specific group of patients, for example symptom burden related to CKD or related to immunosuppressive treatment after kidney transplantation.[10, 11] By tailoring to particular conditions, specific PROMs are usually better able to detect smaller or more specific differences or changes in PROs (e.g. a change in intensity or type of itching). Hence, specific PROMs are particularly suitable for comparisons within a population, but not for comparisons across populations. A disadvantage of a specific PROM is that relevant outcomes may be missed due to the focus on a certain disease or condition, for instance in heterogeneous populations with multiple comorbid conditions.

Whether a generic or a specific PROM is suitable depends on various aspects, including which PRO you aim to measure (e.g. disease specific symptoms or general functional status), the setting and purpose of measuring the PRO (e.g. is comparison within or also across populations of interest?), the diversity and characteristics of the population of interest (e.g. heterogeneity of the population), and the availability and quality of instruments (i.e. are high-quality and validated generic and/or specific PROMs available?). In practice, a combination of generic and specific PROMs is often used; either combined into one PROM such as the 36-item Kidney Disease Quality of Life (KDQOL-36) measuring generic HRQOL and kidney disease specific burden[12, 13], or as separate PROMs for instance a combination of the SF-12 to measure generic HRQOL[14] and the Dialysis Symptom Index (DSI)[10] to measure kidney disease specific symptoms. The latter combination is used since 2018 in Dutch dialysis care[3], for which the selection of the DSI has been described in detail elsewhere.[15]

Scoring systems of PROMs

A standard PROM scoring system or scale does not exist, not even when PROMs are measuring the same PRO. In contrast to other measures (e.g. temperature and distance) that can be measured on the same scale (e.g. Celsius and meters), PROMs use varying scales and scoring methods.

Table 2 presents an example of three PROMs that measure HRQOL (PROMIS Profile-29), symptoms (DSI) or both (KDQOL-36) to illustrate the variety in measurement characteristics across PROMs. The PROMs differ for many features, such as the domains being measured (also for the same PRO, i.e. HRQOL), the number of questions, response options, scales and scoring methods. As a result of the differences in features, PROMs often also differ in the interpretation of scores. For example: although the DSI and the KDQOL-36 both measure disease specific symptoms, PROM scores are not directly comparable due to different scoring systems (e.g. score range, method and direction; Table 2). A KDQOL-36 symptom burden score of 71 represents a reasonable health status similar to that of an average patient with CKD.[13, 16] However, a DSI symptom burden score of 71 represents an extremely high symptom burden that is twice as high as in an average dialysis patient.[3]

Measurement properties of PROMs

Measurement properties such as validity and reliability provide essential information about the quality of the PROM in certain populations and settings. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) taxonomy describes which aspects should be considered to judge the quality of the PROM.[7] Good measurement properties are a prerequisite for PROMs to be useful and reasonably interpretable. However, measurement properties such as validity and reliability itself provide insufficient insight into the meaning of scores, i.e. the interpretation of PROM-scores.

Interpretation of PROM scores

The interpretability of a PROM has been defined as "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".[7] The interpretability can be considered a characteristic of the PROM, meaning that one PROM may be easier to interpret than another PROM. The interpretation of PROM scores can be challenging, for instance due to the complexity of the PRO (e.g. HRQOL, which includes various physical, mental and social domains) or the PROM (e.g. a

complex scoring method). Luckily, there are some intuitive methods that may facilitate the interpretation of PROM scores which will be discussed below.

First, a discussion on PROs between the patient and the professional may provide insight into the individual's view on certain aspects of health, e.g. what is important to the patient and what is his frame of reference. The PROM items and also the overall PROM scores may facilitate this conversation, for instance by serving as a checklist or as a reason to start the conversation about (difficult) subjects.[3]

Second, group-level data may facilitate the interpretation of individual PROM-scores by providing insight into what is 'normal' and what may be expected. Descriptive information such as the mean, standard deviation and range in the population of interest gives an indication of the variability of scores (i.e. should scores be expected across the whole scale or on a smaller range?) and of what is 'normal' (e.g. is the score of a patient low, average or high as compared to other patients?). Comparison to norm- or reference scores of a general population or a population with a certain condition or treatment can be highly informative. For example: comparing a 65-year old dialysis patient's HRQOL-score of 40 to the average Dutch dialysis population (mean score: 36 (SD 11))[3] and the general 60-69 year old Dutch population (mean score: 51 (SD 9))[18] gives an idea of how the patient addresses his outcome in comparison to the reference population. Furthermore, descriptive information about floor- or ceiling effects, meaning that many individuals score at the lower (i.e. floor) or upper (i.e. ceiling) end of the scale, may be informative because differences below or above these limits cannot be observed. This may be valuable information to take into account when interpreting individual patient scores.

Third, it is insightful to compare PROM scores to scores of other measures. Since most PROMs are relatively new to clinical care, most users (both patients and healthcare professionals) are not yet sufficiently familiar with PROM scores. By comparing PROM scores to well-known (clinical) measures such as kidney function or laboratory measures and to patient- or disease characteristics, one may become more experienced with the scores and get a feeling for which scores are common for certain patients, conditions and situations (i.e. the scores get 'clinical or commonly understood connotations').

Finally, the interpretability of PROM scores may automatically improve over time when patients and professionals become more experienced in using and discussing PROM scores. In addition to these more intuitive aspects of interpreting PROM scores, there are also methodological concepts, i.e. benchmarks, that are relevant to the interpretability of *changes* in PROM scores, which will be discussed below.

Table 2. Illustration of variation in characteristics across different patient-reported outcome measures

	PROMIS Profile-29	KDQOL-36	DSI
PRO	HRQOL	Disease burden and HRQOL	Symptom burden
Target population[$]	People with or without (chronic) illness	Patients with kidney disease	Haemodialysis patients
Type	Generic	Disease specific and generic[#]	Disease specific
Domains	Depression Anxiety Physical function Pain interference Fatigue Sleep disturbance Ability to participate in social roles and activities Pain intensity	Disease specific: Symptoms/problems Effects of kidney disease Burden of kidney disease Generic[#]: SF-12 Physical Health Composite SF-12 Mental Health Composite	Symptom burden
Number of questions	29, or tailored to the patient[§]	36	30
Recall period	In general/1 week	In general/4 weeks	1 week
Rating scale	5-point Likert scale, 0-10 scale (for pain intensity only)	Various scales: Yes/no, 3-, 5- or 6-point scale	Yes/no (presence of symptoms), 5-point Likert scale (severity)
Item score	1 to 5 points or vice versa, so that a higher score represents more of the domain being measured.	Item-scores are transformed to a 0-100 possible range. E.g. the 5-point scale has 0/25/50/75/100 points.	0 points if symptom is not present; 1 to 5 points for severity[^]
Total score (range)	T-score (roughly 0-100)	0-100	0-150[^]
Scoring method	IRT-based scoring	Disease specific: average score Generic[#]: norm-based scoring algorithm	Sum score[^]

Meaning of score direction	Higher scores represent more of the domain being measured. E.g. a higher score on fatigue means a worse fatigue, and a higher score on physical function means a better physical function.	Higher scores represent a more favourable health state. E.g. a higher score on symptoms means a lower symptom burden, and a higher score on physical health means a better physical health.	Higher scores represent a higher symptom burden.
Norm- or reference standard	General US population: mean 50, SD 10	Disease specific: n/a. Generic[#]: General US population: mean 50, SD 10	N/a

[$] The target population is the population for which the PROM was originally developed and is not necessarily the only population for which the questionnaire is used and considered suitable.

[§] PROMIS questionnaires can be applied as Computerized Adaptive Test (CAT) per domain, whereby the computer selects items based on the patient's responses to previous questions. The number of questions usually depends on a predetermined threshold for the precision of the measurements and may therefore vary across patients and measurements.

[#] The generic part of the KDQOL-36 is the 12-item short form (SF-12) health survey.

[^] In the original development paper of the DSI[10], a 0-4 scale was used for severity and no guidance for an overall score was provided. Therefore, the symptom burden score is often calculated according to the method presented in this table, which was previously described by Abdel-Kader et al. (2009).[17]

Abbreviations: DSI, Dialysis Symptom Index; KDQOL-36, 36-item Kidney Disease Quality of Life; PROMIS, Patient-Reported Outcomes Measurement Information System; IRT, Item Response Theory; n/a, not available.

Minimal detectable change (MDC)

Suppose that a patient with advanced CKD fills in the Short Form-36 (SF-36) twice with a 6 months interval between the two measurements. The HRQOL results show a decrease of 5 points at the physical component score (hereafter called 'physical HRQOL') and a decrease of 2 points at the mental component score (hereafter called 'mental HRQOL'). Can we then speak of a real deterioration in HRQOL? In other words, do we observe an actual change or is it possibly just random variation? To answer this question we need to know whether the observed change is larger than the minimal detectable change (MDC), also known as the smallest detectable change or the minimal real change. The MDC is a parameter of reliability and is defined as the "smallest change in score that can be detected beyond

measurement error".[8] Thus, the MDC reflects the threshold at which a change in score can be considered statistically significant.

The MDC should be estimated in persons who have *not* changed over time (e.g. clinically stable patients) using a test-retest design, because this demonstrates the random variation (i.e. measurement error) in score *within* persons (see Box 1 for the method to calculate the MDC). In patients with conservatively managed stage 5 CKD, Erez et al. (2016) found an MDC of 4.2 and 7.0 for the SF-36 physical and mental HRQOL, respectively.[19] Using these thresholds in our example, the observed change of 5 points for physical HRQOL is larger than the MDC and can therefore be considered a statistically significant change. The observed change of 2 points in mental HRQOL is smaller than the MDC and can therefore not be distinguished with 95% confidence from no change – i.e. the change in mental HRQOL may be due to random variation and thus cannot be considered a true change.

Taken together, the MDC helps with the interpretation of PROM scores over time by distinguishing real changes from what is probably random variation. Although some literature is available[19, 20], more research on MDC is needed to facilitate interpretation of changes in PROM scores for different PROMs and in different patients and settings within nephrology care.[21]

Minimal important change (MIC)

If the observed change in our example of 5 points on physical HRQOL is likely a true change, can we than assume that this change is relevant to patients? And, if a decrease of 2 points does not demonstrate a real change in mental HRQOL, can we then also assume that this change is not meaningful for patients? To answer this question we need to know whether the observed change is larger than the minimal important change (MIC) or minimal clinically important change, in the literature also referred to as the minimal (clinically) important difference. MIC has been defined as "the smallest change in score in the construct to be measured which patients perceive as important".[8]

There are several methods for estimating the MIC, some of which are briefly discussed in Box 2. The MIC is not a fixed characteristic of a PROM and can vary across populations and settings. For example: characteristics of the population (e.g. mild or severe conditions), the direction of change (i.e. improvement or deterioration) and the study design and analysis used to estimate MIC (e.g. different anchors or definitions of importance) can influence the MIC.[8] Some literature is available that can provide a cautious indication of the MIC of some PROMs (e.g. SF-36) that might be used in nephrology care.[19, 22] However, in order to interpret changes in PROM scores clearly, more information is needed about the MIC in patients with

Box 1. Measuring minimal detectable change (MDC)

The MDC is a statistical parameter based on the measurement error (Standard Error of Measurement; SEM). The MDC can be determined in individuals who have not changed using a test-retest design, and can be calculated using the following formula: $1.96 * SD_{change,}$ which equals $1.96 * \sqrt{2} * SEM$.[8]

Box 2. Measuring minimal important change (MIC)

The MIC can be assessed using an anchor-based approach, for which several methods exist. In the literature also distribution-based approaches have been described[23]; however, these methods do not involve the *importance* of change and are therefore considered less suitable. In this box, we briefly touch upon the most common (anchor-based) methods to define MIC.

With an anchor-based approach the MIC is determined by comparing the changes in the PROM score to another measure that defines a clinical relevant change (i.e. the anchor). For PROMs usually the patient's general rating of change serves as an anchor, in which the minimal relevant change is explicitly defined by the patient.[8, 23]

A relatively easy method to determine the MIC is the mean change method. With this method the MIC is defined as the mean change in PROM score in patients who consider themselves to be minimally importantly changed, according to the anchor (e.g. in patients who rate their health as 'slightly improved').[8, 23]

Another method to determine the MIC is by use of receiver operating characteristic (ROC) analysis. The method is similar to the method known from diagnostic test research, whereby the PROM score is considered the diagnostic test and the anchor serves as a gold standard. The optimal ROC cutoff point gives the smallest chance of misclassifying importantly improved and not-improved patients and is therefore considered the MIC.[8, 24]

Furthermore, predictive modeling can be used. The outcome in this analysis is being either improved or not improved, which is defined based on the anchor. The change in PROM score is used as the predictor variable. The MIC is then determined using logistic regression analysis and is defined at the point where the change in PROM score is associated with a likelihood ratio of 1. An example of this method has been described in detail by Terluin et al. (2015).[24]

CKD in different stages and settings, and receiving different treatments.[21]
In patients with conservatively managed stage 5 CKD, Erez et al. (2016) report a MIC of 6.3 for the SF-36 score on physical HRQOL and 8.7 for the SF-36 score on mental HRQOL.[19] Comparing these thresholds to the observed changes in scores in our example of 5 and 2 for physical and mental HRQOL, respectively, shows that both observed scores are smaller than the MIC and are thus, on average, not considered important by patients. This example can be seen as a desirable situation: although statistically there is a decline in physical HRQOL, patients most likely do not perceive it as a relevant deterioration in their HRQOL.
However, the MIC gives an indication of what is *on average* considered important by an individual and should therefore be considered as a probability-threshold to interpret individual changes: if an individual change is larger than the MIC, the probability that this change is perceived important by the patient is greater than the probability that this change is perceived as not important.[25] The fact that the interpretation of the MIC involves probabilities, also indicates that this threshold may not apply to all individuals and that patients differ in which change they perceive as important. Therefore, it may be of added value to discuss the changes to gain insight into what is perceived important by the individual. On the other hand, the MIC may also facilitate the conversation, for example: it may be informative to the patient to explain which change in HRQOL may be expected (e.g. after kidney transplantation) and whether this change is, on average, considered important by patients.
Taken together, the results from our example can be considered positive with regard to both the MIC and the MDC: the MIC is larger than the MDC (6.3 > 4.2 and 8.7 > 7.0 for physical and mental HRQOL, respectively[19]) and thus, both the physical and mental HRQOL scales of the SF-36 seem to be able to detect changes that are, on average, important to patients. If the MIC would be smaller than MDC, the PROM may not be able to distinguish with high certainty relevant changes from random variation. Consequently, important changes might be missed and it may thus be advisable to use a different PROM or to improve the initial PROM in such way that it has a smaller MDC (i.e. by reducing the measurement error), for purposes where a high certainty is important (e.g. evaluation of treatment strategies).

Response shift

Another concept that is important for the interpretation of PROM scores is response shift, which refers to a change in the *meaning* of one's evaluation of the PRO (e.g. HRQOL) over time. This means that patients' answers to PROM questions change over time, not only because their health or HRQOL has changed, but

also because they might have changed their perception on what health or HRQOL means to them. For example: when Jason (male, 62y) started dialysis treatment, he experienced a deterioration in his health condition. Jason had to deal with vascular access problems and anemia, and it took several months to reach a hemoglobin level within the target range. Starting dialysis also had a major impact on his daily life: the sudden change in his schedule affected his ability to work and to participate in social activities. One might expect that such changes would impact Jason's HRQOL. However, contrary to what one might expect, after a couple of months Jason reported a HRQOL that was only slightly lower compared to his HRQOL at the start of dialysis. In this example a 'response shift' has occurred, that has been defined as "a change in the meaning of one's self-evaluation, which can be a result of *recalibration*, *reprioritization* and/or *reconceptualization* of the PRO".[9] Below these response shift inducing concepts are described and illustrated by means of Jason's example.

Recalibration refers to a change in an individual's frame of reference. In the example of Jason, his daily schedule and social life have changed considerably: since Jason started with dialysis treatment, he became more engaged in social comparison by talking to and sharing experiences with other patients treated with dialysis. Insights into the experiences of other patients, changed Jason's internal definition (i.e. his reference standard) of a poor HRQOL and consequently, Jason rates the HRQOL he had when he started dialysis higher now than he did before. Thus, new information and experiences can lead to a change in where a person positions himself on the scale, i.e. recalibration.

Reprioritization refers to a change in personal values. In Jason's case, acceptance of not being able to work and positive experiences with peer support could have encouraged Jason to shift his focus towards other aspects in life and set new life goals. Prior to dialysis, Jason mainly focused on professional accomplishments but after starting dialysis treatment, family relationships and being able to help others became more important to Jason. This illustrates how experiences can change people's self-evaluation and the value of certain aspects in life, and thus in *the extent to which aspects contribute* to a PRO, i.e. reprioritization.

Reconceptualization is a redefinition of the concept of interest. In the example of Jason, this could mean that his personal meaning of HRQOL has changed. By accepting the new daily routine and by appreciating a different way of participation in society, Jason may have realized that other factors determine his HRQOL. For Jason, being able to offer support to less fortunate peers contributes to a good HRQOL and having a certain employment status does no longer determine his HRQOL and consequently, his definition of HRQOL has changed. Hence, new ex-

periences can induce a change in *which aspects* contribute to a PRO and thus in one's definition of the PRO, i.e. reconceptualization.

Changes in internal standards, personal values or conceptualization of PROs may result in a response shift and thus in an experienced HRQOL that differs from what would be expected based on one's change in clinical health status, that is, for instance, based on clinical parameters (i.e. a decline in health status does not automatically imply a decrease in HRQOL). Changes may be induced by certain health- or life-changing events (e.g. getting a diagnosis, the start of a treatment or the loss of a loved one) and can also occur more gradually over time, for instance in chronic diseases.[9, 26-28] It is proposed that changes in health or in life may interact with the patient's characteristics (e.g. personality) and with mechanisms such as coping and social comparison, and consequently influence response shift.[9]

In the past decade, response shift has been investigated particularly in HRQOL research, but can occur in any PRO and when using any PROM as they all concern subjective self-evaluations. Nevertheless, PROs or PROMs that leave more room for personal interpretation are more sensitive to response shift compared to PROs or PROMs that are more unambiguously defined. For example: the question 'How is your sleep quality in general?' requires more consideration and evaluation from the patient than the question 'In the past week, did you sleep through the night without interruptions?', and the first question is therefore more prone to different interpretations over time.[8]

Response shift can complicate the interpretation of PROM scores over time. Therefore, it is important to know that this phenomenon exists, as it may explain unexpected findings (e.g. a stable HRQOL while clinical outcomes clearly show a deterioration in health). Response shift itself may also be a treatment goal, for instance in a treatment aimed at improved coping and self-management. Herein, response shift provides insight into the ability to adapt to a certain change in health. Furthermore, at the individual patient level, further investigation of and discussion about changes in internal standards, values and conceptualizations may help to interpret the patient's scores and guide decision-making.[28]

At a group level, it may also be informative to gain insight into response shift for instance by comparing treatment effects to inform decision making.[26] For example: let's compare HRQOL scores of patients treated with hemodialysis (HD) and peritoneal dialysis (PD) at several time-points during the first year of treatment (Figure 1). Theoretically, one may expect that HD impacts health status (e.g. based on clinical parameters) and HRQOL more severely compared to PD (e.g. due to the hospital visit 3 to 4 times a week). However, it is possible that PD patients will try to maintain their old way of life, while HD patients will try to adapt to their treat-

ment and to their new life. This may result in larger changes in internal standards, values and conceptualizations in HD patients compared to PD patients. As a result, HD patients may perceive a better HRQOL after some time (e.g. T2 in Figure 1), despite having a lower health status compared to PD patients. Such information is important for patients and professionals when drawing conclusions about treatment effects.

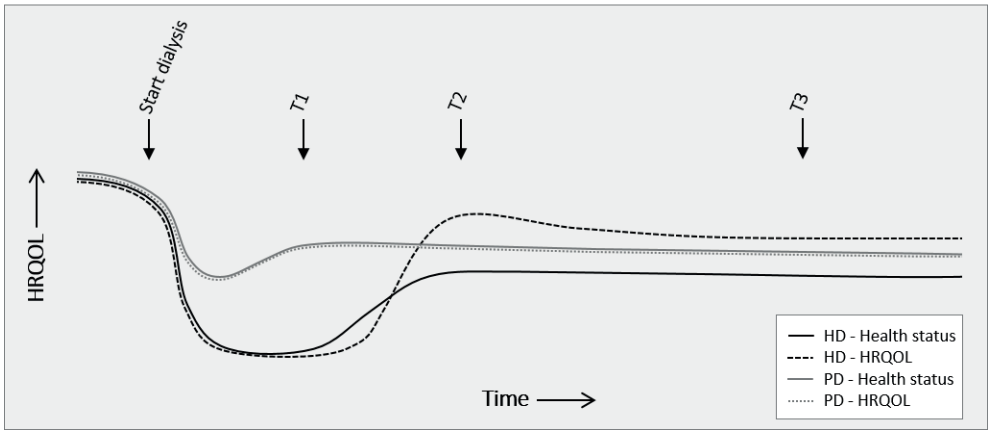


Figure 1. Theoretical example of trajectories of health status and HRQOL in patients receiving HD and PD. *A response shift occurs in the HD patient between T1 and T2. Abbreviations: HRQOL, health-related quality of life; HD, haemodialysis, PD, peritoneal dialysis.*

Furthermore, information about PRO-trajectories over time is also important when evaluating a patient's treatment, for example: the time-point at which the PRO was assessed could be informative to the interpretation of the PROM score.[26] Based on the trajectory comparison between HD and PD in Figure 1, different conclusions can be drawn, depending on the moment PROs are measured (start of dialysis, T1 or T2/T3). This example shows that a response shift may also occur later in the trajectory (e.g. between T1 and T2 in HD), and not directly after the life-changing event (e.g. start of dialysis).

Insight into the size and direction of the response shift can be informative, not only to explain unexpectedly small (or large) changes in PROM scores, but also to gain insight into the psychological change that may have occurred and the patient's ability to adapt. Several methods exist to determine response shift[29], some of which are briefly discussed in Box 3.

Box 3. Measuring response shift

Several methods exist to assess whether, how and to what extent response shift occurred. Barclay-Goddard et al. (2009) provided an overview of the methodologies to address response shift.[29] In this box, we briefly highlight some of the main approaches.

The most commonly used method is the *then-test*. In this method, the patient is asked to complete a PROM about his health status at two time-points, for instance at baseline (pre-test) and after 6 months (post-test). Additionally, the patient is asked at the post-test time-point to also complete the PROM for his health status at baseline (then-test). Since both the post-test and the then-test are completed at the same time-point, it may be assumed that the patient applied the same standards, values and concepts. Therefore, response shift can be assessed by comparing the pre-test and the then-test, and the difference between the post-test and then-test gives the response shift adjusted change (Figure 2).[8, 29]

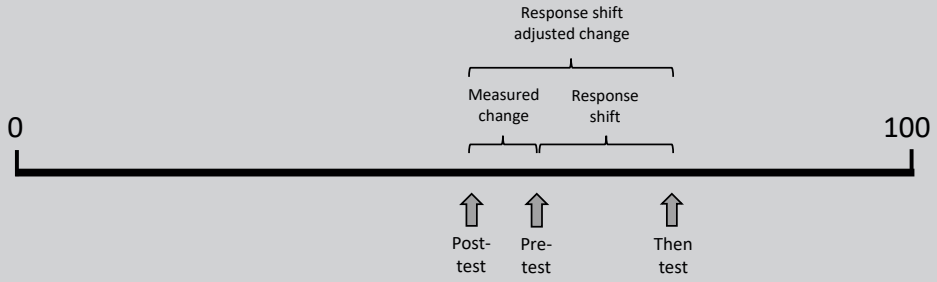


Figure 2. Then-test

The then-test has also been applied in combination with *qualitative methods* (e.g. using an interview) to explore response shift.[30] An advantage of combining these methods is that both numerical value of the response shift (using the then-test) and in-depth insight into the patient's thoughts and considerations regarding his standards, values and concepts are assessed. Qualitative methods can also be applied independently to investigate mechanisms of reconceptualization, reprioritization and recalibration that induce response shifts, as was done by Elliott et al. (2014) in dialysis patients.[28]

Another method to gain insight into changes in the patient's standards, values and concepts is by the use of a *questionnaire* that enables patients to define their own meaning of the construct (e.g. HRQOL), such as the Schedule for the Evaluation of Individual Quality of Life (SEIQOL).[30, 31] Changes over time in the

patient's reference standard, or in which and to what extent domains contribute to the patient's HRQOL may indicate a response shift.
Furthermore, response shift can be investigated using a *statistical approach*, such as confirmatory factor analysis. With this method, the three response shift inducing concepts can be identified by comparing the factor structure of the PROM pre- and post-measurement, namely: recalibration (apparent from a mean change in the variables), reprioritization (by means of a change in importance – i.e. factor loadings – of domains) and reconceptualization (by means of a change in the number of identified domains).[8, 29]

Conclusion

PROMs are instruments to assess aspects of the patient's perceived health, such as HRQOL or symptom burden. Different types of PROMs exist and knowledge about the characteristics of the PROM is necessary to interpret PROM scores and change scores. Information about the average and distribution of PROM scores in a reference population or in comparison to more familiar outcomes (e.g. laboratory measures) are indispensable to interpret and get used to PROM scores. Furthermore, the MDC and MIC are important to inform us about statistically and clinically relevant changes, respectively. Besides, one must be aware that response shift may occur, which may explain unexpectedly small (or large) changes in PROM scores. Finally, communication is important to interpret individual PROM scores; the best manner to interpret individual PROM scores and changes in PROM scores is through a discussion between the patient and the healthcare professional, in which the measures discussed in this paper (i.e. MDC, MIC and response shift) may have a facilitating role. Ideally, such measures are integrated into a dynamic report with individual PROM scores over time, enabling both patients and professionals to easily oversee which outcomes require attention and possibly intervention, and to evaluate treatment strategies at individual level. This will potentially increase the usability of PROMs in nephrology care for both patients and healthcare professionals.

References

1. Porter ME. What is value in health care? *The New England journal of medicine* 2010;363(26):2477-81

2. Black N. Patient reported outcome measures could help transform healthcare. *BMJ (Clinical research ed)* 2013;346:f167-f

3. van der Willik EM, Hemmelder MH, Bart HAJ, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. *Clinical Kidney Journal* 2021;14(6):1535–44

4. Noonan VK, Lyddiatt A, Ware P, et al. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 3: patient-reported outcomes can facilitate shared decision-making and guide self-management. *J Clin Epidemiol* 2017;89:125-35

5. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* 2016;34(6):557-65

6. Kotronoulas G, Kearney N, Maguire R, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. *J Clin Oncol* 2014;32(14):1480-501

7. 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

8. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide. Cambridge University Press: 2011;

9. Sprangers MA, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med* 1999;48(11):1507-15

10. Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *J Pain Symptom Manage* 2004;27(3):226-40

11. Dobbels F, Moons P, Abraham I, Larsen CP, Dupont L, De Geest S. Measuring symptom experience of side-effects of immunosuppressive drugs: the Modified Transplant Symptom Occurrence and Distress Scale. *Transpl Int* 2008;21(8):764-73

12. Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Development of the kidney disease quality of life (KDQOL) instrument. *Qual Life Res* 1994;3(5):329-38

13. Kidney Disease Quality of Life Instrument (KDQOL) Santa Monica, CA, RAND Health Care. https://www.rand.org/health-care/surveys_tools/kdqol.html. Accessed July 17, 2020

14. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-33

15. van der Willik EM, Meuleman Y, Prantl K, et al. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease - a four-phase mixed methods study. *BMC nephrology* 2019;20(1):344-

16. Hays RD, Kallich JD, Mapes DL, et al.: Kidney Disease Quality of Life Short Form (KDQOL-SF), Version 1.3: A Manual for Use and Scoring, Santa Monica, CA, RAND, 1997 Contract No.: P-7994

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

18. Mols F, Pelle AJ, Kupper N. Normative data of the SF-12 health survey with validation using postmyocardial infarction patients in the Dutch population. *Qual Life Res* 2009;18(4):403-14

19. Erez G, Selman L, Murtagh FE. Measuring health-related quality of life in patients with conservatively managed stage 5 chronic kidney disease: limitations of the Medical Outcomes Study Short Form 36: SF-36. *Qual Life Res* 2016;25(11):2799-809

20. Ware J, Ma K, Keller SD. SF-36 Physical and Mental Health Summary Scales: a User's Manual. 1993;8:23-8

21. 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-e

22. Finkelstein FO, van Nooten F, Wiklund I, Trundell D, Cella D. Measurement properties of the Short Form-36 (SF-36) and the Functional Assessment of Cancer Therapy - Anemia (FACT-An) in patients with anemia associated with chronic kidney disease. *Health Qual Life Outcomes* 2018;16(1):111

23. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56(5):395-407

24. Terluin B, Eekhout I, Terwee CB, de Vet HCW. Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. *Journal of Clinical Epidemiology* 2015;68(12):1388-96

25. de Vet HC, Terluin B, Knol DL, et al. Three ways to quantify uncertainty in individually applied "minimally important change" values. *J Clin Epidemiol* 2010;63(1):37-45

26. Hartog ID, Willems DL, van den Hout WB, et al. Influence of response shift and disposition on patient-reported outcomes may lead to suboptimal medical decisions: a medical ethics perspective. *BMC Med Ethics* 2019;20(1):61

27. Barclay-Goddard R, King J, Dubouloz CJ, Schwartz CE. Building on transformative learning and response shift theory to investigate health-related quality of life changes over time in individuals with chronic health conditions and disability. *Arch Phys Med Rehabil* 2012;93(2):214-20

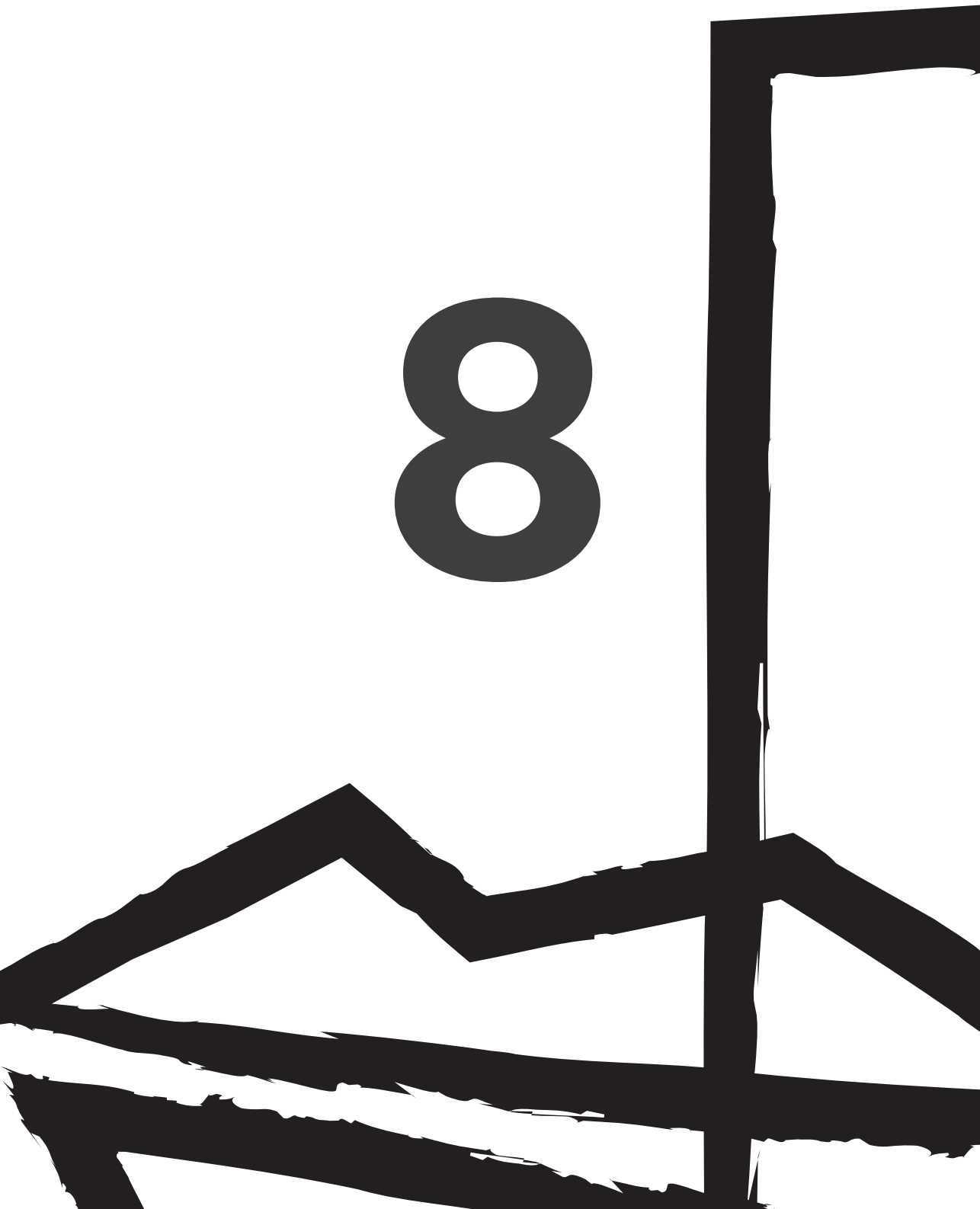
28. Elliott BA, Gessert CE, Larson PM, Russ TE. Shifting responses in quality of life: people living with dialysis. *Qual Life Res* 2014;23(5):1497-504

29. Barclay-Goddard R, Epstein JD, Mayo NE. Response shift: a brief overview and proposed research priorities. *Qual Life Res* 2009;18(3):335-46

30. Westerman MJ, Hak T, Sprangers MA, Groen HJ, van der Wal G, The AM. Listen to their answers! Response behaviour in the measurement of physical and role functioning. *Qual Life Res* 2008;17(4):549-58

31. O'Boyle C. The Schedule for the Evaluation of Individual Quality of Life (SEIQoL). *International Journal of Mental Health* 1994;23:3-23

8



Discussing results of patient-reported outcome measures (PROMs) between patients and healthcare professionals in routine dialysis care: a qualitative study

Esmee M. van der Willik, Jet Milders, Johannes A.J. Bart, Willem Jan W. Bos, Frans J. van Ittersum, Marc A.G.J. ten Dam, Marc H. Hemmelder, Friedo W. Dekker, Yvette Meuleman

BMJ Open. 2022;12:e067044.

Abstract

Objectives: Patient-reported outcome measures (PROMs) provide insight into patients' experienced health and needs, and can improve patient-professional communication. However, little is known about how to discuss PROM-results. This study aimed to provide in-depth knowledge of patients' and healthcare professionals' experiences with and perspectives on discussing PROM-results as part of routine dialysis care.

Design: A qualitative study was performed using an interpretive description approach. Individual semi-structured interviews were conducted with 22 patients and healthcare professionals. Interviews focused on general and specific situations (e.g. addressing sensitive topics or when no medical treatment is available). Interviews were transcribed verbatim and analysed inductively using thematic analysis.

Setting: Participants were purposively sampled from 8 dialysis centres across The Netherlands.

Participants: Interviews were conducted with 10 patients receiving dialysis treatment and 12 healthcare professionals (nephrologists and nurses).

Results: Patients and healthcare professionals provided practical guidance for optimal discussion about PROM-results. First, patients and healthcare professionals emphasized that PROM-results should always be discussed and indicated how to create a suitable setting, adequately prepare, deal with time constraints and use PROMs as a tool for personalised holistic consultations. Second, patients should actively participate and healthcare professionals should take a guiding role. A trusting patient-professional relationship was considered a prerequisite and patient-professional interaction was described as a collaboration in which both contribute their knowledge, experiences and ideas. Third, follow-up after discussing PROM-results was considered important, including evaluations and actions (e.g. symptom management) structurally embedded into the multidisciplinary treatment process. These general themes also applied to the specific situations, for example: results should also be discussed when no medical treatment is available. Though, healthcare professionals were expected to take more initiative and a leading role when discussing sensitive topics.

Conclusions: This study provides insight into how to organize and conduct conversations about PROM-results and lays the foundation for training healthcare professionals to optimally discuss PROM-results in routine nephrology care. Further research is needed to provide guidance on follow-up actions in response to specific PROM-results.

Introduction

People with advanced chronic kidney disease (CKD) experience numerous physical and emotional disease-related symptoms, which have a major impact on their lives and health-related quality of life (HRQOL).[1-3] Although HRQOL and symptom burden have been regarded as highly important by patients and healthcare professionals[4-6], these outcomes often remain undiscussed and un(der)treated in regular practice.[3, 6, 7] This may in part be explained by the fact that patients do not share everything by themselves, for instance because they do not talk easily about sensitive topics or assume their symptoms are not CKD-related or cannot be treated.[8-10] Additionally, healthcare professionals may consider it challenging to inquire about patients' wide range of symptoms and needs, for example due to time limitations, and remain largely unaware of symptom burden.[7, 11]

Patient-reported outcome measures (PROMs) are questionnaires that assess aspects of patients' perceived health, such as HRQOL and symptom burden, and are reported by the patients themselves.[12] PROMs can play an important role in solving the under-recognition of patient-relevant outcomes by providing insight into and facilitating communication about patients' HRQOL, symptoms and needs.[2, 7, 13-16] Integration of PROMs into routine care has the potential to contribute to a more person-centred approach, incorporating patients' experiences and needs complementary to traditional clinical measures. This provides patients and healthcare professionals with a more complete perspective on how the patient is really doing and could enhance shared decision-making about treatment choices.[2, 13-18]

In the Netherlands, PROMs are implemented into routine dialysis care since 2018.[2] Patients are invited 1-2 times a year by their care team to complete PROMs on symptom burden and HRQOL.[2, 19] After completing PROMs, patients and healthcare professionals receive a patient's individual PROM-report and are encouraged to discuss PROM-results during the upcoming consultation.[2] Previous research suggests that using PROMs facilitates patient-professional communication by providing overview and insights, and by prompting discussions about topics that are important to patients.[2, 20] However, it also reveals several challenges in organization and conversations about PROM-results, including: lack of incorporation of PROM-results into electronic health records, limited time during consultations, and lack of knowledge on how to interpret and discuss PROM-results.[21-25] PROMs are relatively new in routine dialysis care, and although research emphasizes the need to discuss PROM-results, little is known about the most optimal way to have this conversation.[2, 14, 26] Therefore, this study aims to provide in-depth knowledge of patients' and healthcare professionals' experiences with and perspectives on discussing PROM-results as part of routine dialysis care.

Methods

Design and participants

A qualitative study was performed using an interpretive description approach. An interpretive description methodology captures the experiences and perspectives from practice, and intends to gain a deep understanding of the topic at hand and to generate knowledge that can enhance clinical practice.[27-29] Individual semi-structured interviews were conducted between April and September 2021, with patients receiving dialysis treatment and healthcare professionals involved in dialysis care – all experienced with completing and discussing PROM-results. Purposive sampling was applied to capture perspectives from a heterogeneous patient-sample based on age, gender and dialysis modality (haemodialysis [HD] or peritoneal dialysis [PD]) and healthcare professional-sample based on age, gender and occupation (nephrologists and nurses). Participants were recruited from 8 dialysis centres across The Netherlands (Provinces Gelderland, Noord Brabant, Noord Holland, Overijssel, Zuid Holland, Utrecht) participating in the Dutch PROMs-registry. Healthcare professionals were recruited through the researcher (EvdW) and patients through their care team, until data saturation was reached (i.e. when little or no new concepts arose from subsequent interviews).[30] All participants received written study-information and provided written informed consent. The study was approved by the Medical Research Ethics Committee (METC-LDD, N20.097). Recommended guidelines and checklists (e.g. the COnsolidated criteria for REporting Qualitative research [COREQ]) were used to conduct and report this study.[31, 32]

Interview and data collection

An interview guide (Supplementary Table S1) was developed based on previous research[2, 6, 14, 23, 26], discussion among the research team, and in collaboration with four patients from regional and national kidney patients associations (Diavaria and NVN). Individual interviews were performed through videocalls (Zoom)[33] or by telephone for patients without access or skills to use electronic devices. Interviews were carried out by two female researchers: EvdW (health scientist, PhD-candidate) and JM (medical scientist, PhD-candidate), who were trained by an experienced qualitative researcher, YM (medical psychologist, PhD). No relationship existed between participants and interviewers. Interviews were digitally recorded and field notes on non-verbal communication were made.

Prior to interviews, participants completed a brief questionnaire to collect sociodemographic information. Additionally, patients completed the PROMs used in Dutch dialysis care[2]: 12-item Short Form Health Survey (SF-12)[34] to assess generic

HRQOL and Dialysis Symptom Index (DSI)[19, 35] to assess CKD-specific symptom burden. Patients received their personal PROM-report and healthcare professionals received a mock PROM-report.

During interviews, participants were asked about: experiences with and perspectives on optimal ways to discuss PROM-results (part A), and perspectives on dealing with specific situations that might complicate discussions about PROM-results (part B). All questions were open-ended and responses were further explored using additional questions and probes. Participants received a summary of the main findings at the end of the study.

Analysis

Interviews were transcribed verbatim. Transcripts were analysed inductively using thematic analysis.[36, 37] Open line-by-line coding of transcripts was done by two researchers (EvdW, JM). Coding strategies and interpretations were discussed with a third researcher (YM) to ensure consistency. Subsequent analyses were performed in two parts. Part A: experiences and perspectives on optimal ways to discuss PROM-results were analysed. Axial coding was applied by constant comparison, grouping similar concepts into themes and organizing them hierarchically. Preliminary themes were discussed within the multidisciplinary research team to ensure triangulation. Part B: data about specific situations were compared with identified themes from part A; in other words, did the same or different themes apply to these specific situations? Analyses were performed using Atlas.ti (GmbH, Berlin, version 9). Finally, illustrative quotes were selected and translated from Dutch to English by one researcher (JM) and translated back (i.e. from English to Dutch) by a second researcher (EvdW) to ensure accuracy.

Patient and public involvement

Four patients from regional and national kidney patients associations (Diavaria and NVN) were involved in the development of the interview guide. One patient representative (HB) is part of the research team and was involved in designing the study, interpreting the results and revising the manuscript.

Table 1. Participant [a] and interview characteristics

	Patients (n=10)	Healthcare professionals (n=12)
Gender, male	5 (50.0)	4 (33.3)
Age, years	62.1 ± 14.5	52.0 ± 7.0
Marital status		
Married/partnered	7 (70.0)	10 (83.3)
Widowed/single	3 (30.0)	2 (16.7)
Ethno-cultural group [b], Dutch	10 (100.0)	12 (100.0)
Educational level [c]		
Low	2 (20.0)	
Middle	5 (50.0)	
High	3 (30.0)	12 (100.0)
Employment		
Full-time	1 (10.0)	8 (66.7)
Part-time	1 (10.0)	4 (33.3)
No, retired	5 (50.0)	
No, disabled due to health	3 (30.0)	
Years since diagnosis	17.5 (10.8-24.8)	
Years since dialysis initiation	4.5 (2.9-16.5)	
Dialysis modality		
HD, centre	6 (60.0)	
HD, home	1 (10.0)	
PD	3 (30.0)	
Dialysis sessions per week		
2 sessions (HD)	1 (10.0)	
3 sessions (HD)	4 (40.0)	
4 sessions (HD)	2 (20.0)	
7 sessions (PD)	3 (30.0)	
PROM scores		
Physical HRQOL (0-100)	37.6 ± 9.0	
Mental HRQOL (0-100)	51.3 ± 10.3	
Number of symptoms (0-30)	8.7 ± 4.1	
Symptom burden score (0-150)	22.0 (15.8-27.3)	

Healthcare profession		
Nephrologist		8 (66.7)
Nurse practitioner		2 (16.7)
Nurse		2 (16.7)
Years involved in dialysis care		20.2 ± 7.7
Experience with PROMs-results, times discussed [d]	2.0 (1.0-3.0)	20.0 (12.0-35.0)
Interview duration, minutes	63.2 ± 16.5	58.4 ± 13.5
Interview setting		
Face-to-face (videocall)	6 (60.0)	12 (100.0)
Telephone	4 (40.0)	

Descriptive statistics were conducted using SPSS (IBM, Armonk, NY, USA, version 25.0): participant characteristics were presented as number (proportion), mean ± standard deviation or median (interquartile range) where appropriate.
[a] Additional information regarding non-participation: two more female nurses provided informed consent, but cancelled the interview and withdrawn from the study due to too busy schedules (data not shown).
[b] Self-reported ethno-cultural group: "What ethnic group do you consider yourself belonging to?"
[c] Educational level according to International Standard Classification of Education (ISCED) levels 2011, classified as low: primary, lower secondary or lower vocational education; middle: upper secondary or upper vocational education; high: tertiary education (college/university).
[d] During the interview, participants were asked to estimate the number of times they had discussed PROM-results as part of routine dialysis care.
Abbreviations: HD, haemodialysis; PD, peritoneal dialysis; PROM, patient-reported outcome measure; HRQOL, health-related quality of life.

Results

Participant and interview characteristics

All characteristics are shown in Table 1. In total, 22 interviews were conducted with 10 patients and 12 healthcare professionals. Interviews lasted on average 61±15 minutes. Patients' mean age was 62±15 years, half of them was male, seven patients received HD and three PD. Eight nephrologists and four nurses participated, with a mean age of 52±7 years and four of them were male.

Part A: Experiences and perspectives on optimal ways to discuss PROM-results

Three overarching themes were identified and will be discussed below, with corresponding subthemes and illustrative quotations. Additional quotations are provided in Table 2.

Table 2. Illustrative quotations of patients and healthcare professionals for identified themes and subthemes (part A)

1. Organization and basic principles of discussing PROM-results
1.1 Suitable setting for discussing PROM-results
That she says "well, we will plan a consultation and we will actually discuss those subjects". That makes that you come to the consultation with a different attitude. *(Patient, age: 70-80y)*
Maybe it should be told more clearly in advance that the PROMs will be discussed during the annual consultations, and that the PROMs determine what will be discussed during the consultations. *(Nephrologist, age: 50-60y)*
What I find difficult with PROMs is that patients also participate in research and already get so much questionnaires. […] And because they get the same questions from different angles, they get the idea that nothing is happening with it. *(Nurse, age: 50-60y)*
I think the annual consultation is actually ideal, because then you are alone with the patient instead of in the shared dialysis room. It is easier then to discuss such sensitive topics. *(Nephrologist, age: 50-60y)*
Yes, I believe that the nephrologist is the right person. Why? Because they know the person. And two, they know the numbers of the patient. The lab values, because they have them, so they can give their medical opinion on it. *(Patient, age: 40-50y)*
A dialysis patient is never actually sick by themselves, their surrounding suffer from it too. And this is also an opening to talk about that. *(Nephrologist, age: 40-50y)*
Well once a year, I think, is more than enough. Unless someone indeed has a lot of issues, maybe then it is interesting to do it more often. *(Patient, age: 40-50y)*
1.2 Preparations by patients and healthcare professionals
I prepare myself for this consultation, but I also expect that from you [the healthcare professional]. *(Patient, age: 70-80y)*
Well, maybe, they [patients] underestimate their input, also in the preparation. […] So, like "this may help you to prepare for the conversation and it would be useful if you select the point of attention that you would like to discuss, that you find important". Then you put the balance and interests a bit towards the patient. But my experience is that my patients don't come into the consultation room that way… *(Nephrologist, age: 40-50y)*

For people who may find it difficult or not dare to say it at the moment: write it down beforehand. Write down what bothers you and what you want to talk about. *(Nephrologist, age: 40-50y)*

Leading up to the conversation I have a piece of paper at hand and write down all the questions that come to my mind. She [nephrologist] laughs when I take out the piece of paper and says "let's go". *(Patient, age: 70-80y)*

The conversations are better, because you already know the patient's experience. You are prepared, you can better come to a solution and the patient feels better heard. *(Nephrologist, age: 30-40y)*

1.3 Dealing with time constraints

We are always puzzling with our time, so we have to prioritize. (Nurse, age: 50-60y)

The advantage is immediately also a disadvantage, because you now get a complete picture, but then you have to make a selection again of the things we can do something about and what we are going to work on. (Nephrologist, age: 50-60y)

Look, now it is also a time issue, a certain time has been set aside for it and then you cannot go through the entire questionnaire in detail. (Patient, age: 70-80y)

It is also much more efficient for me as a doctor. […] For patients that did not complete the PROMs, I do not ask 30 questions about symptoms, there is no time for that. (Nephrologist, age: 30-40y)

1.4 PROMs as tool for personalised holistic consultations

Well, I find it extremely insightful and useful. New things come up, even when you see someone weekly and ask them how they are doing. If you look at the Dialysis Symptom Index especially, sometimes complaints and things come up that bother people a lot, and which they otherwise would have never told about, and which I apparently have not asked them about specifically before. I think that is very valuable. *(Nephrologist, age: 50-60y)*

Well, it makes you think about your own situation. That is the positive thing of it. *(Patient, age: 70-80y)*

These [PROMs] also highlight questions that you, and maybe even a doctor, would not normally ask. *(Patient, age: 40-50y)*

I think it is a great addition to the conversation, because you now hear directly from the patient what he/she experiences. So it gives you the opportunity to primarily focus on what the patient reports as most burdensome. *(Nephrologist, age: 50-60y)*

Well, to me it is very important that it comes from the patient. Do not necessarily only look at the numbers. The colours can help you to identify where the biggest problems occur, but you should try to ask the patient what he thought of filling it in and what he thinks are the most important outcomes. *(Nurse, age: 50-60y)*

I use it as a topic of conversation, and do not show 'you score very low on this'. It feels different then, as if they get some sort of grade, like 'you did a bad job'. *(Nephrologist, age: 50-60y)*

I am not here for the kidney. I am here for the patient. *(Nephrologist, age: 30-40y)*

That you want to see the progression of PROMs over time. So those PROMs should be in the electronic health record, in a kind of dashboard, so that you can quickly see what is going on. *(Nephrologist, age: 50-60y)*

1.5 Always discuss PROM-results

You should not have them fill in the questionnaires and leave it at that. If that happens the questionnaires may not be filled in year after year. If you have been mentioning something for years and nothing is being done with it… Yes… *(Nurse, age: 50-60y)*

My role is to make clear everything can be discussed, that the patient feels invited to share complaints, that I am open to talking about that too. […] And then the patient decides how long and how extensively we talk about it. *(Nephrologist, age: 40-50y)*

There is nothing as annoying as coming in with questions and still leaving with questions. *(Patient, age: 60-70y)*

Maybe sometimes as a doctor you get that feeling of "there is nothing I can do about it, why would I discuss it then?", but for a patient it can be enough to know that they are not the only one, that it is part of the disease, or that it is because of their diabetes and vascular disease and dialysis. The doctor cannot do anything about it, but at least I understand why. *(Nephrologist, age: 40-50y)*

When it is continuously being made discussable it diminishes the burden I think. […] They can explain what is happening, why something is not possible and why something is not necessary, and what may be the consequences. Well, that can be very reassuring for the patient, I think. *(Patient, age: 70-80y)*

The nurse can often reassure you, or well, to the extent that the subject remains discussable. I think having a listening ear is just important. *(Patient, age: 60-70y)*

2. Roles of and interaction between patients and healthcare professionals

2.1 Patient's role: active participant

You want the patient to take a more directive stance, but we are not there yet. That will gradually become the case. I mean, the moment that they are more familiar with it and can compare with previous times, then I think it will go the way you want it to. *(Nephrologist, age: 60-70y)*

Well I think the nephrologist should [take the initiative]. I am the one who has ticked the boxes, but that is to indicate how I am doing. I do not see that I should then take the lead in the conversation and say "let's talk about the PROMs" or "I have filled this in for this and this reason". *(Patient, age: 70-80y)*

That he/she [the patient] initiates what to discuss, and also thought about ideas… on how problems can be solved, or what is supportive. Together, together searching for the best solution. That is, I think, also with compliance, when the patient feels he is cooperating and that you decided together how to continue, that it works better. *(Nurse, age: 50-60y)*

Well, things about dialysis and kidneys; she [healthcare professional] knows more about that. But everything beyond that, the complaints I have, then it often comes more from my side than before. *(Patient, age: 30-40y)*

But I was always the one that had the last word and I enjoy that because I am the boss of my own body and not the doctor, just to be clear. *(Patient, age: 40-50y)*

2.2 Healthcare professional's role: guide

Uhm yes, I think that as a healthcare professional I am in the lead, because the patient comes into the conversation with an open mind and does not feel a clear purpose in that conversation. So I think it is quite an open approach, in which I am more in the lead than the patient is. It is not that I am not open to the patient being more in control, but yeah, they do not see the purpose of it in advance. *(Nephrologist, age: 40-50y)*

And I notice that it is quite difficult to assign that role to the patients, because before you know it, if the patient does not pick it up, you are back in your old role. *(Nephrologist, age: 60-70y)*

The opening is then with the nephrologist, who says "I see that you have filled this in". That is how I would do it. Asking "How did you get that answer?", "What is on your mind" and " What is going on?". *(Patient, age: 70-80y)*

Because nowadays the doctor more often has an advisory role, instead of a decisive role. And yes… It has actually always been that way. We have always though that they know much better, so we will just do what they say. But in the meantime people are becoming more stubborn and it might also be useful to improve communication in that regard. *(Patient, age: 30-40y)*

I think it is good to articulate that you see what is happening, because you cannot decide for someone if they do not want to talk about it or if they just find it hard to talk about it. *(Nephrologist, age: 40-50y)*

I think sometimes we talk too much, but that above all we should also listen very carefully. We should also listen to non-verbal signals and uhm, really connect, try to feel, to understand what that patient wants from you. *(Nephrologist, age: 40-50y)*

2.3 Trusting patient-professional relationship

I think in that sense we do have a trusting relationship with the patient, where they have to feel the freedom to discuss everything. *(Nephrologist, age: 40-50y)*

How the doctor and patient interact with each other. I know nephrologists in the hospital and I even know what they have for dinner so to speak. That is how the communication is. But for the patient that is of course simply amazing. Because if something is up, you can bring it up much more easily. I think that that works much better for the patient, also in the treatment plan. *(Patient, age: 60-70y)*

Firstly, there must already be a trusting relationship between the healthcare professional and the patient. That is extremely important. I always think about how I would experience something. And if the doctor was a complete stranger to me, I would also think "well, never mind.". So I think trust is extremely important. *(Nurse, age: 50-60y)*

And I also expect that from the nephrologist, when you know each other already for 10-12 years. This guiding through the trajectory… until, yeah, when you don't need to go to the hospital anymore… *(Patient, age: 70-80y)*

There needs to be a relationship over a longer period, otherwise you cannot discuss these topics. […] Because she [nephrologist] can tell me 'you can always talk with social work', but then I feel like 'ok, but I have not known them for 10 years'. Of course, it could be [helpful]… but, the doctor knows me and my situation. *(Patient, age: 70-80y)*

Something was bothering me once, and immediately there was a social worker right in front me. I really do not want all that. *(Patient, age: 60-70y)*

This is much more human. How do you experience this, what does it mean for you and your environment? [...] So now when a PD nurse calls me by my name, I notice that. I like that. But I remain the patient and she remains the healthcare professional. The roles do not change, but the relationship becomes a bit more human. I experience that to be very pleasant. *(Patient, age: 70-80y)*

2.4 Patient-professional collaboration

It is actually always a two-way. Like, I myself or the doctor of course asks a question, and search for a solution… The doctor gives options and then I respond what I think is best, what fits me best, because I am still the boss of my own body and my lifestyle. *(Patient, age: 40-50y)*

Well I think just 50/50… Sometimes I have something to add and sometimes she has something to add. And yes, just have a conversation from person to person. *(Patient, age: 50-60y)*

Well, I think it is important that as a patient you do no put yourself in the underdog position. I think it is best if you feel a little bit equal to each other. In that manner you might actually remove some barriers for the patient. *(Patient, age: 60-70y)*

A new integration into your consultations, you have to find a mode in that. *(Nurse, age: 50-60y)*

I think the biggest barrier is that there are… yeah, I call them 'old-fashioned-doctors' or 'my-will-is-law-doctors'… That's were the biggest hurdle is, I think. *(Patient, age: 30-40y)*

Not everyone takes the floor. That also has to do with generations. Young people will do that, that is no problem. But older people, yeah, they find it difficult. *(Nephrologist, age: 60-70y)*

Sometimes there are cultural differences. Something that we may consider normal to talk about, can be something you do not talk about for them. I think that at some point there has to be respect for that too. *(Nurse, age: 50-60y)*

We have a multicultural group of patients. So there is a language barrier and yeah, then discussing something like sexuality with an interpreter in the middle… So certain topics may remain underrecognized, just because of the circumstances. *(Nurse, age: 50-60y)*

3. Follow-up on PROM-results discussions

3.1 Follow-up actions and evaluations embedded into treatment process

I think that action points should be formulated, so that is what I do. There should also be some sort of feedback towards the patient, so I write that down as well and I often bring that up during the dialysis visits. So I build in these kinds of moments of evaluation. *(Nephrologist, age: 30-40y)*

Uhm yes, those are often more supportive types of treatment. I believe that it is not about the major decisions surrounding dialysis and transplantation, but more focused on things like do we have to lower the phosphate levels, is the patient anaemic, should we do something about the fluid retention because the patient experiences shortness of breath? So it is more about the finetuning than the bigger picture. *(Nephrologist, age: 50-60y)*

My doctor often looks at what is important and what to discuss. And where… for them it is more like where the most benefit can be gained. *(Patient, age: 30-40y)*

Yeah, quality of life, I find more difficult to act upon. I actually look into what is the problem and is this something we can solve. […] And that has to do, I think, with that we as doctors like to do something. *(Nephrologist, age: 50-60y)*

And it doesn't have to be an answer that solves all my issues. That is not what I mean. More like, is there an answer? *(Patient, age: 40-50y)*

Ideally you would have a treatment or guideline available for every complaint, so you can treat the patient and evaluate whether the symptoms are decreasing. *(Nephrologist, age: 40-50y)*

3.2 Multidisciplinary process

Kidney patients often have a multitude of problems. I think that for a doctor, or for a nephrologist in this case, it is very important to acknowledge that there are other problems and to refer the patient if necessary. *(Patient, age: 60-70y)*

I am now kind of the playmaker that says "this should go there, and that should go there". *(Nephrologist, age: 50-60y)*

Look, when you are on dialysis the nephrologist is your central doctor. So he/she is the starting point from which you can go in different directions. *(Patient, age: 30-40y)*

I talk with colleagues, I email other healthcare professionals: can you answer this, is it ok if I do this or give that, so I ask for help. Discuss. *(Nurse, age: 50-60y)*

We have of course multidisciplinary consultations, with social work, dieticians, and sometimes we invite spiritual caretakers or a psychologist, or we invite the general practitioner to join online. And of course with the nurses and other colleagues, and then with that group we look into what we can improve and where we can adapt the treatment plan. *(Nephrologist, age: 50-60y)*

THEME 1: ORGANIZATION AND BASIC PRINCIPLES OF DISCUSSING PROM-RESULTS

Patients and healthcare professionals described experiences and advised on how to organize conversations about PROM-results, such as organizing a suitable setting, adequate preparation and dealing with time constraints. Additionally, they described that PROMs should be considered a tool for personalised holistic consultations and PROM-results should always be discussed.

1.1. Suitable setting for discussing PROM-results

Patients and healthcare professionals gave a univocal description of how conversations about PROM-results should be organized. First, PROM-results should be available on time: they recommended to invite patients to complete PROMs 2-6 weeks in advance and announce that PROM-results will be discussed, in what way and for what purpose, to ensure patients do not think "nothing is being done with it". Second, patients and healthcare professionals stated that PROM-results should be discussed face-to-face in a consultation room that ensures a safe and private setting. PROM-results were mostly discussed during annual consultations, which was considered very suitable. They advised that patients discuss PROM-results with the nephrologist or nurse who knows the patient best. Many recommended to invite the patient's partner or a close relative, as this may provide valuable insights for all attendees. Third, all patients and healthcare professionals considered a conversation about PROM-results once per year sufficient. Some patients and healthcare professionals suggested to schedule additional conversations when many symptoms or important changes in treatment or health occur.

1.2. Preparations by patients and healthcare professionals

All patients and healthcare professionals advised that both prepare for conversations about PROM-results. They suggested that patients think beforehand about their own results, whether they have questions, and what they would like to discuss and receive support on. They advised healthcare professionals to examine which topics emerge from the PROM-results and which treatment-options are available; this helps to inform patients and quickly get to what is important to them:

> *"You should not open the list when you are already there with your patient. Take some time to prepare it, to discuss it, and then it really is an addition to the annual consultation." (Nurse/50-60y)*

Additionally, healthcare professionals indicated that they always prepared discussions on PROM-results. Patients and healthcare professionals indicated that in practice patients often did not prepare, for example because it was not clear to patients that, when and how PROM-results would be discussed.

1.3. Dealing with time constraints

Most healthcare professionals and some patients experienced time constraints. Often, more topics emerged from PROMs than can be discussed during one consultation. Healthcare professionals indicated that they solve this by prioritizing and

focusing on the most important 3-4 topics, identified by patients or themselves. Some healthcare professionals indicated that PROMs also save time, because you get to the point more quickly: more information is gained, but with PROMs there is no need to spend time asking questions about symptoms that patients do not have:

> *"People who use PROMs are especially happy that they can work in a more focused way and that they do not lose time." (Nephrologist/50-60y)*

1.4. PROMs as tool for personalised holistic consultations

Many healthcare professionals and some patients argued that PROMs are a tool to improve communication during consultations:

> *"Discussing PROM-results is not an aim in itself, it is a tool to conduct the conversation well and to do justice to what is important to the patient." (Nephrologist/40-50y)*

Many patients and healthcare professionals stated that using PROMs during consultations helped to focus on what is important to patients. They shared that PROM-results can serve as a 'conversation-starter' to discuss topics that were otherwise not brought up – PROM-topics but also overarching topics such as treatment continuation or death. Additionally, some patients and healthcare professionals described that PROM-results can stimulate self-reflection and awareness in patients, which may result in better understanding and confirmation of their experienced health, although it can also be confronting:

> *"At times I find it difficult to hear that my condition is less good than that of a seventy-eighty year old. On the other hand, it is also something that I clearly run into: it really bothers me more than others, so it also validates that I didn't just make it up." (Patient/30-40y)*

Furthermore, many healthcare professionals and some patients indicated that the overview of PROM-results and comparison of scores with other dialysis patients, provides a sense of 'how someone is doing' and insight into 'what is normal'. Some of them emphasized that this comparison should only be used as a tool to better understand PROM-results, and not as a grade or treatment-target. Although most patients and healthcare professionals indicated that using PROMs provided a more complete picture, many healthcare professionals also mentioned technical

barriers: PROM-results are not yet integrated into electronic health records, which makes it more challenging to get a complete overview in combination with biomedical results and to evaluate changes over time. Implementation into electronic health records was considered an important next step to support personalised holistic consultations.

1.5. Always discuss PROM-results

Almost all patients and healthcare professionals indicated that discussing PROM-results is essential, and some even argued that PROMs should only be collected on condition that results are discussed:

> *"If you have indicated that you experience certain problems, then it must be discussed, otherwise such a questionnaire is useless." (Patient/60-70y)*

Many patients and healthcare professionals indicated that discussing PROM-results is important "to make clear that everything can be discussed, now or at any given moment", and that there is an opportunity to start a conversation about it. Furthermore, patients and healthcare professionals indicated that discussing PROM-results in itself, can already help: patients do not always expect or want something to be done about symptoms, but they do sometimes want to talk about it with their healthcare professional – to inform them or to feel seen, heard and understood by them:

> *"A listening ear from the doctor is most important. Whatever complaint you have, it can be small for the doctor, but big for you. That you are taken seriously." (Patient/60-70y)*

Additionally, patients and healthcare professionals indicated that explanation and clarity (e.g. about causes, prognosis and treatment-options for symptoms) is important for patients, because better understanding can help to accept and cope with the disease.

THEME 2: ROLES OF AND INTERACTION BETWEEN PATIENTS AND HEALTHCARE PROFESSIONALS

Patients and healthcare professionals described the patient's role as active participant and the healthcare professional's role as guide. They indicated that a trusting patient-professional relationship is a requirement and affects what conversations about PROM-results yields. They described the interaction as a collaboration and considered this the optimal approach.

2.1. Patient's role: active participant

Patients and healthcare professionals described the patient's role as active participant in the discussion. Healthcare professionals would like patients to take more control when it comes to PROM-results, but acknowledged that this is not easy for all patients. Contrary, most patients indicated that they usually wait for the healthcare professionals' initiative: even when they feel empowered enough to articulate their experiences and needs, they believe it is not their role to bring up PROM-results or to structure conversations:

> *"During a consultation, I think I will feel empowered enough to take the initiative. But still, the doctor needs to give you enough space to do something with it." (Patient/60-70y)*

Both patients and healthcare professionals indicated that, ultimately, patients decide whether or not something is discussed and follow-up actions are taken.

2.2. Healthcare professional's role: guide

Patients and healthcare professionals described a guiding role for healthcare professionals in taking initiative and providing structure in PROM-results conversations. Healthcare professionals are expected to ask patients to expand on their experiences and needs to gain a deeper understanding, and to facilitate prioritizing outcomes and actions. Some healthcare professionals indicated that they would prefer to be less steering in conversations and find it challenging to facilitate rather than direct conversations, especially when patients have a 'wait-and-see' attitude. Patients indicated they are used to the healthcare professionals' leading role and this also worked well in discussing PROM-results:

> *"The healthcare professional should facilitate and guide the conversation. So take someone by the hand and ask questions." (Patient/40-50y)*

Many patients and healthcare professionals considered a personal approach as essential, for example: healthcare professionals should verify whether they have drawn correct conclusions from PROM-results and patients' explanations, and whether they understand patients' needs.

2.3. Trusting patient-professional relationship

Patients and healthcare professionals described that the regular contact in dialysis care is a strong facilitator to feel connected and build a trusting relationship, which

was considered a prerequisite to conduct adequate PROM-result conversations:

> *"Depending on trust, you tell someone more or less." (Patient/70-80y)*

Some patients mentioned that the close relationship can also give rise to expectations regarding healthcare professionals' engagement, which goes beyond nephrology care. Some patients and healthcare professionals mentioned that feeling connected can be a reason why patients sometimes prefer to talk about PROM-results with their nephrologist or nurse rather than other healthcare professionals (e.g. social worker or psychologist), while also acknowledging differences in expertise.

2.4. Patient-professional collaboration

Patients and healthcare professionals indicated that discussing PROM-results provides most benefits when approached as a collaboration, in which patients and healthcare professionals contribute their own knowledge, experiences and ideas, and together consider what suits the patient best:

> *"We have an agreement that I will already think about the solution myself. Then I come to her and explain: 'This is the complaint and I think this is the solution for this reason'. Then she says what she thinks the solution is. Then we discuss it and ultimately arrive at the best solution together." (Patient/30-40y)*

Many patients and healthcare professionals believed that this person-centred collaboration is now more adopted compared to the past. Some healthcare professionals explained that discussing PROM-results takes some practice to become familiar with and integrate into consultations:

> *"In the beginning, it was a bit uncomfortable, because you get a lot of information and you don't know how to process and discuss it with the patient." (Nephrologist/30-40y)*

Patients and healthcare professionals indicated that clear, open and honest communication from both sides is very important, and considered it important to acknowledge and verify non-verbal communication. Many healthcare professionals and some patients mentioned that interpersonal differences, for example in cultural or religious beliefs, can make discussing PROM-results more challenging. It

then takes more effort to understand each other and to articulate needs and suggestions:

> *"When the patient's view is very different and it goes against my gut feeling. That are the most difficult conversations." (Nephrologist/40-50y)*

Some patients and healthcare professionals shared that consequently, certain topics may remain undiscussed, for example sexual dysfunction in the presence of a family member.

THEME 3: FOLLOW-UP ON PROM-RESULTS DISCUSSIONS

Patients and healthcare professionals stated that discussing PROM-results, including follow-up actions and evaluation, should be part of routine care and integrated into the multidisciplinary treatment process.

3.1. Follow-up actions and evaluations embedded into treatment process

Patients and healthcare professionals indicated that it does not stop after discussing PROM-results: follow-up actions and continuous evaluation over time are part of this process:

> *"That she [nephrologist] asks at the next consultation 'Has it improved sir?'... So that you do not get a feeling like 'the conversation is done, the results end up in a drawer and the person [patient's perspective] does not matter'." (Patient/70-80y)*

Several healthcare professionals stated there is a need for guidance on how to act on PROM-results: they now tend to focus on symptoms for which they know treatment exist and less on medically 'non-treatable' symptoms or HRQOL, because they like to have "something to offer". Many patients indicated that they desire explanations and clarity about symptoms and their treatments, even when there are only few or no treatments available. Most patients and healthcare professionals described the influence of PROMs as 'fine-tuning' of treatment, since information and conversation about PROM-results help to personalise treatment:

> *"With PROMs, you can support patients where they experience problems. We used to look at laboratory values and blood pressure, but sometimes patients do not benefit from that. But if we work on the limited energy, we also build on a better quality of life." (Nurse/50-60y)*

3.2. Multidisciplinary process

Patients and healthcare professionals described follow-up actions on PROM-results as a multidisciplinary process, guided by patients' attending nephrologist and/or nurse practitioner – also called "the playmaker" in this context. This healthcare professional consults with colleagues within (e.g. during multidisciplinary meetings) and outside the care team, and refers patients for appropriate care:

> *"There may be things that require other healthcare professionals and then you refer [...] to the general practitioner, other medical specialists and paramedics." (Nephrologist/50-60y)*

Patients and healthcare professionals considered referral an appropriate response to certain PROM-results: patients should be able to discuss all topics with their attending nephrologist or nurse practitioner, but this healthcare professional is not expected to act upon all PROM-results.

Part B: Specific situations

Patients' and healthcare professionals' perspectives on specific situations are described below, while highlighting themes (by means of the letter 'T') from part A that were most strongly reflected (see Table 3 for illustrative quotations). For all situations, similar themes emerged as in part A. However, for sensitive topics, the roles of and interaction between patients and healthcare professionals slightly differed: compared to other topics, healthcare professionals were expected and advised to take more initiative and a leading role in conversations.

SITUATION 1: NO CHANGES OR NEW TOPICS

Patients and healthcare professionals emphasized that when PROM-results have not changed or revealed new topics, it should still be discussed (T1.5), to verify interpretation of PROM-results (e.g. because the patient's situation or opinion may have changed [T1.4]) and whether the patient wants to further discuss it (T2). Many mentioned that, due to the frequent contact between patients and healthcare professionals (T2.3), topics may have been discussed already. They stated that, in most cases, it is then sufficient to only briefly discuss PROM-results, simply to confirm the situation is stable, that the healthcare professional is aware of the PROM-results and that everything can be discussed (T1.5, T2).

SITUATION 2: SENSITIVE TOPICS

Patients and healthcare professionals indicated that PROMs can serve as a tool,

'a conversation starter', to discuss sensitive topics such as sexual problems, depression and death, that probably remained undiscussed otherwise (T1.4-1.5). In contrast to less sensitive topics (compared to general topics from part A), both patients and healthcare professionals experienced and advised healthcare professionals to up the leading role and take the initiative (T2.2) to talk about sensitive topics, "because patients have already taken the first step by reporting their complaints". Many stated that the conversation should be initiated by healthcare professionals as if it is any other topic, suggesting that "it is normal to talk about these topics" (T1.5), but with additional care, for example: use open questions, so that patients can decide what and how much they want to share (T2.1). Healthcare professionals should acknowledge and verify non-verbal communication, and ask whether patients want to talk about it at all (T2), at that moment or later (to ensure a safe environment [T1.1] and feeling prepared [T1.2]), with this or another healthcare professional (feeling connected and trust are highly important [T2.3]), and with or without others present (e.g. partner or interpreter) (T1.1, T2.4). Furthermore, all patients and healthcare professionals indicated that topics such as sexual problems, depression and death can be difficult to discuss for the patient. Although all patients recognized that such topics can be challenging for healthcare professionals as well, only few healthcare professionals confirmed this – most indicated that "no topic should be difficult for them".

SITUATION 3: NO MEDICAL TREATMENT

Patients and healthcare professionals indicated that also outcomes for which there is no medical treatment available should be discussed, since patients experienced and reported these complaints. They described that healthcare professionals can also support patients on these topics by listening to them (T1.5) and referring them to other healthcare professionals for non-medical support (e.g. regarding coping or lifestyle). Yet many healthcare professionals also described their discomfort when discussing symptoms for which they feel there is no solution. Most patients, however, indicated that they do not expect a solution for all symptoms, but do want explanation and clarity on potential causes, prognoses and options that may provide relief (e.g. things they can do themselves and/or with support from other healthcare professionals [medical or paramedical]) (T3).

SITUATION 4: FACTORS NOT RELATED TO KIDNEY DISEASE OR DIALYSIS TREATMENT

All patients and healthcare professionals emphasized that everything can be discussed, whether or not it is directly related to CKD or dialysis treatment (e.g. co-

morbidities or life events) (T1.5). The nephrologist or nurse is considered the initial person to discuss PROM-results, as they are the patient's first point of contact, who have a complete overview and understanding of the patient's situation, and can refer to the proper healthcare professionals for additional support (T1.4, T2-3). Although most healthcare professionals acknowledge that everything can be discussed with them (T1.5), they also strongly felt the need to define their responsibilities and manage patients' expectations towards them (T2-3). Many patients and healthcare professionals described that identifying potential causes and responsible (para)medical fields also helps patients to understand the symptoms and treatment options (T3).

Table 3. Illustrative quotations of patients and healthcare professionals for specific situations in discussing PROM-results (part B)

1. No changes or new topics

I actually have few or no complaints, so yes, then I think that it does not need to be discussed. Well, it can always be mentioned that "I have looked at it and I see that you have no complaints or that you are doing fine with that." *(Patient, age: 60-70y)*

Not much I think. If you have already discussed it extensively and nothing changes and the needs of the patient have not changed either.. Yes, then I take it for notice. *(Nephrologist, age: 50-60y)*

And you can of course, certainly for quality of life, raise the topic in general; what keeps the patient busy. And sometimes it happens that you notice something and think hey, we should talk about that a bit further. *(Nurse, age: 50-60y)*

So then I would suggest that the doctor decides together with the patient, do you feel the need or do we skip it this time? *(Patient, age: 40-50y)*

The answers may be the same, but the circumstances can vary. Yes, then you will have to have the same conversation, and maybe it does not take as long. Yes, that is basically what happens medically as well. *(Patient, age: 70-80y)*

2. Sensitive topics

I do think that these are things that you should always discuss, yes. Because the more we discuss these kinds of things, the more the taboo disappears, the easier it becomes to discuss these kinds of things. *(Patient, age: 30-40y)*

Yes, and maybe also to indicate that it's OK to talk about this. That you indicate as a healthcare professional that you find it normal and that more people have problems with this. So that it also feels safe to talk about this. […] Because yes, if it bothers them, it should be possible to discuss it. *(Nurse, age: 50-60y)*

Yes, you know, [sexual dysfunction] it's at the top of the list as the most distressing symptom in male dialysis patients, yes… I mean, the idea of that PROM is that you get insight into what bothers patients. Yeah, if this comes up then I think you can't ignore that. No matter how uncomfortable the subject may be for many. *(Nephrologist, age: 40-50y)*

There are quite a few points that you say well, that is a bit more sensitive, perhaps also to the nephrologist. But I do think you should be able to discuss that. […] The patient experiences that and brings it up. So he [healthcare professional] has to meet him and indicate "I would rather have that you discuss this with a colleague… I'm having some trouble with it." So… Well, I'd only respect that. *(Patient, age: 70-80y)*

Well, discreetly. The patient must of course feel safe, there must already be a trusting relationship, that is extremely important. I always think for myself, how would I experience that? If the doctor were a complete stranger, I would also think well never mind. And also mention that it may be a difficult topic, but that you will figure it out together. And the healthcare professional may also sometimes say that he finds it difficult, and show a certain vulnerability. That can of course also comfort the patient. *(Nurse, age: 50-60y)*

It was about death, and a social worker was with me and stammered and I said hey, I don't know what you want to say, but just tell me… Then you notice that not everyone is used to discussing certain topics. *(Patient, age: 40-50y)*

Well I think that the healthcare professional should first determine whether it is important to discuss; Do you think it is important to talk about it? Do you want to talk about it? And what do you want to say about it? *(Patient, age: 70-80y)*

I just openly ask if they want to discuss it and if they say no I respect that. So there is also a kind of shared decision-making in which topics you want to discuss together or not. We also discuss the possibility of discussing it with someone else and then it is up to the patient… a patient also has the right not to discuss things with me. *(Nephrologist, age: 50-60y)*

Discuss, discuss. Don't make your taboo the patient's taboo. You know, taboos arise when you make a fuss about it. So, just be open, honest… and also clearly give space. And there is no need to talk about it, that is not an end in itself. The PROMs aren't there because you need to talk about everything, but you do need to see if the patient wants to talk about it, and if he doesn't, then don't force it. Because you are there for the patient, and the patient is not there for you. *(Nephrologist, age: 40-50y)*

3. No medical treatment

Yes, listening, I think.. to the inconvenience. And more… more often ask how these things are going. Because I think talking about it will help that patient. *(Patient, age: 70-80y)*

Sexuality is also a topic that we doctors usually do not bring up. […] The problem is that I can't do much either. I can of course refer… if they wish. […] The point is, I can bring up a topic, but then I cannot offer a solution. Yes, I find that difficult. *(Nephrologist, age: 60-70y)*

I think you just have to be honest with this; at the moment we don't have a medical solution for you, but we might be able to look into things beyond that. Perhaps there are things that you can solve with fitness or food, for example. *(Patient, age: 30-40y)*

But in any case discuss why nothing could be done about it. I mean, are there no pills, is there no ointment or is there any other way, things like that. That it is explained why nothing can be done about it. I would consider that very important. *(Patient, age: 70-80y)*

If the patient indicates that they are bothered by this, then they should be able to discuss it. That can also be a relief, that people are listening. I think it should be discussed and then you can also confirm, even if we can't do anything with it. I think that also does something for the patient. *(Nurse, age: 50-60y)*

And, what I also mention is that we cannot change some complaints in that it is present, but we can change the way you look at it and how it hinders you in your daily life. Together we can look at how we can best approach this. And sometimes social work can provide support, or medical psychology, or dietetics. So I try to help them in that way. *(Nephrologist, age: 40-50y)*

The patient often does not expect me to solve the problem. That is also the reason why he often does not discuss it at all, because he really feels that I am not going to change that. But, often just that short explanation that you explain I have seen that you suffer from it, that may be difficult for you, and this is the reason why it occurs. That often is enough. *(Nephrologist, age: 40-50y)*

You can't be good at everything. No… I mean, it's nice if there's a solution, but I can imagine that she can't solve everything. *(Patient, age: 60-70y)*

4. Factors not related to kidney disease or dialysis treatment.

But certainly if it is in that area, that they are struggling, are sad, worrying, yes then I realize… I'm still the nephrologist, so it's important that I let them know that I've heard it and that I'm taking it seriously and want to do something about it, but I'm not going to solve it. So yes, I will talk to the patient; do you want to do something with this, and what do you want me to do with it? And that could indeed be a social worker, general practitioner or… someone in the family you talk to. *(Nephrologist, age: 40-50y)*

Yes say a good talk with the nephrologist about the social aspects, yes... But she might say yes, but that's not my area of expertise... […] And there can of course be things of which she says I know too little about this, I would…, that is what I mean by guidance, here I would say well, you could contact the social worker. *(Patient, age: 70-80y)*

And, that would of course be very short-sighted if I were to say I have my own subject and I stick to that. I can't imagine a nephrologist looking at his profession that way. *(Nephrologist, age: 50-60y)*

It is of course influenced by many more things than just your illness... the environment also determines how you feel, how you react to other people and how people react to you. So your environment also matters. *(Patient, age: 70-80y)*

And it is also important that the doctor finds out whether there may be other causes for the complaints, in addition to kidney disease in this case. And then look again where the appropriate support can be found for those complaints. *(Patient, age: 60-70y)*

Often you cannot make that distinction. I mean, many complaints are multifactorial, and by mentioning that, they also understand that some things won't go away completely. *(Nephrologist, age: 40-50y)*

Look, sometimes you don't know why you feel something… Is it because of the cancer that I feel insecure, or is it because of my kidneys? I do not always know that. And I would also not know where to go with these questions. *(Patient, age: 70-80y)*

Discussion

This qualitative study investigated patients' and healthcare professionals' experiences with and perspectives on optimal ways to discuss PROM-results in routine dialysis care. Three overarching themes were identified describing the organization and basic principles of discussing PROM-results, the roles of and interaction between patients and healthcare professionals, and follow-up after PROM-results discussions.

The majority of existing literature is theoretical and focuses on potential benefits and mechanisms explaining *why* PROMs contribute to patient-professional communication.[14, 16] Our study adds to this literature by providing in-depth knowledge on *how* to discuss PROM-results based on experiences and insights from routine nephrology care, resulting in practical guidance that is directly applicable to real-world practice. Our results showed similarities to the findings in the literature reviews of Greenhalgh et al. and Schick-Makaroff et al., for example: PROMs help to start conversations through patients' self-reflection, the confirmation that everything can be discussed or "permission to raise issues", by providing a tool for discussion and articulation of experiences (e.g. sensitive topics such as sexual problems and depressive symptoms) and needs, and by increasing healthcare professionals' insight into and awareness of patient-relevant outcomes.[14, 16] Moreover, in line with and building on their findings[14, 16], we found that a good patient-professional relationship is crucial for discussing personal concerns, needs and PROM-results, and that the frequent contact in dialysis care was considered a strong facilitator in building this trusting relationship. In The Netherlands, it is common practice that patients have one primarily responsible nephrologist and nurse, and therefore most patients already have established patient-professional relationships when discussing their PROM-results. Our results suggest that this is a good foundation for getting the most out of discussing PROM-results in routine dialysis care and potentially also in other chronic care settings.

Notable in our research is the difference between what is pursued and experienced, when it comes to patient-professional interaction. Healthcare professionals in particular expressed they would like patients to have more control, while at the same time acknowledging they struggle with not being too dominant and directing in conversations. Although healthcare professionals expressed preference for a guiding role, no suggestions were made on what or how they could change their own approach in practice. The preferred patient-professional interaction aligns with the deliberative model as described by Emanuel & Emanuel (1992).[38] Although our findings and literature suggest that patients' autonomy and values are increasingly given place in practice, further guidance may be needed to create a partnership as in the deliberative model.[39] Currently, healthcare professionals may not be aware of or actively choose a communication style, and literature also shows that it may be challenging for health-

care professionals to change their usual approach despite new insights and awareness.[14] We believe that recognition is the first step and change may require further guidance, training and practice.[40] Though, our research already provides some suggestions to enhance patients' initiative, for example: healthcare professionals could inform patients beforehand about the aim and relevance of PROMs, and about when and how they can share their experiences, concerns and questions about PROM-results; this may encourage patients' preparation, awareness that everything can be discussed and empower them to raise topics during consultations.

Another remarkable finding is that healthcare professionals tend to prioritize topics for PROM-discussions based on perceived 'treatability of symptoms', while at the same time emphasizing that PROM-results and *all* topics should always be discussed, irrespective of the situation (e.g. whether or not there is a medical treatment available or it is CKD-related), and that discussing PROM-results in itself can already help. This underlines the importance of asking, articulating and verifying what is important to patients and what their needs are, and the central role (i.e. the "playmaker" role) assigned to the attending nephrologists and/or nurses in dialysis care (i.e. discussing all topics and, if needed, referring to other [para]medical healthcare professionals when necessary); this role in nephrology may be broader than in other medical fields, for example: a study in oncological care found that discussions and further exploration based on PROM-results were limited to cancer-related issues.[41]

Most of the barriers identified in this study are in line with and add to literature on PROM-implementation into routine care.[21-25] However, our study shows a different nuance regarding time constraints, namely: using PROMs itself is not time consuming (some stated that it even saves time), but the fact that it reveals new topics that would otherwise remain unnoticed and undiscussed. This nuance does not discard that time management can be challenging, but it primarily confirms a major advantage of using and discussing PROMs namely that we gain a more complete, person-centred overview of how patients are really doing. Moreover, we expect time challenges to diminish over time: it will most likely take less time once patients and healthcare professionals have become familiar with discussing PROM-results[12], and topics will be discussed gradually over the CKD-trajectory, resulting in less topics to discuss compared to the first PROM-discussion.[42]

Our findings confirm that using PROMs can facilitate person-centred care[14-17, 43], and form the foundation for training healthcare professionals in discussing PROM-results. See also Box 1 for a brief overview of the practical guidelines and considerations. These practical guidelines can be used to develop novel or adapt existing training-tools (e.g. the PROmunication tool developed in cancer care by Skovlund et al.[15]) for the nephrology care setting. Training healthcare professionals in how to use and

discuss PROM-results can improve shared decision-making, since both the information obtained with PROMs and a good patient-professional communication, enhance this process.[44]

Further research is needed to investigate whether and how discussing PROM-results can improve actual patients' outcomes (e.g. HRQOL).[45] Moreover, healthcare professionals expressed the desire for guidance on actions that can be taken in response to PROM-results. Literature also suggests that healthcare professionals are not always aware of all available treatment-options to improve symptom burden[11, 24], hereby underlining the need for further research on and development of a guide to improve symptom-management and HRQOL.[46]

Our study elucidates in-depth knowledge of patients' and healthcare professionals' experiences with and perspectives on optimal ways to discuss PROM-results in routine dialysis care. Participants had experience with completing and discussing PROM-results, resulting in knowledge that is directly applicable to real-world practice. We achieved saturation and demonstrated robustness of our results by exploring specific situations in addition to the general dialysis context. Limitations include that, despite efforts to achieve heterogeneous samples, our results are shaped by people who are willing to participate in interviews, which may limit the transferability of the findings to the total population. For example: participants may have a more positive attitude towards research, healthcare and PROMs. Patients and healthcare professionals who are struggling with or who experienced little benefit from using PROMs, may not have felt addressed, motivated or comfortable enough to participate in the interviews. Though, our participants also shared critical notes, explained that discussing PROM-results took practice, and provided examples of patients and healthcare professionals who believe or behave otherwise. Moreover, all participants were Dutch and hence, topics such as cultural differences and language barriers were primarily uncovered via second-hand information (e.g. healthcare professionals' perspectives). Finally, we should be aware that our own experiences and preconceptions may have coloured our results (i.e. research reflexivity)[32], even though multiple researchers with different backgrounds have performed interviews and interpreted data.

In conclusion, this study provides in-depth knowledge into patients' and healthcare professionals' experiences with and perspectives on the organization of, the roles of and interactions between patients and healthcare professionals in, and follow-up on PROM-results discussions. Our findings form the foundation for training of healthcare professionals regarding optimal ways to discuss PROM-results in routine dialysis care. Further research is needed to provide guidance on follow-up actions in response to PROM-results to ultimately improve patients' outcomes.

Box 1. Brief overview of practical guidelines and considerations for optimal discussion about PROM-results between patients and healthcare professionals.

1. **Organization and basic principles of discussing PROM-results**

1.1 Suitable setting for discussing PROM-results

- Invite patients to complete PROMs 2-6 weeks in advance.
- Inform patients that, how and why the PROM-results will be discussed.
- Discuss PROM-results face-to-face, in a safe and private setting, e.g. during annual consultations (once per year is sufficient).

1.2 Preparations by patients and healthcare professionals

- Both prepare the discussion about PROM-results: think beforehand about what to discuss, if and what support is needed, and which options exist.

1.3 Dealing with time constraints

- PROMs provide additional information on how the patient is really doing. As a patient and healthcare professional: together prioritize and discuss the 3-4 most important topics.

1.4 PROMs as tool for personalised holistic consultations

- Use PROMs as a tool that provides guidance on what is important to the patient.
- Use PROM-results as a 'conversation-starter' to discuss topics that otherwise remain undiscussed (e.g. sensitive topics such as sexuality).
- Use PROMs to gain insight and overview through patients' self-reflection, comparative information and a more complete picture.

1.5 Always discuss PROM-results

- Discussing PROM-results is an essential part of using PROMs in routine care.
- Discuss PROM-results to make clear that any topic can be discussed and to create an opportunity to start a conversation about these topics.
- Discuss PROM-results as discussing itself also helps patients to inform or to feel seen, heard and understood by their healthcare professional.

2. **Roles of and interaction between patients and healthcare professionals**

2.1 Patient's role: active participant

- As a patient: take the initiative and articulate your experiences and needs.
- As a healthcare professional: realize that patients often wait for their

healthcare professional's initiative and, if so, explicitly provide them the space to share experiences and needs.

2.2 Healthcare professional's role: guide

- As a healthcare professional: take up a guiding role and provide structure in the conversation; ask patients to expand on their experiences and needs, help to prioritize topics and actions, and verify your interpretation and conclusions.

2.3 Trusting patient-professional relationship

- The nephrologist and/or nurse who knows the patient well should conduct the conversation; feeling connected and a trusting relationship is a prerequisite to discuss PROM-results and to share personal concerns, experiences and needs.

2.4 Patient-professional collaboration

- Both contribute your own knowledge, experiences and ideas, and together consider what suits the patient best.
- Interpersonal differences can be a barrier; take the time and effort to understand each other and articulate needs and suggestions.

3. **Follow-up on PROM-results discussions**

3.1 Follow-up actions and evaluations embedded into treatment process

- Follow-up actions and evaluations are part of the process of discussing PROM-results.
- As a healthcare professional: provide patients with explanation and clarity about their symptoms and treatment options; patients do not expect a solution for all symptoms.
- Take PROM follow-up actions for fine-tuning of treatment; information and conversation about PROM-results help to personalise treatment.

3.2 Multidisciplinary process

- Incorporate PROM follow-up actions into the multidisciplinary process, guided by the nephrologist or nurse practitioner.
- Discuss all topics, regardless of the situation (e.g. when no medical treatment is available or experiences are not directly related to CKD) and refer to other (para-)medical healthcare professionals when needed and desired by patients; referral is an appropriate response to PROM-results.

References

1. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

2. van der Willik EM, Hemmelder MH, Bart HAJ, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. *Clin Kidney J* 2021;14(6):1535-44

3. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

4. Manns B, Hemmelgarn B, Lillie E, et al. Setting research priorities for patients on or nearing dialysis. *Clin J Am Soc Nephrol* 2014;9(10):1813-21

5. Urquhart-Secord R, Craig JC, Hemmelgarn B, et al. Patient and Caregiver Priorities for Outcomes in Hemodialysis: An International Nominal Group Technique Study. *Am J Kidney Dis* 2016;68(3):444-54

6. de Jong Y, van der Willik EM, Milders J, et al. Person centred care provision and care planning in chronic kidney disease: which outcomes matter? A systematic review and thematic synthesis of qualitative studies : Care planning in CKD: which outcomes matter? *BMC Nephrol* 2021;22(1):309

7. Weisbord SD, Fried LF, Mor MK, et al. Renal provider recognition of symptoms in patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2007;2(5):960-7

8. Pugh-Clarke K, Read SC, Sim J. Symptom experience in non-dialysis-dependent chronic kidney disease: A qualitative descriptive study. *J Ren Care* 2017;43(4):197-208

9. Aresi G, Rayner HC, Hassan L, et al. Reasons for Underreporting of Uremic Pruritus in People With Chronic Kidney Disease: A Qualitative Study. *J Pain Symptom Manage* 2019;58(4):578-86.e2

10. Flythe JE, Dorough A, Narendra JH, Forfang D, Hartwell L, Abdel-Rahman E. Perspectives on symptom experiences and symptom reporting among individuals on hemodialysis. *Nephrology Dialysis Transplantation* 2018;33(10):1842-52

11. Feldman R, Berman N, Reid MC, et al. Improving symptom management in hemodialysis patients: identifying barriers and future directions. *J Palliat Med* 2013;16(12):1528-33

12. van der Willik EM, Terwee CB, Bos WJW, et al. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. *Nephrology (Carlton)* 2021;26(5):391-9

13. Etkind SN, Daveson BA, Kwok W, et al. Capture, transfer, and feedback of patient-centered outcomes data in palliative care populations: does it make a difference? A systematic review. *J Pain Symptom Manage* 2015;49(3):611-24

14. 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. *J Patient Rep Outcomes* 2018;2:42

15. Skovlund PC, Ravn S, Seibaek L, Thaysen HV, Lomborg K, Nielsen BK. The development of PROmunication: a training-tool for clinicians using patient-reported outcomes to promote patient-centred communication in clinical cancer settings. *J Patient Rep Outcomes* 2020;4(1):10

16. Schick-Makaroff K, Levay A, Thompson S, et al. An Evidence-Based Theory About PRO Use in Kidney Care: A Realist Synthesis. *Patient* 2022;15(1):21-38

17. Noonan VK, Lyddiatt A, Ware P, et al. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 3: patient-reported outcomes can facilitate shared decision-making and guide self-management. *J Clin Epidemiol* 2017;89:125-35

18. Carfora L, Foley CM, Hagi-Diakou P, et al. Patients' experiences and perspectives of patient-reported outcome measures in clinical care: A systematic review and qualitative meta-synthesis. *PLoS One* 2022;17(4):e0267030

19. van der Willik EM, Meuleman Y, Prantl K, et al. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease - a four-phase mixed methods study. *BMC Nephrol* 2019;20(1):344

20. Greenhalgh J, Dalkin S, Gooding K, et al. Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. Health Services and Delivery Research. NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK., Southampton (UK): 2017.

21. Foster A, Croot L, Brazier J, Harris J, O'Cathain A. 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

22. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *Bmj* 2015;350:g7818

23. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23(6):508-18

24. Evans JM, Glazer A, Lum R, et al. Implementing a Patient-Reported Outcome Measure for Hemodialysis Patients in Routine Clinical Care: Perspectives of Patients and Providers on ESAS-r:Renal. *Clin J Am Soc Nephrol* 2020;15(9):1299-309

25. Anderson NE, McMullan C, Calvert M, et al. Using patient-reported outcome measures during the management of patients with end-stage kidney disease requiring treatment with haemodialysis (PROM-HD): a qualitative study. *BMJ Open* 2021;11(8):e052629

26. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? *Soc Sci Med* 2005;60(4):833-43

27. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a non-categorical qualitative alternative for developing nursing knowledge. *Res Nurs Health* 1997;20(2):169-77

28. Thompson Burdine J, Thorne S, Sandhu G. Interpretive description: A flexible qualitative methodology for medical education research. *Med Educ* 2021;55(3):336-43

29. Thompson S, Schick-Makaroff K. Qualitative Research in Clinical Epidemiology. In: Parfrey PS, Barrett BJ, editors. Clinical Epidemiology: Practice and Methods. Springer US, New York, NY: 2021; 369-88.

30. Morse JM. The significance of saturation. *Qual Health Res* 1995;5(2):147-9

31. 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* 2007;19(6):349-57

32. Amir N, McCarthy HJ, Tong A. Qualitative Research in Nephrology: An Introduction to Methods and Critical Appraisal. *Kidney360* 2021;2(4):737-41

33. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants. *International Journal of Qualitative Methods* 2019;18:1609406919874596

34. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34(3):220-33

35. Weisbord SD, Fried LF, Arnold RM, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: the Dialysis Symptom Index. *J Pain Symptom Manage* 2004;27(3):226-40

36. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;3(2):77-101

37. Green J, Thorogood N. Thematic content analysis. Qualitative methods for health research. 4th ed. SAGE: 2018; 258-68.

38. Emanuel EJ, Emanuel LL. 4 Models of the physician-patient relationship. *Jama-Journal of the American Medical Association* 1992;267(16):2221-6

39. Borza LR, Gavrilovici C, Stockman R. Ethical models of physician - patient relationship revisited with regard to patient autonomy, values and patient education. *Medical-Surgical Journal-Revista Medico-Chirurgicala* 2015;119(2):496-501

40. Damman OC, Verbiest MEA, Vonk SI, et al. Using PROMs during routine medical consultations: The perspectives of people with Parkinson's disease and their health professionals. *Health Expect* 2019;22(5):939-51

41. Greenhalgh J, Abhyankar P, McCluskey S, Takeuchi E, Velikova G. How do doctors refer to patient-reported outcome measures (PROMS) in oncology consultations? *Qual Life Res* 2013;22(5):939-50

42. Takeuchi EE, Keding A, Awad N, et al. Impact of patient-reported outcomes in oncology: a longitudinal analysis of patient-physician communication. *J Clin Oncol* 2011;29(21):2910-7

43. Feldthusen C, Forsgren E, Wallström S, et al. Centredness in health care: A systematic overview of reviews. *Health Expectations* 2022;25(3):885-901

44. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med* 2012;27(10):1361-7

45. Johnson JA, Al Sayah F, Buzinski R, et al. A cluster randomized controlled trial for the Evaluation of routinely Measured PATient reported outcomes in HemodialYsis care (EMPATHY): a study protocol. *BMC Health Serv Res* 2020;20(1):731

46. Meuleman Y. Integrating Patient-Reported Outcome (PRO) measures into Dutch dialysis care: Toward a PRO Treatment Guide to achieve optimal multidisciplinary and personalized dialysis (Dutch Kidney Foundation projectcode: 21OM+002). 2022.

Supplemental Material for Chapter 8

Table S1. Main topics with example questions* for individual semi-structured interviews with patients and healthcare professionals

Part A

Experiences with discussing PROM-results

- *How often have you discussed the PROM-results, and in what setting?*
- *How did you experience the discussion about the PROM-results?*
- *What went well when discussing the PROM-results, and what could be improved?*
- *What was the role of the healthcare professional/patient, and of yourself, during the conversation?*
- *What were your expectations for discussing the PROM-results, and what did it bring you?*

Perspectives on optimal ways to discuss PROM-results

- *What should be the purpose of discussing the PROM-results?*
- *How can it be ensured that what is important to the patient emerges during the conversation?*
- *How would you describe the ideal conversation about the PROM-results?*
- *What are barriers for having this ideal conversation, and what could be done to overcome these barriers?*
- *How can a conversation about PROM-results contribute to the treatment or wellbeing of the patient?*
- *What should be done after discussing the PROM-results?*

Part B

Perspectives on dealing with specific situations in discussing PROM-results

- *How should the PROM-results be discussed if:*
  - *no changes or new topics emerge?*
  - *it is about a sensitive topic that someone doesn't talk about easily?*
  - *there is no medical treatment for the symptoms that emerge?*
  - *factors not related to kidney disease or dialysis treatment (e.g. comorbidities or life events) may have caused the decreased quality of life or symptoms that emerge?*

* *Example interview questions are shown. During the interviews responses were further explored using additional questions and probes.*

9

Summary and general discussion

Summary and general discussion

This dissertation aimed to provide insight into and practical knowledge of the implementation and use of patient-reported outcome measures (PROMs) in routine nephrology care. We performed research within the different steps for implementation of PROMs into routine care, including the selection of PROMs, pilot testing PROMs, and nationwide implementation and use of PROMs, using a broad variety of quantitative and qualitative research methods. We investigated the use of PROMs both at individual patient level and at population level, with the potential to facilitate personalised treatment and evaluation of healthcare quality. In this chapter, we summarize our main findings, discuss the implications of our main findings, and provide suggestions for future research, and for further implementation of PROMs into routine care.

Summary of main findings

Implementation of PROMs into routine nephrology care

SELECTION OF PROMS

Based on existing literature and in collaboration with patient representatives and healthcare professionals, we identified generic health-related quality of life (HRQOL) and disease-specific symptom burden as important outcomes to measure at individual and population level in routine nephrology care **(Chapter 1)**. The next step was to select PROMs to assess these patient-reported outcomes (PROs). In **Chapter 2**, we described the selection of the best suitable existing PROM to assess disease-specific symptom burden for routine assessment in nephrology care. We conducted this study in four phases. In the first two phases, we searched and build on the existing literature, from which we identified 28 potentially suitable symptom questionnaires and 10 symptom clusters. During the third phase, the questionnaires were evaluated based on predefined criteria regarding the relevance (e.g., applicable to CKD population), completeness (e.g., 90% cluster coverage) and comprehensibility (e.g., appropriate length, and straightforward and clear questions). Two questionnaires met the criteria: the Dialysis Symptom Index (DSI) and Palliative Care Outcome Scale-Renal Version (IPOS-Renal). In the fourth phase, these questionnaires were reviewed by 2 panels of in total 151 patients who were randomly assigned to a questionnaire, and 1 panel of 6 experts (i.e., experienced questionnaire assessors) who compared both questionnaires. Patients reported more symptoms using the DSI compared to the IPOS-Renal (12 and 8 symptoms, respectively), and needed less time to complete the DSI (5.4 and 7.5 minutes to complete the DSI and IPOS-Renal, respectively). Both the patients and

experts panels assessed the DSI as the most complete, specific and comprehensible symptom questionnaire. Therefore, the DSI was selected as PROM to assess disease-specific symptom burden in routine nephrology care.
The 12-item Short-Form Health Survey (SF-12) is a validated and commonly used PROM to assess generic HRQOL, and was recommended by an European expert consensus group for use in routine nephrology care.[1] In addition, the SF-12 was – similarly to the DSI – selected by our patients and experts panels as suitable PROM to assess generic HRQOL in routine nephrology care. Later in time, the Patient-Reported Outcomes Measurement Information System (PROMIS) was selected as one of the recommended PROMs to measure generic HRQOL in patients with CKD by a consensus group of the International Consortium of Health Outcomes Measurement (ICHOM).[2] PROMIS instruments can also be administered as computerized adaptive tests (CATs), that are expected to deliver similar or even more precise measurements with fewer questions compared with fixed (i.e., non-adaptive) PROMs.[3] To explore this relatively novel method in healthcare, we examined and compared the psychometric properties of seven PROMIS CATs compared with the SF-12 in patients with advanced chronic kidney disease (CKD) in **Chapter 3**. We performed a content comparison between the seven PROMIS CATs (assessing physical function, pain interference, fatigue, sleep disturbance, anxiety, depression, and the ability to participate in social roles and activities) and the SF-12, and examined the construct validity and test-retest reliability. We found evidence for sufficient construct validity of all seven PROMIS CATs. Furthermore, the PROMIS CATs, the SF-12 summary scores and most SF-12 domains, and the DSI showed sufficient test-retest reliability. Overall, PROMIS CATs showed better reliability, resulting in a lower minimal detectable change (MDC), compared with the SF-12. However, seven PROMIS CATs required 45 items (10 minutes), which is 3 to 4 times the length of the SF-12 (12 items; 3 minutes). These results show evidence for sufficient construct validity and a better test-retest reliability of seven PROMIS CATs, but requiring more items, compared with the SF-12. Moreover, these results do not address the suitability and feasibility of PROMIS CATs in routine nephrology care and therefore, the SF-12 is retained for now.

PILOT TESTING AND NATIONWIDE IMPLEMENTATION OF PROMS

Chapter 4 described the experiences and results of the first introduction of PROMs into Dutch routine nephrology care, in the form of a pilot study in 16 dialysis centres across the Netherlands. We used quantitative and qualitative research methods to explore these first experiences. In total, 512 patients receiving dialysis treatment completed 908 PROMs across three time points. The quantitative

part showed that there is room for improvement in patients' PROM-scores: patients receiving dialysis treatment experienced a substantially decreased physical HRQOL and a high symptom burden, with on average 11 different symptoms of moderate burden. The variation between the symptom frequency and burden suggests that the most common symptoms are not necessarily the most burdensome for patients. Furthermore, this first introduction of PROMs in routine dialysis care showed a low average response rate of 36%, which varied from 6% to 70% among centres. The high variability across centres underlines that achieving high response rates is feasible, but challenging and may require extra encouragement of patients and healthcare professionals. In the qualitative part, we explored patients' and healthcare professionals' experiences and preferences regarding the use of PROMs in clinical practice. Patients appreciated the content, length and structure of the PROMs (DSI and SF-12) and the online completion of PROMs. Individual feedback should be presented in a relevant context (e.g., with reference scores of similar patients) and can contribute to and in preparation for a consultation. Furthermore, some patients already discussed individual PROM-results with their healthcare professional. Patients and healthcare professionals indicated that discussing HRQOL and symptom burden scores was highly insightful and valuable, and individual feedback on PROM-scores was considered crucial. These first experiences with discussing PROM-results were promising.

Building on the findings from the pilot study, the PROMs infrastructure was further optimized for nationwide implementation and use of PROMs in routine dialysis care. For example, improvements were made to broaden the applicability (e.g., PROMs came available in four languages), to support implementation (e.g., a webpage about PROMs with information and hand-outs with tips and tricks to guide implementation), and to facilitate the use of PROMs (e.g., reports with individual PROM-results were provided to patients and their healthcare professionals directly after completing the PROMs). The PROMs became available to all dialysis centres in the Netherlands through Nefrovisie (as part of the renal registry Renine), and centres were invited to implement using PROMs into routine dialysis care.

Use of PROMs in routine nephrology care

USE OF PROMS AT POPULATION LEVEL

At aggregated population level, PROM-results can be used to evaluate healthcare quality and to inform patients and healthcare professionals about the effects and course of disease or treatment. Funnel plots can be used to evaluate healthcare quality by comparing hospital performances on certain outcomes. In **Chapter 5**, we explained the use and interpretation of funnel plots by presenting an overview

of the basic principles, pitfalls and considerations when applied to PROs, using examples from Dutch routine dialysis care. A funnel plot is a graphical method to evaluate healthcare quality and has several advantages, including clearly visualized precision, detection of volume-effects, discouragement of ranking hospitals and easy interpretation of results. However, without sufficient knowledge of underlying methods, it is easy to stumble into pitfalls, such as overinterpretation of standardized scores, incorrect direct comparisons of hospitals and assuming a hospital to be in-control (i.e., to perform as expected) based on underpowered comparisons. Furthermore, application of funnel plots to PROs is accompanied by additional challenges related to the multidimensional nature of PROs and difficulties with measuring PROs. To enable relevant and fair comparisons of PROs, high and consistent response rates, adequate case mix correction and high-quality PRO measures are required. These challenges need to be addressed before using PRO data for healthcare quality evaluations, for instance by using funnel plots. In **Chapter 6,** we showed an example of aggregated PROM-results that can be used to inform patients and healthcare professionals about the course and effects of disease and outcomes. In this chapter, we investigated the impact of itching on HRQOL and interactions with sleep problems and psychological symptoms in patients receiving dialysis treatment. We performed cross-sectional and longitudinal analyses in 2978 patients who completed the PROMs between 2018 and 2020. Our results showed that half of the patients experienced itching and in 70% of them, itching was persistent over time. Patients with itching experienced a 3 to 4 points lower physical and mental HRQOL compared with patients without itching, which remained stable during 2 years of follow-up. Furthermore, we found that sleep problems (70% versus 52%) and psychological symptoms (36% versus 19%) were more common in patients with itching. These symptoms had an additional negative effect on physical and mental HRQOL but did not interact with itching (i.e., the combination of both symptoms did not result in a significantly lower or higher HRQOL than the sum of individual effects). The high prevalence and persistence of itching, its impact on HRQOL over time and the additional effect on HRQOL of the often co-occurring sleep problems and psychological symptoms emphasize the need for recognition and effective treatment of itching to reduce symptom burden and improve HRQOL in patients receiving dialysis treatment.

USE OF PROMS IN INDIVIDUAL PATIENTS

For optimal use of PROMs in individual patients, knowledge on how to interpret and discuss PROM-results is needed. In **Chapter 7**, we explained the different types and characteristics of PROMs and provide guidance on how to interpret

individual PROM-scores and changes in PROM-scores over time. In this chapter, we introduced types and characteristics such as generic and specific PROMs, and scoring systems of PROMs. We explained that intuitive measures such as information about the average and distribution of PROM-scores in a reference population or in comparison to more familiar outcomes (e.g., laboratory measures) are indispensable to interpret and get used to PROM-scores. Furthermore, methodological concepts such as the MDC and minimal important change (MIC) are important to inform us about statistically and clinically relevant changes, respectively. Besides, one must be aware that response shift may occur, which refers to a change in the meaning of the patient's evaluation of the PRO over time (e.g., a change in one's perception on HRQOL). A response shift may explain unexpectedly small (or large) changes in PROM-scores. Finally, having a conversation with the patient is important to interpret individual PROM-scores. The best manner to interpret individual PROM-scores and changes in PROM-scores is through a discussion between the patient and the healthcare professional, in which measures like MDC, MIC and response shift may have a facilitating role. For example: the MIC provides an indication of which changes in PROs are likely considered relevant at group level and the discussion of individual results reveals what changes are important to this specific patient, to what extent, and in which manner.

In **Chapter 8**, we investigated how to optimally discuss PROM-results by conducting semi-structured interviews with 22 patients receiving dialysis treatment and healthcare professionals about their experiences with and perspectives on discussing PROM-results in routine dialysis care. Interviews focused on general situations and specific situations (e.g., addressing sensitive topics or when no medical treatment is available). Patients and healthcare professionals (nephrologists and nurses) highly appreciated the use of PROMs, as it provides insight and overview of how the patient is doing and feeling, and contributes to patient-professional communication. Furthermore, patients and healthcare professionals provided practical guidance for optimal discussion about PROM-results. First, patients and healthcare professionals emphasised that PROM-results should always be discussed and indicated how to create a suitable setting, adequately prepare, deal with time constraints and use PROMs as a tool for personalised holistic consultations. Second, patients should actively participate and healthcare professionals should take a guiding role. A trusting patient-professional relationship was considered a prerequisite and patient-professional interaction was described as a collaboration in which both contribute their knowledge, experiences and ideas. Third, follow-up after discussing PROM-results was considered important, including evaluations and actions (e.g., symptom management) structurally embedded

into the multidisciplinary treatment process. These general themes also applied to the specific situations, for example: results should also be discussed when no medical treatment is available. Interesting to note is that healthcare professionals were expected to take more initiative and a leading role when discussing sensitive topics. This study provided in-depth knowledge and practical guidance on how to organise and conduct conversations about PROM-results in routine nephrology care.

General discussion

The findings of this dissertation provide insights into and practical knowledge of the implementation and use of PROMs in routine nephrology care. Specific considerations and implications regarding each study have been discussed in the corresponding chapters (Chapter 2-8). In this part, we discuss the implications for clinical practice, and future directions for research and further implementation of PROMs, based on our overall findings and experiences with PROMs in routine nephrology care.

Implications for clinical practice

IMPLEMENTATION IS AN ITERATIVE PROCESS THAT TAKES COLLABORATION, TIME, AND EFFORT

The added value of PROMs is to a great extent determined by how well the PROMs are integrated into healthcare.[4, 5] Therefore, a structured and carefully prepared approach to implement PROMs into routine care is necessary.[6-8] Based on our experiences with implementing PROMs into routine nephrology care, we would like to highlight several important aspects to facilitate optimal implementation of PROMs into a routine care setting.

Firstly, collaborate with all stakeholders in all phases: from designing the project to implementing PROMs into routine care. Literature shows that patients are often not or only partly involved, even when new PROMs are being developed.[9, 10] We believe that here is great room for improvement. For example: in our project, a patient representative was one of the initiators and part of our research team, and was involved in all phases of the project. Furthermore, patients played an indispensable role in the implementation of PROMs into routine dialysis care, for instance by making sure that the PROMs, setting, timing, interpretation and feedback fits the patients' needs (Chapter 2, 4 and 8). Based on our experience, collaboration with patients and patient representatives supports both research and practice, and we hope that our inclusive approach will encourage more initiatives

to collaborate with patients and patient representatives. Furthermore, healthcare professionals play an important role in all phases to ensure that the use of PROMs fits the workflow (e.g., which timing, setting and healthcare professionals' roles are suitable) and that PROMs also provide added value for the healthcare professionals themself (e.g., when and how PROMs can provide insights and serve as a tool to conduct the conversation). Indeed, our results showed that engagement of healthcare professionals is an important facilitator both to implement PROMs (e.g., a coordinator on-site resulted in higher response rates) and to get the most benefits out of using PROMs (e.g., optimal discussion about individual PROM-results) (Chapter 4 and 8). A passionate professional may be a role model and important motivator for colleagues to optimally use PROMs.[11, 12] Besides, the healthcare quality institute of nephrology care (Nefrovisie Foundation) was an important facilitator and shows that implementation of PROMs at a national level is feasible, for example through use of the existing network and ICT infrastructure of the national registry.[8, 11, 13]

Secondly, carefully design and prepare the implementation of PROMs.[6] Designing includes defining the steps to be taken and which studies should be performed to inform the next steps, exploring the setting and purposes, and selecting the PROs and PROMs. Our research showed that the psychometric quality of the PROM but also factors related to the feasibility and suitability given the setting and purposes are important, for instance: the questionnaire length (i.e., number of items and time to complete), completeness and comprehensibility (Chapter 2). These factors should be examined for each PROM, setting, purpose and population that is considered. For example, PROMIS CATs showed good psychometric properties in patients with advanced CKD, but the feasibility and suitability in routine nephrology care remains to be explored before the next steps towards implementation of PROMIS CATs can be taken (Chapter 3).[14]

Preparation involves developing an electronic system to invite patients, collect PROMs and obtain individual feedback, and providing instructions and guidance to centres and healthcare professionals on why and how to use PROMs.[6, 15, 16] A great advantage of organising this nationally is that it is structured and similar across all centres, and not dependent on resources of individual centres. However, the manner in which PROMs are collected and used must fit within the workflow of the centres, and therefore, pilot testing is essential (Chapter 4).

Thirdly, treat the implementation as an iterative process of learning and improving, and invest time and effort. As with any other new approach in healthcare, it takes practice to adjust and become familiar with it. For PROMs, both the instrument itself and its use by patients and healthcare professionals are relatively new

in routine healthcare, and thus both the interpretation of PROM-scores and how to use it for personalised treatment takes practice (Chapter 7 and 8). Moreover, the iterative process involves adjusting expectations and assumptions. Barriers that are often reported in the literature[7, 15, 17] – such as 'PROMs take too much time' or 'PROMs create too high expectations in patients' – were nuanced in our studies by patients and healthcare professionals that had some experience with using PROMs, for example: healthcare professionals indicated that time was actually spent more efficiently by focussing on what is important to patients, and patients pointed out that they do not expect their healthcare professional to solve all complaints, only that it is discussed and that they are informed about potential causes, prognoses and treatment options (Chapter 8). Nevertheless, these nuances do not discard that it remains challenging to successfully implement PROMs into routine care. In this dissertation, the first steps taken are described, but the iterative process of optimal implementation of PROMs into routine nephrology care is still ongoing. Continuous evaluation of experiences with using PROMs in routine care, further research, training and guidance is needed to keep learning and improving.

IMPROVE RESPONSE RATES TO ENHANCE THE USE OF PROMS AT POPULATION LEVEL

The low average response rate of 36% with high variability across centres at first introduction of PROMs in routine dialysis care (Chapter 4) is an important finding of our pilot study, but also one of the main limitations when using PROMs at population level (Chapter 5 and 6). Although PROMs are now used in all Dutch dialysis centres and response rates are still increasing (45% in 2021)[18], the response rate remains an important point of attention. High and consistent response rates are needed to ensure that information at population level (i.e., in each centre and at national level) is of sufficient quality. Consistency in PROMs response is needed for multiple aspects, for example (unmeasured) characteristics of responders should not differ from non-responders; responders should be representative of the entire population of interest. In addition, the reason and timing of collecting PROs should be consistent. Patients should have an equal chance of being invited and being able to complete PROMs. Even though at individual level it may be reasonable to complete PROMs at indication (e.g., when someone has many symptoms), this should not be the main recruitment strategy, as this likely results in a biased or incomplete picture at population level (and also risk of under-recognition at individual level).[19, 20] Moreover, deciding on the right timing to collect PROs in nephrology care may be challenging since there is often no clear starting point in chronic care (e.g., prevalent dialysis patients). However, different timing in the

trajectory of the disease or treatment may demonstrate different PROM-results[21, 22], for example: it may matter whether someone has just started dialysis treatment or is already receiving dialysis treatment for a year (Chapter 7). The timing should thus be taken into account and preferably, comparable timepoints of PROMs completion over the entire disease and treatment trajectory are used for all patients.
Thus, recruitment strategies that yield high and consistent response rates are needed to enhance the use of PROMs at population level. Although the validity of the data strongly depends on the randomness of the (non-)response (i.e., representativeness of the responders), thresholds of 60-80% have been proposed in the literature as adequate response rates.[20, 23, 24] Patients and healthcare professionals considered discussing individual PROM-results essential in using PROMs in routine care (Chapter 4 and 8), and we believe this may be the most important facilitator in reaching high response rates. Furthermore, additional training of healthcare professionals (e.g., preparing healthcare professionals on how to invite and inform patients, and how to use PROMs) and support of patients (e.g., availability of tablets onsite or help with completing PROMs online) may improve response rates.[6, 25, 26] In addition, further development of the ICT infrastructure could improve response rates, for instance incorporation of PROMs into the electronic health record including automated invitations (e.g. 2-6 weeks prior to patients' upcoming annual consultation) and reminders to complete PROMs.[6, 27]

START WITH USING PROMS WHERE THEY DIRECTLY PROVIDE ADDED VALUE: AT INDIVIDUAL LEVEL

Our results show that already at first introduction of PROMs into routine dialysis care, the use of PROMs could facilitate the conversation about symptom burden and HRQOL between patients and healthcare professionals (Chapter 4 and 8). These findings confirm that using PROMs can improve patient-professional communication and support shared decision making by providing a tool to start and conduct a conversation and by providing a more complete picture and awareness of patient-relevant outcomes. These benefits of using PROMs contribute to a more person-centred healthcare and are directly achievable at individual patient level. Discussing individual PROM-results is already possible as soon as the patient has completed the PROMs, and is not dependent on population-level factors like high response rates or full integration into the electronic health record. However, these factors can further improve the use of PROMs at individual level; high quality information of similar patients – also known as 'patients-like-me' information – can facilitate the interpretation of individual PROM-results and shared decision making, and integration into the electronic health record provides insight into individual

results over time and comparison with clinical measures.[28, 29] Hence, our results imply that PROMs should directly be used at individual patient level and at the same time be further implemented at population level, so that the added value of completing and using PROMs is directly experienced by patients and healthcare professionals and will further improve over time.

Despite the fact that there are still steps to be taken, there are already some examples in the literature showing that PROs can be of added value in healthcare quality evaluation.[30-32] Future research should demonstrate whether this also applies to nephrology care. Our findings show that patients and healthcare professionals particularly consider the individual use of PROs of great added value, and individual use may therefore be the primary purpose of collecting PROs in routine nephrology care. However, we should keep in mind that individual and aggregated use often go together and may strengthen each other, for example, aggregated PROM-results can inform patients and healthcare professionals about prognosis, treatment and factors influencing PROs.[33, 34] Use at individual level is expected to improve response rates, which in turn results in better quality of aggregated information. Finally, the ultimate aim of collecting PROs is to improve patient-relevant outcomes and healthcare quality, and in order to evaluate whether the use of PROMs at individual level indeed results in these improvements, data on an aggregated level is required.

IMPROVE PATIENT-RELEVANT OUTCOMES

Our findings confirm the high symptom burden and decreased HRQOL that patients receiving dialysis treatment experience.[35-38] The high average number of 11 symptoms, the broad range of physical and psychosocial symptoms experienced by patients, and the fact that the most common symptoms are not necessarily the most burdensome (Chapter 4), may partly explain why symptoms were missed in routine nephrology care when not systematically assessed and discussed.[37, 39] Furthermore, by further investigating a common symptom in dialysis patients (i.e., itching) we found that this symptom is persistent in many patients, often co-occurs with other burdensome symptoms (i.e., sleep problems and psychological symptoms) and has a high impact on HRQOL (Chapter 6). These findings highlight the need for recognition, discussion and effective treatment to reduce symptom burden and improve HRQOL. Literature suggests that the use of PROMs may contribute to better symptom management, and that this in turn might result in improved patient-relevant outcomes such as a better HRQOL, less hospitalizations and lower mortality.[40-45] By implementing the PROMs into routine dialysis care, the first steps have been taken to provide insight into and facilitate discussion about

patients' symptoms and needs. However, to actually improve patient-relevant outcomes, follow-up actions in response to PROM-results are needed. The need for guidance to adequately respond to and act upon patients' symptoms and needs was also emphasized by healthcare professionals who already use and discuss individual PROM-results (Chapter 8).

In addition, using PROMs during consultations can help to focus on what is important to patients (Chapter 8). This may be of added value in decision making about starting, stopping or fine-tuning a treatment, for example: to what extent should anaemia be treated when the patient does not experience fatigue or other burdensome anaemia-related symptoms?[46] Until when is dialysis treatment beneficial, given the impact on someone's daily life?[47] Or to what extent is it beneficial to increase the prescription of phosphate binders when the patient experiences a high pill burden?[48] The patient's perspective is important in answering such questions and should be taken into account, just like clinical and biomedical values routinely are. The use of PROMs can already contribute to this more person-centred approach, and a personalised and holistic treatment. This may require a different approach from healthcare professionals and takes time and learning in practice to become familiar with.

Future directions for research

This dissertation provides grounds for further research into improving PROs and how to optimally use PROMs. Based on our findings, we would like to highlight some suggestions for further research.

FEASIBILITY AND SUITABILITY OF PROMIS CATS IN ROUTINE NEPHROLOGY CARE

Recently, PROMIS instruments have been recommended as generic PROMs for all medical specialist care in the Netherlands.[49] Our research showed sufficient validity and good test-retest reliability of PROMIS CATs (Chapter 3). However, in contrast to the SF-12 (and DSI), we did not investigate the suitability of PROMIS CATs within the setting and purpose of using PROMs in routine nephrology care. Therefore, further research is needed to explore the feasibility and suitability in routine nephrology care. In addition to characteristics such as questionnaire length (i.e., number of items and time to complete), comprehensibility and completeness, some additional factors specific for PROMIS CATs should be further examined: first, as the PROM is adapted to the patient's ability, questions will vary across patients and over time. Our research showed that particularly the specific items and not the overall scores are being used in clinical practice when discussing individ-

ual PROM-results (Chapter 8). Therefore, research is needed to explore whether and how PROMIS CATs can be used and provide (similar) added value during consultations in routine nephrology care. Second, PROMIS CATs can only be completed digitally. Currently, some dialysis centres are (also) using paper-based versions to reach the entire population and enable completion by the patient himself (without help). Further research could provide insight into whether it is feasible and desirable to shift towards only digital completion of PROMs (e.g., regarding accessibility to all patients and response rates, and digital resources and availability of help). Research findings regarding the feasibility and suitability of PROMIS CATs in routine nephrology care may help in determining the next steps in the iterative process of implementation and continuing to improve the use of PROMs.

RESEARCH TO ENABLE HEALTHCARE QUALITY EVALUATIONS BASED ON PROS

Using PROMs to evaluate healthcare quality requires not only further implementation at population level (e.g. higher and more consistent response rates), but also further research on the association between healthcare quality and PROs, and relevant case mix factors (Chapter 5). A PRO that is important to patients is not necessarily a suitable PRO for healthcare quality evaluation.[50, 51] To evaluate healthcare quality, an association between the PRO and healthcare quality must be plausible or established. To make relevant comparisons, there should also be room for improvement (i.e., variation across hospitals) and actionable care plans must exist.[50, 52] For most PROs, these associations have not yet been investigated. In addition, adequate case mix correction is required to enable fair comparisons and to draw conclusions about differences in healthcare quality. PROs and clinical outcomes may have different underlying mechanisms and also different case mix factors playing a role.[53] Identifying a sufficient set of case mix factors may be more challenging for PROs given the complexity and multidimensional character of PROs such as HRQOL.[54, 55] More research on which factors and through which mechanisms PROs are influenced may contribute to the selection of an adequate set of case mix factors.

FURTHER RESEARCH ON HOW TO IMPROVE PROS

The ultimate aim of using PROMs is to improve patient-relevant outcomes and healthcare quality. Building on our findings, future research should focus on investigating to what extent and how PROMs can be successfully incorporated and used in routine care to actually achieve these improvements. As part of this, our results emphasize the importance to explore how common and highly burdensome

symptoms (e.g., itching) can be improved, and suggest that better symptom management may also improve HRQOL. Our findings regarding the impact of itching, and the often co-occurring sleep problems and psychological symptoms, on HRQOL also highlighted the need for effective (para)medical treatment options (Chapter 6). Healthcare professionals also expressed the research priority that actionable care plans must be identified or developed to respond to individual PROM-results (Chapter 8). In Dutch nephrology care, the first steps towards such actionable care plans have been taken with the recently started research project 'Integrating Patient-Reported Outcome (PRO) measures into Dutch dialysis care: Toward a PRO Treatment Guide to achieve optimal multidisciplinary and personalized dialysis care' (PRO-GUIDE), in which a communication and treatment guide (and a supplementary generic toolbox for using PROMs in clinical practice) will be developed, in order to reduce the symptom burden and improve HRQOL.[56]

Future directions for further implementation of PROMs

TRAINING AND SUPPORT TO IMPROVE IMPLEMENTATION AND USE OF PROMS

Our research shows that there is a need for guidance and training for the implementation and use of PROMs in routine care. This applies to all steps: from getting the centres ready and inviting patients, to discussing individual PROM-results and taking and monitoring follow-up actions. Our results already provide a part of the necessary guidance and form the foundation for development of training and tools, including for instance practical implementation guidelines, training tools to support the interpretation of PROM-scores and to improve patient-professional communication about individual PROM-results.[15, 26, 56, 57] Training and tools can be of added value to both healthcare professionals and patients, especially when it comes to patient-professional communication. Guidance and training can help to optimize the use of PROMs and integrate it as a standard approach into healthcare.[26, 57] This will increase the added value and will help to reach the potential effects of using PROMs, contributing to person-centred care. In addition, to stimulate and facilitate improvements in using PROMs, a platform to easily share experiences and facilitating factors among centres may be of added value, to readily learn from each other.

Moreover, this dissertation comprises the implementation and use of PROMs in routine care until the discussion about individual PROM-results between patients and healthcare professionals. Although discussing individual PROM-results was regarded as an essential part in using PROMs and the main aim of completing PROMs, it does not stop after this step; using PROMs is an ongoing process and

also includes follow-up actions and monitoring. Guidelines are needed to adequately take this next step after discussing the results, and are yet to be developed as part of the PRO-GUIDE project.[56]

IMPROVE ACCESSIBILITY OF PROMS AND PROM-RESULTS

PROMs and PROM-results should be easily accessible for all patients and their healthcare professionals. Despite our efforts to optimize the implementation process (e.g., individual PROM-results were directly provided and PROMs were available in four languages), some challenges regarding technical issues and inclusivity remain. For example, an important and often mentioned barrier is that PROMs and PROM-results are not yet integrated into the electronic health record.[58] Patients and healthcare professionals need to work with a separate electronic system and PROM-results are therefore not easily combined with clinical measures and monitored over time. We expect a higher added value of PROMs in personalised and holistic treatment when fully incorporated into the standard workflow, and therefore, PROMs should be integrated into the electronic health record. Preferably as part of a clinical dashboard and in such a way that all steps are included: PROMs invitations and reminders are send (automatically) and results are collected and visualised within the electronic health record. Visualisation (for example through dynamic dashboards) should enable monitoring of individual PROM-scores over time in comparison to relevant clinical measures and PROM-results from similar patients (i.e., 'patients-like-me' information). Moreover, additional information (e.g., MIC and MDC) and explanations (e.g., colour-indications and meaning of scores) can support the interpretation of PROM-results, and may help both patients and healthcare professionals to identify outcomes that may require attention and discussion during consultations.

Furthermore, the PROMs and the methods of using PROMs in routine care setting must be further developed to improve the accessibility for all patients, including those with low (health) literacy, poor digital skills and language barriers. PROMs can help to start conversations about experiences and needs that may otherwise remain undiscussed, and this is perhaps of most added value for people who have difficulties with expressing themselves and with self-management.[59] However, these people might be those who also experience difficulties to complete PROMs.[60] This should be taken into account when taking the next steps in the implementation of PROMs into routine care. For example: simplify the language and layout with help of experts in this field[61], on short term where possible (e.g., PROM instructions and feedback) and after validation where necessary (e.g., PROM itself). In addition, when integrating PROMs into the electronic health record, one

should think of alternatives or extra help for people with poor digital skills. Explore for instance the possibilities to complete PROMs onsite (e.g., provide tablets and offer help to open the PROM), use of image answer options (e.g., smileys), and build-in read-aloud functions or video-instructions. Ideally, completing PROMs is facilitated in such a way that all patients can answer the questions by themselves, to ensure the results reflect their own perspective without any interpretation of others.[37, 39, 62]

BROADENING THE IMPLEMENTATION OF PROMS TO TOTAL NEPHROLOGY CARE AND BEYOND

During this research project, the PROMs were implemented into routine dialysis care. In addition, it laid the foundation for implementation of PROMs into the entire population in nephrology care. First steps have been taken to introduce PROMs for kidney transplant recipients (e.g., the 'PROs: Input of Valuable Endpoints' [POSITIVE] study[63]) and patients with CKD prior to kidney failure (e.g., the 'Patient-Relevant Outcomes in CKD' [PRO-CKD] project and the Dutch Outcome-based Healthcare program[64]), and will be included into the infrastructure of the Dutch renal registry Renine in the near future. This enables monitoring and individual follow-up of PROs – in addition to the already available clinical measures – over the entire disease and treatment trajectory (i.e., from advanced CKD to kidney failure and including different types of kidney replacement therapy and comprehensive conservative care). This will provide a more complete picture of disease patterns and can inform treatment decisions.

Literature shows that PROMs are also of added value in other fields, such as oncology, neurology and orthopaedics.[16, 41, 45, 65] Currently, different PROMs and separate workflows are often applied across the medical specialties. We hope that in the future, generic PROMs are no longer integrated into each medical specialty separately, but really organised around the patient, corresponding to the person-centred approach. This is especially of great added value in multimorbid populations, like patients with CKD in which comorbidities like diabetes and cardiovascular disease are common. Using the same PROM across medical specialties can lower the questionnaire burden and facilitate multidisciplinary use of the information, since the same PROM-results can be used by multiple specialists (i.e., reuse of information; there is no need to complete a PROM for each medical specialty separately). Particularly generic PROMs (e.g., SF-12 or PROMIS) are suitable for this broad application. First steps have already been taken to agree on suitable generic PROMs (e.g., PROMIS instruments) and crosswalks are being developed to enable transition between PROMs without losing historical PROM-data.[49, 66] Furthermore, the integration of PROMs into the electronic health record – although one of the biggest challenges – would bring

the person-centred collection of PROMs one step closer. Nevertheless, we expect that a certain generic approach will provide a good standard, but will not always be sufficient. Our results show that especially the Dialysis Symptom Index (DSI) – a disease specific PROM – provided valuable and additive insights, and was considered highly important by patients and healthcare professionals for the discussion and actionability of individual PROM-results (i.e., improvement of symptom management). PROMs specific for a certain disease or treatment may thus remain of great importance and an valuable addition to a generic PROM in routine care.

Conclusions

This dissertation provides insight into and practical knowledge of the implementation and use of PROMs in Dutch routine nephrology care. The introduction of PROMs into routine dialysis care corroborates the importance of a structured approach involving all relevant stakeholders (especially patients) and careful preparation (e.g., selection of PROMs and facilitate infrastructure to collect and use PROMs). Although further improvements in the implementation are required to enable valuable use of PROMs at aggregated level – for instance high and consistent response rates must be achieved – our results also show that PROMs are directly suitable and of added value for use at individual patient level. The high symptom burden (e.g., itching, sleep problems and psychological symptoms) and its impact on HRQOL in patients receiving dialysis treatment, highlights the need for recognition, discussion and effective treatment of PROs. Discussing individual PROM-results between patients and healthcare professionals is an essential part of using PROMs and facilitates patient-professional communication and shared decision making. Our results form the foundation for training and guidance for healthcare professionals and patients, and for further development (e.g., regarding ICT facilities and inclusivity) to optimize the use of PROMs in routine care. The ultimate aim of using PROMs is to improve patient-relevant outcomes, and to achieve this follow-up actions (i.e., monitoring over time and improve symptom management) in response to PROM-results are required. Finally, based on our highly positive and valuable experiences with PROMs in routine dialysis care, we continue optimizing the implementation of PROMs and expand on it by also including care for patients with advanced CKD and kidney transplant recipients. The presented approach can serve as an example and we hope that our results and lessons learned provide guidance to other researchers, policy makers, healthcare professionals and patients, within and beyond (inter)national nephrology care, regarding the implementation and use of PROMs in routine care.

References

1. Breckenridge K, Bekker HL, Gibbons E, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. *Nephrol Dial Transplant* 2015;30(10):1605-14

2. Verberne WR, Das-Gupta Z, Allegretti AS, et al. Development of an International Standard Set of Value-Based Outcome Measures for Patients With Chronic Kidney Disease: A Report of the International Consortium for Health Outcomes Measurement (ICHOM) CKD Working Group. *Am J Kidney Dis* 2019;73(3):372-84

3. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010;63(11):1179-94

4. Schick-Makaroff K, Wozniak LA, Short H, et al. How the Routine Use of Patient-Reported Outcome Measures for Hemodialysis Care Influences Patient-Clinician Communication: A Mixed-Methods Study. *Clin J Am Soc Nephrol* 2022

5. Flythe JE. Integrating PROMs in Routine Dialysis Care - The Devil Is in the (Implementation) Details. *Clinical Journal of the American Society of Nephrology* 2022:CJN.10840922

6. Foster A, Croot L, Brazier J, Harris J, O'Cathain A. 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

7. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23(6):508-18

8. van der Veer SN, Couchoud C, Morton RL. The role of kidney registries in expediting large-scale collection of patient-reported outcome measures for people with chronic kidney disease. *Clinical Kidney Journal* 2021;14(6):1495-503

9. Wiering B, de Boer D, Delnoij D. Patient involvement in the development of patient-reported outcome measures: a scoping review. *Health Expect* 2017;20(1):11-23

10. Tillett W, Adebajo A, Brooke M, et al. Patient involvement in outcome measures for psoriatic arthritis. *Curr Rheumatol Rep* 2014;16(5):418

11. Calvert M, Kyte D, Price G, Valderas JM, Hjollund NH. Maximising the impact of patient reported outcome assessment for patients and society. *Bmj* 2019;364:k5267

12. Jochemsen-van der Leeuw HG, van Dijk N, van Etten-Jamaludin FS, Wieringa-de Waard M. The attributes of the clinical trainer as a role model: a systematic review. *Acad Med* 2013;88(1):26-34

13. Bartlett SJ, Witter J, Cella D, Ahmed S. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 6: creating national initiatives to support development and use-the PROMIS example. *J Clin Epidemiol* 2017;89:148-53

14. Terwee CB, van der Willik EM, van Breda F, et al. Responsiveness and Minimal Important Change of seven PROMIS computerized adaptive tests (CAT) in patients with advanced chronic kidney disease. *Submitted*

15. Bausewein C, Simon ST, Benalia H, et al. Implementing patient reported outcome measures (PROMs) in palliative care--users' cry for help. *Health Qual Life Outcomes* 2011;9:27

16. Damman OC, Verbiest MEA, Vonk SI, et al. Using PROMs during routine medical consultations: The perspectives of people with Parkinson's disease and their health professionals. *Health Expect* 2019;22(5):939-51

17. Aiyegbusi OL, Kyte D, Cockwell P, et al. Patient and Clinician Perspectives on Electronic Patient-Reported Outcome Measures in the Management of Advanced CKD: A Qualitative Study. *Am J Kidney Dis* 2019;74(2):167-78

18. Hoekstra T, Heemskerk M, Berger S, Hemmelder M, de Vries A, ten Dam M: Nefrovisie jaarboek 2022: Trends in nierfunctievervanging in Nederland, Nefrovisie, 2022

19. Hutchings A, Neuburger J, van der Meulen J, Black N. Estimating recruitment rates for routine use of patient reported outcome measures and the impact on provider comparisons. *BMC Health Serv Res* 2014;14:66

20. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care* 1998;10(4):311-7

21. de Rooij ENM, Meuleman Y, de Fijter JW, et al. Symptom Burden before and after Dialysis Initiation in Older Patients. *Clin J Am Soc Nephrol* 2022;17(12):1719-29

22. de Rooij ENM, Meuleman Y, de Fijter JW, et al. Quality of Life before and after the Start of Dialysis in Older Patients. *Clin J Am Soc Nephrol* 2022;17(8):1159-67

23. Fincham JE. Response rates and responsiveness for surveys, standards, and the Journal. *Am J Pharm Educ* 2008;72(2):43

24. Rolfson O, Bohm E, Franklin P, et al. Patient-reported outcome measures in arthroplasty registries Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. *Acta Orthop* 2016;87 Suppl 1(Suppl 1):9-23

25. Flythe JE, Narendra JH, Hilliard T, et al. Cultivating a Research-Ready Dialysis Community. *J Am Soc Nephrol* 2019;30(3):375-80

26. Hardee JT, Rehring TF, Cassara JE, Weiss K, Perrine N. Effect and Durability of an In-depth Training Course on Physician Communication Skills. *Perm J* 2019;23

27. Nakash RA, Hutton JL, Jørstad-Stein EC, Gates S, Lamb SE. Maximising response to postal questionnaires--a systematic review of randomised trials in health research. *BMC Med Res Methodol* 2006;6:5

28. van der Horst DEM, van Uden-Kraan CF, Parent E, et al. Optimizing the use of patients' individual outcome information - Development and usability tests of a Chronic Kidney Disease dashboard. *Int J Med Inform* 2022;166:104838

29. Engels N, van der Nat PB, Ankersmid JW, et al. Development of an online patient decision aid for kidney failure treatment modality decisions. *BMC Nephrol* 2022;23(1):236

30. Brønserud MM, Iachina M, Green A, Groenvold M, Jakobsen E. Patient reported outcome data as performance indicators in surgically treated lung cancer patients. *Lung Cancer* 2019;130:143-8

31. Varagunam M, Hutchings A, Black N. Do patient-reported outcomes offer a more sensitive method for comparing the outcomes of consultants than mortality? A multilevel analysis of routine data. *BMJ Qual Saf* 2015;24(3):195-202

32. Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the Swedish National Quality Registers. *J Intern Med* 2016;279(2):141-53

33. Damman OC, Jani A, de Jong BA, et al. The use of PROMs and shared decision-making in medical encounters with patients: An opportunity to deliver value-based health care to patients. *J Eval Clin Pract* 2020;26(2):524-40

34. Schick-Makaroff K, Levay A, Thompson S, et al. An Evidence-Based Theory About PRO Use in Kidney Care: A Realist Synthesis. *The Patient - Patient-Centered Outcomes Research* 2022;15(1):21-38

35. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. *J Ren Care* 2013;39(3):140-50

36. Fletcher BR, Damery S, Aiyegbusi OL, et al. Symptom burden and health-related quality of life in chronic kidney disease: A global systematic review and meta-analysis. *PLoS Med* 2022;19(4):e1003954

37. Raj R, Ahuja KD, Frandsen M, Jose M. Symptoms and their recognition in adult haemodialysis patients: Interactions with quality of life. *Nephrology (Carlton)* 2017;22(3):228-33

38. Murtagh FE, Addington-Hall J, Higginson IJ. The prevalence of symptoms in end-stage renal disease: a systematic review. *Adv Chronic Kidney Dis* 2007;14(1):82-99

39. Weisbord SD, Fried LF, Mor MK, et al. Renal provider recognition of symptoms in patients on maintenance hemodialysis. *Clin J Am Soc Nephrol* 2007;2(5):960-7

40. McLaren S, Jhamb M, Unruh M. Using Patient-Reported Measures to Improve Outcomes in Kidney Disease. *Blood Purif* 2021;50(4-5):649-54

41. Basch E, Deal AM, Kris MG, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* 2016;34(6):557-65

42. Weisbord SD, Mor MK, Green JA, et al. Comparison of symptom management strategies for pain, erectile dysfunction, and depression in patients receiving chronic hemodialysis: a cluster randomized effectiveness trial. *Clin J Am Soc Nephrol* 2013;8(1):90-9

43. Valderas JM, Kotzeva A, Espallargues M, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008;17(2):179-93

44. Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? A systematic review. *Qual Life Res* 2013;22(9):2265-78

45. Kotronoulas G, Kearney N, Maguire R, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? A systematic review of controlled trials. *J Clin Oncol* 2014;32(14):1480-501

46. Fishbane S, Ross DW, Hong S. Anemia in Non-Dialysis-Dependent CKD: To Treat or Not to Treat? *American Journal of Kidney Diseases* 2019;73(3):297-9

47. Verberne WR, van den Wittenboer ID, Voorend CGN, et al. Health-related quality of life and symptoms of conservative care versus dialysis in patients with end-stage kidney disease: a systematic review. *Nephrol Dial Transplant* 2021;36(8):1418-33

48. Chiu YW, Teitelbaum I, Misra M, de Leon EM, Adzize T, Mehrotra R. Pill burden, adherence, hyperphosphatemia, and quality of life in maintenance dialysis patients. *Clin J Am Soc Nephrol* 2009;4(6):1089-96

49. Oude Voshaar M, Terwee CB, Haverman L, et al. Development of a standard set of PROs and generic PROMs for Dutch medical specialist care. *Quality of Life Research* 2023

50. Umeukeje EM, Nair D, Fissell RB, Cavanaugh KL. Incorporating patient-reported outcomes (PROs) into dialysis policy: Current initiatives, challenges, and opportunities. *Semin Dial* 2020;33(1):18-25

51. Greenhalgh J, Dalkin S, Gooding K, et al. Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Health Services and Delivery Research* 2017

52. Finkelstein FO, Finkelstein SH. Time to Rethink Our Approach to Patient-Reported Outcome Measures for ESRD. *Clin J Am Soc Nephrol* 2017;12(11):1885-8

53. Oemrawsingh A, van Leeuwen N, Venema E, et al. Value-based healthcare in ischemic stroke care: case-mix adjustment models for clinical and patient-reported outcomes. *BMC Medical Research Methodology* 2019;19(1):229

54. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q* 2014;92(4):754-75

55. Dennison CR. The role of patient-reported outcomes in evaluating the quality of oncology care. *Am J Manag Care* 2002;8(18 Suppl):S580-6

56. Meuleman Y. Integrating patient-reported outcome (PRO) measures into Dutch dialysis care: toward a PRO treatment guide to achieve optimal multidisciplinary and personalized dialysis The Netherlands, Dutch Kidney Foundation (project code: 21OM+002), 2022. https://www.nieren.nl/nieuws/3644-vermoeidheid-meten-en-verbeteren-het-pro-guide-project

57. Skovlund PC, Ravn S, Seibaek L, Thaysen HV, Lomborg K, Nielsen BK. The development of PROmunication: a training-tool for clinicians using patient-reported outcomes to promote patient-centred communication in clinical cancer settings. *J Patient Rep Outcomes* 2020;4(1):10

58. Anderson N, Kyte D, McMullan C, et al. Electronic patient-reported outcomes in chronic kidney disease. *Nat Rev Nephrol* 2022;18(12):739-40

59. Billany RE, Thopte A, Adenwalla SF, March DS, Burton JO, Graham-Brown MPM. Associations of health literacy with self-management behaviours and health outcomes in chronic kidney disease: a systematic review. *Journal of Nephrology* 2023

60. Nimmo A, Bell S, Brunton C, et al. Collection and determinants of patient reported outcome measures in haemodialysis patients in Scotland. *Qjm* 2018;111(1):15-21

61. Pharos. Sneltest: test je vragenlijst op begrijpelijkheid.

https://www.pharos.nl/kennisbank/test-je-vragenlijst-op-begrijpelijkheid/

62. Xiao C, Polomano R, Bruner DW. Comparison between patient-reported and clinician-observed symptoms in oncology. *Cancer Nurs* 2013;36(6):E1-e16

63. Wang Y, Snoep JD, Hemmelder MH, et al. Outcomes after kidney transplantation, let's focus on the patients' perspectives. *Clin Kidney J* 2021;14(6):1504-13

64. Programma Uitkomstgerichte Zorg – Lijn 1 'Meer inzicht in uitkomsten': Eindrapport aandoeningswerkgroep Chronische Nierschade. Nederland 2022.

65. Makhni EC. Meaningful Clinical Applications of Patient-Reported Outcome Measures in Orthopaedics. *J Bone Joint Surg Am* 2021;103(1):84-91

66. Schalet BD, Lim S, Cella D, Choi SW. Linking scores with patient-reported health outcome instruments: a validation study and comparison of three linking methods. *Psychometrika* 2021;86(3):717-46

9

Appendices

Nederlandse samenvatting

Dutch summary

In de gezondheidszorg is er steeds meer aandacht voor het patiëntenperspectief, gericht op een persoonlijke en holistische behandeling die aansluit bij de voorkeuren en behoeften van de patiënt. Het gebruik van patiënt-gerapporteerde uitkomstmaten (PROM's) in de praktijk kan bijdragen aan deze meer patiëntgerichte benadering. Maar hoe integreer je PROM's in de gebruikelijke zorg en hoe gebruik je PROM's om tot deze persoonlijke en holistische behandeling te komen? Dit proefschrift biedt inzicht en praktische kennis over de implementatie en het gebruik van PROM's binnen de nefrologische zorg.

Achtergrond

Mensen met chronische nierschade hebben een verminderde nierfunctie. Vaak gaat de nierfunctie in de loop van de tijd verder achteruit, waardoor nierfalen kan ontstaan. Mensen staan dan voor de keuze tussen het starten van nierfunctievervangende behandelingen of conservatieve behandeling. Er zijn twee soorten nierfunctievervangende behandelingen, namelijk niertransplantatie of dialyse. Dialyse is een intensieve behandeling die mensen met nierfalen drie keer per week tot dagelijks in het ziekenhuis of thuis ondergaan. Chronische nierschade en in het bijzonder dialyse hebben een grote impact op het dagelijks leven. Patiënten ervaren vaak fysieke en emotionele gezondheidsklachten, zoals vermoeidheid, jeuk, spierkrampen, slaapproblemen, seksuele problemen en depressieve klachten, met een grote impact op de ervaren kwaliteit van leven.

De nefrologische zorg (nierziekten zorg) richt zich van oudsher op medische uitkomsten zoals sterfte, nierfunctie, bloeddruk en hart- en vaatziekten. Hoewel patiënt-gerapporteerde uitkomsten (PRO's), zoals gezondheidsklachten en kwaliteit van leven, belangrijk worden gevonden door patiënten en zorgverleners, worden deze uitkomsten vaak onvoldoende herkend of genoemd in de spreekkamer, en blijven daarmee onbesproken en onbehandeld in de gebruikelijke zorg. Het systematisch meten van PRO's kan ervoor zorgen dat deze uitkomsten vaker en sneller herkend worden in de praktijk. PRO's kunnen systematisch worden gemeten door middel van vragenlijsten, ook wel PROM's genoemd (naar de Engelse term 'patient-reported outcome measures'). PROM's zijn specifiek ontwikkeld om de ervaren gezondheid vanuit het perspectief van de patiënt te meten.

PROM's kunnen een completer beeld geven van hoe het met de patiënt gaat of wat een patiënt kan verwachten van de behandeling of het verloop van de ziekte, in aanvulling op de medische uitkomsten. Je kan hierbij gebruik maken van de re-

sultaten op het niveau van de individuele patiënt en op het niveau van de gehele of een groter deel van de patiëntengroep. Op individueel niveau kan het gebruik van PROM's bijvoorbeeld bijdragen aan het 'samen beslissen' door de patiënt en de zorgverlener in de spreekkamer. Een beter inzicht in de uitkomsten die de patiënt ervaart kan ondersteuning bieden bij het gesprek over wat belangrijk is voor de patiënt en welke (behandel)keuzes bij hem of haar passen. Op groepsniveau kunnen PROM's informatie opleveren over welke uitkomsten patiënten kunnen verwachten en over welke factoren en behandelingen de PRO's kunnen beïnvloeden. Daarnaast zouden PROM's-resultaten op groepsniveau ook gebruikt kunnen worden voor de evaluatie van de kwaliteit van zorg door resultaten van zorginstellingen te vergelijken.

Ondanks dat er steeds meer aandacht en draagvlak is voor een meer patiëntgerichte benadering, werden PROM's nog niet breed toegepast in de nefrologische zorg. In Nederland zijn we daarom een project gestart voor de ontwikkeling en landelijke implementatie van PROM's in de nefrologische zorg (het PROMs-NNL project). Het PROMs-NNL project is een initiatief van Nierpatiënten Vereniging Nederland en wordt uitgevoerd in nauwe samenwerking met onder andere patiënten, zorgverleners (Nederlandse Federatie voor Nefrologie), onderzoekers (Klinische Epidemiologie, Leids Universitair Medisch Centrum) en het kwaliteitsinstituut Nefrovisie. De implementatie richt zich in eerste instantie op de dialyse zorg en omvat de selectie van PROM's, het pilot testen van PROM's in de praktijk en de landelijke implementatie en het gebruik van PROM's in de nefrologische zorg.

Het doel van dit proefschrift is om inzicht en praktisch toepasbare kennis te verkrijgen voor de implementatie en het gebruik van PROM's in de nefrologische zorg.

Implementatie van PROM's in de nefrologische zorg

SELECTIE VAN PROM'S

Op basis van de literatuur en in samenwerking met patiëntvertegenwoordigers en zorgverleners zijn gezondheids-gerelateerde kwaliteit van leven en ziekte-specifieke gezondheidsklachten geïdentificeerd als belangrijke PRO's om te meten binnen de gebruikelijke nefrologische zorg **(Hoofdstuk 1)**. De volgende stap is de selectie van geschikte PROM's om deze PRO's te meten. In **Hoofdstuk 2** beschreven we de selectie van een geschikte PROM om gezondheidsklachten te meten binnen de nefrologische zorg. We hebben deze selectie in vier fasen uitgevoerd. In de eerste twee fasen hebben we gezocht naar en voortgebouwd op de bestaande literatuur, waaruit we 28 potentieel geschikte vragenlijsten en 10 categorieën van gezondheidsklachten identificeerden. Tijdens de derde fase

werden de vragenlijsten beoordeeld op basis van vooraf gestelde criteria met betrekking tot de relevantie (bijv. van toepassing voor mensen met chronische nierschade of nierfalen), volledigheid (bijv. omvat 90% van de categorieën) en begrijpelijkheid (bijv. passende lengte en eenvoudige en duidelijke vragen). Twee vragenlijsten voldeden aan de criteria: de 'Dialysis Symptom Index' (DSI) en de 'Palliative Care Outcome Scale-Renal Version' (IPOS-Renal). In de vierde fase werden deze vragenlijsten beoordeeld door twee panels van in totaal 151 patiënten die willekeurig een vragenlijst kregen toegewezen, en één panel van zes patiënt-experts (d.w.z. ervaren vragenlijstbeoordelaars) die beide vragenlijsten vergeleken. Patiënten rapporteerden meer gezondheidsklachten met de DSI dan met de IPOS-Renal (12 versus 8 gezondheidsklachten) en hadden minder tijd nodig om de DSI in te vullen (5,4 versus 7,5 minuten). Zowel de patiënten als de experts beoordeelden de DSI als de meest complete, specifieke en begrijpelijke vragenlijst. De DSI werd daarom geselecteerd als PROM voor het meten van gezondheidsklachten binnen de nefrologische zorg.

De '12-item Short Form Health Survey' (SF-12) is een gevalideerde en veelgebruikte PROM om gezondheids-gerelateerde kwaliteit van leven te meten. De SF-12 wordt geadviseerd voor gebruik in de nefrologische zorg en werd geselecteerd door onze panels met patiënten en experts. Recentelijk werden ook de 'Patient-Reported Outcomes Measurement Information System' (PROMIS) instrumenten aanbevolen, welke ook als computer adaptieve test (CAT) kunnen worden afgenomen. Een CAT selecteert vragen op basis van eerdere antwoorden van de patiënt, waardoor naar verwachting net zo precies (of preciezer) gemeten kan worden met minder vragen vergeleken met 'standaard' (niet-adaptieve) PROM's. Om deze relatief nieuwe methode te onderzoeken, hebben we in **Hoofdstuk 3** de psychometrische eigenschappen (meeteigenschappen) van zeven PROMIS CAT's vergeleken met de SF-12 in patiënten met chronische nierschade. De zeven PROMIS CAT's bevatten vragen over fysiek functioneren, belemmeringen door pijn, vermoeidheid, slaapstoornissen, angst, depressie en het vermogen om deel te nemen aan sociale rollen en activiteiten. We vergeleken de inhoud van de zeven PROMIS CAT's met de SF-12 en onderzochten de constructvaliditeit en test-hertest betrouwbaarheid. De zeven PROMIS CAT's toonden voldoende constructvaliditeit. Verder was ook de test-hertest betrouwbaarheid van de PROMIS CAT's, de overkoepelende scores van de SF-12, de meeste domeinscores van de SF-12 en de DSI voldoende. Over het algemeen was de betrouwbaarheid van PROMIS CAT's beter dan van de SF-12, wat gepaard gaat met een lagere minimale detecteerbare verandering (MDC). Voor zeven PROMIS CAT's waren echter 45 items (10 minuten) nodig, wat 3 tot 4 keer de lengte is van de SF-12 (12 items; 3 minu-

ten). Kortom, deze resultaten tonen bewijs voor voldoende constructvaliditeit en een betere test-hertest betrouwbaarheid van zeven PROMIS CAT's, maar vereisen meer items in vergelijking met de SF-12. Deze resultaten geven nog geen inzicht in de geschiktheid en haalbaarheid van PROMIS CAT's in de gebruikelijke nefrologische zorg en daarom blijft de SF-12 voorlopig behouden in de Nederlandse nefrologie.

PILOT TESTEN EN LANDELIJKE IMPLEMENTATIE VAN PROM'S

Hoofdstuk 4 beschrijft de eerste ervaringen en resultaten van de introductie van PROM's in de Nederlandse nefrologische zorg in de vorm van een pilotstudie waaraan 16 dialysecentra deelnamen. Door middel van kwantitatieve en kwalitatieve methoden hebben we deze eerste ervaringen onderzocht. In totaal vulden 512 patiënten 908 PROM's in, verspreid over drie meetmomenten. Er was een gemiddeld laag responspercentage van 36%, welke varieerde van 6% tot 70% tussen centra. De grote variatie tussen de centra laat zien dat het behalen van een hoog responspercentage haalbaar maar uitdagend is en mogelijk om meer ondersteuning en training van patiënten en zorgverleners vraagt. Verder lieten de kwantitatieve resultaten zien dat patiënten die dialyseren een substantieel lagere fysieke kwaliteit van leven hebben. Hierbij ervaren zij een hoge belasting van gezondheidsklachten met gemiddeld 11 verschillende gezondheidsklachten waar patiënten 'nogal wat last' van ervaren. De variatie in frequentie en belasting van gezondheidsklachten suggereert dat de meest voorkomende niet per se de meest belastende gezondheidsklachten zijn. In het kwalitatieve deel onderzochten we de ervaringen en voorkeuren van patiënten en zorgverleners met betrekking tot het gebruik van PROM's. Patiënten waardeerden de inhoud, lengte en structuur van de gebruikte PROM's (DSI en SF-12) en het online invullen van de PROM's. Verder kwam naar voren dat individuele terugkoppeling van resultaten in een relevante context moet worden gepresenteerd, bijvoorbeeld door vergelijking met gemiddelde scores van andere dialysepatiënten in Nederland. Ook werd aangegeven dat het gebruik van PROM's kan bijdragen aan (een goede voorbereiding van) het gesprek in de spreekkamer. Tijdens deze pilotstudie bespraken enkele patiënten hun PROM-resultaten al met hun zorgverlener. Patiënten en zorgverleners gaven aan dat het bespreken van de resultaten over kwaliteit van leven en gezondheidsklachten zeer inzichtelijk en waardevol was, en dat het bespreken van de individuele resultaten essentieel is. De eerste indrukken van het bespreken van de PROM-resultaten waren veelbelovend.

Voortbouwend op de bevindingen uit de pilotstudie is de infrastructuur voor de PROM's verder geoptimaliseerd voor landelijke implementatie en het gebruik van

PROM's in de reguliere dialysezorg. Vervolgens werden de PROM's eind 2018 via Nefrovisie beschikbaar voor alle dialysecentra in Nederland en opgenomen in de landelijke kwaliteitsregistratie Renine. Alle dialysecentra werden uitgenodigd om PROM's te implementeren in de dagelijkse dialysezorg.

Gebruik van PROM's in de nefrologische zorg

GEBRUIK VAN PROM'S OP GROEPSNIVEAU

Op groepsniveau kunnen PROM-resultaten worden gebruikt om de kwaliteit van zorg te evalueren en om patiënten en zorgverleners te informeren over de effecten van behandeling en het verloop van de ziekte. Een funnelplot is een grafische methode die gebruikt kan worden om de kwaliteit van zorg in ziekenhuizen te vergelijken. In **Hoofdstuk 5** gaan we in op het gebruik en de interpretatie van funnelplots door een overzicht te geven van de basisprincipes, valkuilen en overwegingen bij toepassing op PRO's, aan de hand van voorbeelden uit de Nederlandse dialysezorg. Een funnelplot heeft verschillende voordelen, waaronder duidelijke visualisatie van de precisie, waarneembaarheid van volume-effecten, ontmoediging van het rangschikken van ziekenhuizen en een eenvoudige interpretatie van resultaten. Echter, zonder voldoende kennis van de onderliggende methode stap je gemakkelijk in valkuilen, zoals over-interpretatie van gestandaardiseerde scores, onjuiste directe vergelijkingen tussen ziekenhuizen en aannemen dat een ziekenhuis naar behoren (d.w.z. zoals verwacht) presteert op basis van vergelijkingen met onvoldoende bewijskracht. Daarnaast gaat de toepassing van funnelplots voor PRO's gepaard met extra uitdagingen gerelateerd aan het multidimensionale karakter van sommige PRO's en uitdagingen bij het meten van PRO's. Om relevante en eerlijke vergelijkingen van PRO's te kunnen maken, zijn hoge en consistente responspercentages, adequate casemixcorrectie en kwalitatief goede PROM's nodig. Deze uitdagingen moeten worden geadresseerd voordat data over PRO's ingezet worden voor evaluatie van kwaliteit van zorg, bijvoorbeeld door gebruik te maken van funnelplots.

Hoofdstuk 6 laat een voorbeeld zien van hoe PROM-resultaten op groepsniveau kunnen worden gebruikt om patiënten en zorgverleners te informeren over het beloop en over de effecten van ziekte en uitkomsten. In dit hoofdstuk onderzochten we de impact van jeuk op gezondheids-gerelateerde kwaliteit van leven en de interacties met slaapproblemen en psychologische gezondheidsklachten bij patiënten die dialyseren. Dit is gedaan met cross-sectionele en longitudinale analyses bij 2978 patiënten die de PROM's tussen 2018 en 2020 hebben ingevuld. Onze resultaten toonde aan dat de helft van de patiënten jeuk ervaarde en dat bij 70% van hen de jeuk over de tijd aanhield. Patiënten met jeuk ervaarden een

3 tot 4 punten lagere fysieke en mentale kwaliteit van leven in vergelijking met patiënten zonder jeuk. Dit verschil bleef gedurende de 2 jaar follow-up stabiel. Verder kwamen slaapproblemen (70% versus 52%) en psychologische klachten (36% versus 19%) vaker voor bij patiënten met jeuk. Deze klachten hadden een bijkomend negatief effect op de fysieke en mentale kwaliteit van leven, maar er was geen interactie met jeuk. Dat betekent dat de combinatie van beide klachten niet in een significant lagere of hogere kwaliteit van leven resulteerde dan de som van de individuele effecten van de klachten. De hoge prevalentie en de aanhoudendheid van de jeuk, de impact van jeuk op kwaliteit van leven over de tijd en het bijkomende effect op kwaliteit van leven van de vaak gelijktijdig voorkomende slaapproblemen en psychologische klachten, benadrukken de noodzaak van herkenning en effectieve behandeling van jeuk. Dit is van belang om de belasting door gezondheidsklachten te verminderen en de kwaliteit van leven te verbeteren bij patiënten die dialyseren.

GEBRUIK VAN PROM'S OP INDIVIDUEEL NIVEAU

Voor optimaal gebruik van PROM's bij individuele patiënten is kennis nodig over hoe PROM-resultaten geïnterpreteerd en besproken kunnen worden. In **Hoofdstuk 7** gaan we in op de verschillende typen en kenmerken van PROM's en bieden we handvatten voor de interpretatie van individuele PROM-scores en veranderingen in PROM-scores over de tijd. In dit hoofdstuk introduceren we typen en kenmerken zoals generieke en specifieke PROM's en scoringsmethoden van PROM's. We leggen uit dat intuïtieve maten zoals informatie over het gemiddelde en de verdeling van PROM-scores in een referentiepopulatie of de vergelijking met meer bekende uitkomsten (bijv. bloedwaarden en bloeddruk) onmisbaar zijn om PROM-scores te interpreteren en bekend te raken met PROM-scores. Verder zijn methodologische maten zoals de MDC en de minimale belangrijke verandering (MIC) belangrijk om ons te informeren over respectievelijk statistisch en klinisch relevante veranderingen. Daarnaast is het belangrijk om ervan bewust te zijn dat een responsverschuiving kan optreden. Dit verwijst naar een verandering in de betekenis van de PRO voor de patiënt (bijv. een verandering in iemands perceptie van kwaliteit van leven). Een responsverschuiving kan onverwacht kleine (of grote) veranderingen in PROM-scores verklaren. Ten slotte is een gesprek met de patiënt van belang om individuele PROM-scores te interpreteren. De beste manier om individuele PROM-scores en veranderingen in PROM-scores te interpreteren, is door middel van een gesprek tussen de patiënt en de zorgverlener, waarbij maten zoals MDC, MIC en responsverschuiving een faciliterende rol kunnen spelen. Bijvoorbeeld: de MIC geeft een indicatie van welke veranderingen in PRO's er

meestal toe doen voor patiënten en uit het gesprek over individuele resultaten komt naar voren welke veranderingen, in welke mate en op welke manier belangrijk zijn voor een individuele patiënt.

In **Hoofdstuk 8** hebben we onderzocht hoe individuele PROM-resultaten het beste besproken kunnen worden in de spreekkamer. We onderzochten dit door middel van semigestructureerde interviews met 22 dialysepatiënten en zorgverleners over hun ervaringen met en ideeën over het bespreken van PROM-resultaten in de dagelijkse dialysezorg. Interviews waren gericht op algemene situaties en specifieke situaties (bijv. het bespreken van gevoelige onderwerpen of wanneer er geen medische behandeling beschikbaar is). Patiënten en zorgverleners (nefrologen en verpleegkundigen) waardeerden het gebruik van PROM's zeer, omdat het inzicht en overzicht geeft van hoe het met de patiënt gaat en omdat het bijdraagt aan de communicatie tussen de patiënt en de zorgverlener. Patiënten en zorgverleners boden praktische handvatten voor een optimaal gesprek over individuele PROM-resultaten. Ten eerste benadrukten patiënten en zorgverleners dat PROM-resultaten altijd besproken moeten worden en gaven ze aan hoe een geschikte setting te creëren, adequaat voor te bereiden, om te gaan met tijdsdruk en PROM's te gebruiken als een hulpmiddel voor persoonlijke holistische zorg. Ten tweede moeten patiënten actief deelnemen en moeten zorgverleners een begeleidende rol aannemen in het gesprek. Een vertrouwensband tussen de patiënt en de zorgverlener werd als een vereiste beschouwd en de interactie tussen patiënt en zorgverlener werd beschreven als een samenwerking waarin beiden hun kennis, ervaringen en ideeën inbrengen. Ten derde werden vervolgstappen na het bespreken van de PROM-resultaten belangrijk gevonden, inclusief het handelen naar de resultaten (bijv. doorverwijzing of behandeling van gezondheidsklachten) en monitoring van de PROM-resultaten en behandeleffecten over de tijd, die structureel zijn ingebed in het multidisciplinaire behandelproces. Deze algemene thema's waren ook van toepassing op de specifieke situaties, bijvoorbeeld: resultaten moeten ook besproken worden als er geen medische behandeling beschikbaar is. Opvallend was dat van zorgverleners verwacht wordt dat zij meer initiatief en een leidende rol nemen bij het bespreken van gevoelige onderwerpen. Dit onderzoek biedt diepgaande kennis en praktische handvatten voor het organiseren en voeren van gesprekken over individuele PROM-resultaten in de reguliere nefrologische zorg.

Conclusie

Dit proefschrift biedt inzicht en praktisch toepasbare kennis over de implementatie en het gebruik van PROM's in de Nederlandse nefrologische zorg. Optimale implementatie van PROM's vraagt om een doorlopend en herhalend proces van ontwikkelen, proberen, leren en verbeteren. De introductie van PROM's binnen de dialysezorg onderschreef het belang van een gestructureerde aanpak met betrokkenheid van alle relevante belanghebbenden (in het bijzonder patiënten) en met een zorgvuldige voorbereiding (bijv. de selectie van PROM's en het faciliteren van de infrastructuur voor het verzamelen en gebruiken van PROM's). Hoewel verdere verbeteringen in de implementatie nodig zijn voor waardevolle informatie op groepsniveau (o.a. hoge en consistente respons is noodzakelijk) laten onze resultaten ook zien dat PROM's direct geschikt en van toegevoegde waarde zijn voor gebruik op individueel niveau. De hoge ervaren last door gezondheidsklachten (bijv. jeuk, slaapproblemen en psychologische klachten) en de impact daarvan op de kwaliteit van leven bij patiënten die een dialysebehandeling ondergaan, benadrukt de noodzaak van herkenning, bespreking en effectieve behandeling van PRO's. Het bespreken van individuele PROM-resultaten tussen patiënten en zorgverleners is een essentieel onderdeel van het gebruik van PROM's en ondersteunt de communicatie tussen patiënt en zorgverlener en draagt bij aan samen beslissen. Onze resultaten vormen de basis voor training en begeleiding van zorgverleners en patiënten en voor verdere ontwikkeling (o.a. op het gebied van ICT-voorzieningen en toegankelijkheid voor laaggeletterden en anderstaligen) om het gebruik van PROM's te optimaliseren. Het uiteindelijke doel van het gebruik van PROM's is het verbeteren van patiëntrelevante uitkomsten. Om dit te bereiken zijn vervolgacties naar aanleiding van de PROM-resultaten nodig, zoals monitoring over de tijd en verbeteren van behandeling van gezondheidsklachten. Ten slotte blijven we de implementatie van PROM's verder optimaliseren op basis van onze zeer positieve en waardevolle ervaringen met PROM's in de dialysezorg. Daarnaast zullen we de implementatie van PROM's uitbreiden door ook de zorg voor patiënten met vergevorderde chronische nierschade en patiënten met een transplantatienier hierin mee te nemen. De aanpak beschreven in dit proefschrift kan als voorbeeld dienen, en we hopen dat onze resultaten en geleerde lessen handvatten bieden aan andere onderzoekers, beleidsmakers, zorgverleners en patiënten, zowel binnen als buiten de (inter)nationale nefrologische zorg, met betrekking tot de implementatie en het gebruik van PROM's binnen de gebruikelijke zorg.

Dankwoord

Acknowledgements

Dit proefschrift is mede tot stand gekomen dankzij de steun, inspiratie en inzet van velen. Graag wil ik iedereen bedanken die hieraan heeft bijgedragen. Een aantal mensen bedank ik graag in het bijzonder.

Allereerst alle patiënten en zorgverleners die het aandurfden om 'iets nieuws' te proberen in de dagelijkse dialysezorg en hun ervaringen en ideeën met ons wilden delen. Jullie inzet, gedrevenheid en positiviteit is niet alleen een heel belangrijke drijfveer in de implementatie van PROM's, maar ook een heel belangrijke motivator voor mij in dit promotietraject.

Mijn promotoren Prof. Friedo Dekker en Prof. Marc Hemmelder, en co-promotor Dr. Yvette Meuleman. Bedankt voor jullie fijne begeleiding, waarin ik altijd bij jullie terecht kon en er ruimte was om te sparren. Friedo, ik heb enorm veel geleerd van onze gesprekken, het brainstormen en het delen van de verschillende perspectieven. Jouw enthousiasme en actieve meedenken zijn heel inspirerend en hebben mij zeer aangemoedigd tijdens mijn promotietraject. Marc, bedankt voor jouw rust en vertrouwen tijdens dit mooie PROM's-avontuur, waarin jij altijd in mogelijkheden denkt. Met jou aan het roer was elke stap in het PROMs-NNL project 'te overzien'. Yvette, zonder jou was dit proefschrift er zeker niet geweest. Ik heb ontzettend genoten van onze samenwerking en had van geen betere begeleiding durven dromen. Bedankt voor alles wat ik van jou heb mogen leren als wetenschapper én als mens, dat is absoluut uniek en enorm waardevol!

Daarnaast wil ik ook de andere leden van het PROMs-NNL team bedanken: Hans Bart, Prof. Frans van Ittersum, Prof. Willem Jan Bos, Marc ten Dam, Martijn Leegte en Karen Prantl. Dit team was en is de fundering voor de implementatie en het gebruik van PROM's binnen de nefrologie, en ik beschouw jullie ook zeker als onderdeel van mijn promotieteam.

Ook veel dank aan de teams van Nefrovisie en NVN, en de vele onderzoekers met wie ik met veel plezier heb mogen samenwerken. Waaronder het Nefro-PROMIS team in het Amsterdam UMC: Prof. Caroline Terwee, Prof. Frans van Ittersum, Dr. Fenna van Breda, Dr. Brigit van Jaarsveld, Marlon van de Putte en Isabelle Jetten. Bedankt voor het warme welkom, jullie inzet en creativiteit om binnen korte tijd en met de uitdagingen van de coronapandemie het project succesvol af te ronden. Jet, bedankt voor de fijne samenwerking in het kwalitatief onderzoek. Ik vond het een zeer inspirerende en leerzame ontdekkingstocht!

Ik wil graag de Epi-promovendi bedanken voor de leuke samenwerking en gezelligheid. De 'koffie en koekjes', pizzalunches en borrels brachten niet alleen veel

&

gezelligheid, maar ook inspiratie voor het onderzoek. Ook wil ik de Nefro Journal Club bedanken voor alle leuke en leerzame discussies over onderzoek. Chava, Ype, Yiman, Linda, Bas, Esther, Carolina, Carlijn en Els, bedankt voor de fijne koffiemomenten met goede gesprekken, het luisteren en meedenken, en de aanmoediging wanneer dit nodig was. Tamara, bedankt voor al je hulp en goede zorgen tijdens mijn promotietraject.

De collega's en oud-collega's van UGZ (ZiNL) en DICA, van wie en met wie ik veel heb geleerd en gelachen. Wat is het leuk om in zulke gedreven en energieke teams te werken aan een betere kwaliteit van zorg, en enorm inspirerend om te beginnen aan en bij het afronden van dit proefschrift.

Verder wil ik graag mijn vrienden bedanken voor alle steun, aanmoediging en gezelligheid. Lieve Bella's Josine, Simone, Ghislaine, Tessa en Merel, we zijn een kleurrijk geheel. Bedankt dat ik met jullie kan lachen, praten, ontspannen en dansen. Simone, bedankt voor jouw creatieve inspiratie en jouw hand aan het design van dit proefschrift. Chava en Daan, bedankt dat jullie mijn paranimfen willen zijn op deze bijzondere dag!

Lieve familie en schoonfamilie, bedankt voor jullie steun, vertrouwen, warmte en gezelligheid. Lieve papa, mama, Tinke, Daan, Pim, Sijmen, Mandy en Sandra, wat heb ik het bijzonder goed met jullie! Liefste zus en broer, bedankt voor jullie warmte, humor, creativiteit en goede gesprekken. Ik geniet ontzettend van alles wat we samen ondernemen! Liefste papa en mama, bedankt dat ik voor werkelijk alles bij jullie terecht kan. Bij jullie is het altijd thuiskomen, bedankt voor jullie onvoorwaardelijke liefde, steun en vertrouwen!

Liefste Dennis, jij bent alles voor mij. Bedankt voor wie jij bent en wie je mij laat zijn. Met jou is het elke dag genieten en kijk ik uit naar elk nieuw avontuur, groot en klein!

Curriculum Vitae

Elisabeth (Esmee) Maria van der Willik werd geboren op 23 maart 1992 te Alkemade. In 2010 behaalde zij haar gymnasium diploma aan het Visser 't Hooft Lyceum in Leiden, waarna zij startte met de studie Gezondheidswetenschappen aan de Vrije Universiteit in Amsterdam. Tijdens haar studie kreeg zij de kans om ervaring op te doen in de dagelijkse zorg in haar werk in een huisartsenpraktijk. Daarnaast heeft zij tijdens de Master samen met medestudenten 'Sam de Waterman' bedacht en ontwikkeld: een methode die basisschooldocenten ondersteunt in het stimuleren en implementeren van een gezond drinkbeleid in peuter- en kleuterklassen, onder andere aan de hand van een voorleesboek. Sam de Waterman is ook na de studie verder ontwikkeld en ingezet op basisscholen in onder andere Amsterdam en Zeeland. In 2014 rondde Esmee haar Master Gezondheidswetenschappen Cum Laude af met een afstudeeronderzoek naar borstvoeding en overgewicht bij baby's en jonge kinderen, wat resulteerde in haar eerste wetenschappelijk publicatie. Met deze positieve ervaring werd bij haar ook de interesse gewekt in het epidemiologisch onderzoek. De interesse in de zorg en het werk in de huisartsenpraktijk maakte dat zij besloot eerst meer ervaring op te willen doen en inzicht te krijgen in wat er in de (klinische) praktijk gebeurt met het onderzoek en resultaten over de zorg. In 2014 startte zij als medisch datamanager op het wetenschappelijk bureau van Dutch Institute for Clinical Auditing (DICA) in Leiden. Hier maakte zij kennis met de evaluatie van kwaliteit van zorg en werkte zij aan het inzichtelijk maken hiervan door middel van terugkoppeling van uitkomsten van zorg aan ziekenhuizen. Binnen DICA groeide zij door tot teamleider van het team datamanagers. Daarnaast is zij zich ook op het gebied van onderzoeksmethoden verder blijven ontwikkelen, onder andere door de cursus 'Onderzoeksdesign en -analyses' van de afdeling Klinische Epidemiologie (Leids Universitair Medisch Centrum, LUMC) op Schiermonnikoog te volgen. Hier kwam zij in contact met Prof. Friedo Dekker, die haar een jaar na de cursus (2017) uitnodigde om te komen praten over een project met betrekking tot de ontwikkeling en implementatie van patiënt-gerapporteerde uitkomstmaten (PROM's) in de nefrologie. Hier kwam het promotietraject uit voort bij de Klinische Epidemiologie in het LUMC in Leiden, waar zij in 2017 mee startte. Hierin deed zij kwantitatief en kwalitatief onderzoek naar de landelijke implementatie en het gebruik van PROM's binnen de nefrologische zorg. Ook was zij in 2019-2021 onderzoeker voor een validatiestudie naar PROM's binnen de nefrologie vanuit de afdeling Epidemiologie en Data Science van de Amsterdam Universitair Medische Centra (AUMC). Het tijdelijk stilvallen van verschillende onderzoeksactiviteiten (o.a. dataverzameling in de zorg) tijdens het eerste jaar van de coronapandemie (2020) bood de mogelijkheid om als on-

derzoeker mee te werken aan de BCG-PRIME studie, en daarmee ervaring op te doen met de uitvoer van een RCT. Verder volgde zij tijdens haar promotietraject de opleiding tot Epidemioloog B bij de Klinische Epidemiologie in het LUMC. Sinds april 2022 is Esmee werkzaam als senior adviseur bij Zorginstituut Nederland in de rol van methodoloog bij het programma Uitkomstgerichte Zorg. Haar eerdere werkervaring op het gebied van samen beslissen en evaluatie van kwaliteit van zorg komen hierin goed van pas. Ook in volgende uitdagingen hoopt zij zich verder te ontwikkelen en haar ervaringen in te kunnen zetten om bij te dragen aan een steeds betere zorg.

PhD portfolio

Courses	Hours	Year
Study design and analysis (Schiermonnikoog)	56	2016
Basic principles of epidemiological research (Rothman)	84	2017
Basic methods and reasoning in biostatistics	42	2018
Statistical aspects of clinical trials	28	2018
Regression analysis	42	2018
Survival analysis	42	2018
Meta-analysis	28	2018
Clinical epidemiology (Grobbee)	84	2018
Advanced epidemiological methods (Poelgeest)	56	2018
Winterschool Dutch Kidney Foundation	28	2018
Working group guidance	12	2018
Analysis of repeated measurements	42	2019
Causal inference (Hernan)	84	2019
Clinimetrics: assessing measurement properties of health measurement instruments	84	2019
BROK course	42	2020
Capita selecta	56	2021

Teaching activities	Hours	Year
Introduction into clinical scientific research (KWO) (Honours programme, BSc Medicine, year 2)	2	2017
Academic and scientific training year 1 (AWVI) (BSc Medicine)	2	2018
Frontiers of science course (FOS)	4	2018
Academic and scientific training year 2 (AWVII) (BSc Medicine)	8	2018
Methods and techniques (BSc Biomedical sciences, year 1)	4	2019
Basic principles of epidemiological research (Rothman) for honours students	32	2019
Clinical research in practice (CRIP) (MSc Biomedical sciences)	4	2019
Academic and scientific training year 2 (AWVII) (BSc Medicine)	8	2019
Academic and scientific training year 1 (AWVI) (BSc Medicine)	28	2020
Clinical research in practice (CRIP) (MSc Biomedical sciences)	5	2020
Academic and scientific training year 2 (AWVII) (BSc Medicine)	4	2020
Critical appraisal of a topic (CAT) (BSc Medicine, year 3)	16	2021
Academic and scientific training year 1 (AWVI) (BSc Medicine)	28	2021
Practical research tools (POV) (MSc Medicine)	4	2021

Student supervision	Hours	Year
Daily supervisor to Sophie van Rees Vellinga, honours student, BSc Medicine	22	2019
Daily supervisor to Isabelle Jetten, research student, BSc Psychology/ Biology	52	2021

Congress presentations	Hours	Year
European Renal Association (ERA-EDTA) Congress – poster presentation	8	2018
Social Scientific Research Nephrology (SWON) Symposium – oral presentation (invited)	12	2018
Working group Epidemiological Research the Netherlands (WEON) Conference – oral presentation	12	2018
Science day of Dutch Kidney Patients Association (NVN), Dutch Kidney Foundation (Nierstichting) and Dutch Federation of Nephrologists (NFN) – oral presentation (invited)	12	2018
Association for Researchers in Psychology and Health (ARPH) Conference – oral presentation	12	2019
Symposium 'Together Strong' of Dutch Kidney Patients Association (NVN) and Dutch Kidney Foundation (Nierstichting) – information booth (invited)	8	2019
Dutch Nephrology Days (NND) – oral presentation (invited)	8	2019
Dialysis Initiatives Nephrologists (DIN) – oral presentation (invited)	8	2019
HAL session of Dutch Kidney Patients Association (NVN) and Dutch Kidney Foundation (Nierstichting) – oral presentation (invited)	8	2019
European Renal Association (ERA-EDTA) Congress – poster presentation	8	2020
Social Scientific Research Nephrology (SWON) Symposium – oral presentation (invited)	12	2020
European Health Psychology Society (EHPS) Congress – oral presentation	12	2021
Science day of Dutch Kidney Patients Association (NVN), Dutch Kidney Foundation (Nierstichting) and Dutch Federation of Nephrologists (NFN) – oral presentation (invited)	12	2021
European Renal Association (ERA) Congress – oral presentation (invited)	12	2022

Other activities	Hours	Year
Researcher BCG-PRIME study	147	2020
Chair weekly nephrology journal club	160	2020/ 2021

List of Publications

Articles included in this thesis

1. **van der Willik EM**, Milders J, Bart JAJ, Bos WJW, van Ittersum FJ, ten Dam MAGJ, Hemmelder MH, Dekker FW, Meuleman Y. Discussing results of patient-reported outcome measures (PROMs) between patients and healthcare professionals in routine dialysis care: a qualitative study. *BMJ Open*. 2022;12:e067044.

2. **van der Willik EM**, van Breda GF, van Jaarsveld BC, van de Putte M, Jetten IW, Dekker FW, Meuleman Y, van Ittersum FJ, Terwee CB. Validity and reliability of Patient-Reported Outcomes Measurement Information System (PROMIS®) using Computerized Adaptive Testing (CAT) in patients with advanced chronic kidney disease. *Nephrology Dialysis Transplantation*. 2022;gfac231.

3. **van der Willik EM**, Lengton R, Hemmelder MH, Hoogeveen EK, Bart HAJ, van Ittersum FJ, ten Dam MAGJ, Bos WJW, Dekker FW, Meuleman Y. Itching in dialysis patients: impact on health-related quality of life and interactions with sleep problems and psychological symptoms – results from the RENINE/PROMs registry. *Nephrology Dialysis Transplantation*. 2022;37(9):1731-1741.

4. **van der Willik EM**, Terwee CB, Bos WJW, Hemmelder MH, Jager KJ, Zoccali C, Dekker FW, Meuleman Y. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. *Nephrology (Carlton)*. 2021;26(5):391-399.

5. **van der Willik EM**, van Zwet EW, Hoekstra T, van Ittersum FJ, Hemmelder MH, Zoccali C, Jager KJ, Dekker FW, Meuleman Y. Funnel plots of patient-reported outcomes to evaluate healthcare quality: Basic principles, pitfalls and considerations. *Nephrology (Carlton)*. 2021;26(2):95-104.

6. **van der Willik EM**, Hemmelder MH, Bart HAJ, van Ittersum FJ, Hoogendijk-van den Akker JM, Bos WJW, Dekker FW, Meuleman Y. Routinely measuring symptom burden and health-related quality of life in dialysis patients: first results from the Dutch registry of patient-reported outcome measures. *Clinical Kidney Journal*. 2020;14(6):1535-1544.

7. **van der Willik EM**, Meuleman Y, Prantl K, van Rijn G, Bos WJW, van Ittersum FJ, Bart HAJ, Hemmelder MH, Dekker FW. Patient-reported outcome measures: selection of a valid questionnaire for routine symptom assessment in patients with advanced chronic kidney disease - a four-phase mixed methods study. *BMC Nephrology*. 2019;20(1):344.

Other publications in nephrology care

8. Terwee CB, **van der Willik EM**, van Breda GF, van Jaarsveld BC, van de Putte M, Jetten IW, Dekker FW, Meuleman Y, van Ittersum FJ. Responsiveness and Minimal Important Change of seven PROMIS computerized adaptive tests (CAT) in patients with advanced chronic kidney disease. *Journal of Patient-Reported Outcomes*. 2023;7(1):35.

9. Lengton R, **van der Willik EM**, Meuleman Y, Hemmelder MH, Dekker FW, Hoogeveen EK; for the Netherlands Cooperative Study on the Adequacy of Dialysis-2 (NECOSAD) Study Group. Effect of Residual Kidney Function and Dialysis Adequacy on Chronic Pruritus in Dialysis Patients. *Nephrology Dialysis Transplantation*. 2022;gfac341.

10. Meuleman Y, **van der Willik EM**. Patiënt-gerapporteerde uitkomsten in de Neder-

landse nefrologie. *Diëtisten Nierziekten Nederland, nieuwsbrief*. Oktober 2022.

11. de Jong Y, **van der Willik EM**, Milders J, Meuleman Y, Morton RL, Dekker FW, van Diepen M. Person centred care provision and care planning in chronic kidney disease: which outcomes matter? A systematic review and thematic synthesis of qualitative studies. *BMC Nephrology*. 2021;22(1):309.

12. de Jong Y, **van der Willik EM**, Milders J, Voorend CGN, Morton RL, Dekker FW, Meuleman Y, van Diepen M. A meta-review demonstrates improved reporting quality of qualitative reviews following the publication of COREQ- and ENTREQ-checklists, regardless of modest uptake. *BMC Medical Research Methodology*. 2021;21(1):184.

13. **van der Willik EM**, Meuleman Y, Dekker FW. PROMs voor inzicht in kwaliteit van leven en gezondheidsklachten bij dialysepatiënten. *Dialyse & Nefrologie Magazine*. December 2020.

14. Fu EL, Janse RJ, de Jong Y, van der Endt VHW, Milders J, **van der Willik EM**, de Rooij ENM, Dekkers OM, Rotmans JI, van Diepen M. Acute kidney injury and kidney replacement therapy in COVID-19: a systematic review and meta-analysis. *Clinical Kidney Journal*. 2020;13(4):550- 563.

15. **van der Willik EM**. Zicht op jeuk; onderzoek vanuit LUMC. *Diavariatie*. September 2019.

Publications in other medical fields

16. Groeneveld L, Terwee CB, **van der Willik EM,** van Ittersum FJ, Langendoen-Gort M, Pals F, Blom M, Beulens JW, Elders PJM, Rutters F. Psychometric properties and implementation of seven PROMIS computerized adaptive tests in people with type 2 diabetes. *Submitted*.

17. Lijftogt N, Vahl AC, Karthaus EG, **van der Willik EM**, Amodio S, van Zwet EW, Hamming JF; Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit, and the Dutch Institute for Clinical Auditing. Effects of hospital preference for endovascular repair on postoperative mortality after elective abdominal aortic aneurysm repair: analysis of the Dutch Surgical Aneurysm Audit. *BJS Open*. 2021;5(3):zraa065.

18. Karthaus EG, Lijftogt N, Vahl A, **van der Willik EM**, Amodio S, van Zwet EW, Hamming JF; Dutch Society for Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing. Patients with a Ruptured Abdominal Aortic Aneurysm Are Better Informed in Hospitals with an "EVAR-preferred" Strategy: An Instrumental Variable Analysis of the Dutch Surgical Aneurysm Audit. *Annals of Vascular Surgery*. 2020;69:332-344.

19. Lijftogt N, Vahl A, **van der Willik EM**, Leijdekkers VJ, Wouters MWJM, Hamming JF; Dutch Society of Vascular Surgery, the Steering Committee of the Dutch Surgical Aneurysm Audit and the Dutch Institute for Clinical Auditing. Toward Optimizing Risk Adjustment in the Dutch Surgical Aneurysm Audit. *Annals of Vascular Surgery*. 2019;60:103-111.

20. Beck N, Hoeijmakers F, **van der Willik EM**, Heineman DJ, Braun J, Tollenaar RAEM, Schreurs WH, Wouters MWJM. National Comparison of Hospital Performances in Lung Cancer Surgery: The Role of Case Mix Adjustment. *Annals of Thoracic Surgery*. 2018;106(2):412-420.

21. van der Werf LR, Dikken JL, **van der Willik EM**, van Berge Henegouwen MI, Nieuwenhuijzen GAP, Wijnhoven BPL; Dutch Upper Gastrointestinal Cancer Audit (DUCA) group. Time interval between neoadjuvant chemoradiotherapy and surgery for oesophageal or junctional cancer: A nationwide study. *European Journal of Cancer.* 2018;91:76-85.

22. Lijftogt N, Karthaus EG, Vahl A, van Zwet EW, **van der Willik EM**, Tollenaar RAEM, Hamming JF, Wouters MWJM; Dutch Society of Vascular Surgery; Steering Committee of the Dutch Surgical Aneurysm Audit; Dutch Institute for Clinical Auditing. Failure to Rescue - a Closer Look at Mortality Rates Has No Added Value for Hospital Comparisons but Is Useful for Team Quality Assessment in Abdominal Aortic Aneurysm Surgery in The Netherlands. *European Journal of Vascular and Endovascular Surgery.* 2018;56(5):652-661.

23. de Neree Tot Babberich MP, **van der Willik EM**, van Groningen JT, Ledeboer M, Wiggers T, Wouters MW. Darmkankerchirurgie sinds het bevolkingsonderzoek. *Nederlands Tijdschrift voor Geneeskunde.* 2017;161:D997.

24. **van der Willik EM**, Vrijkotte TG, Altenburg TM, Gademan MG, Kist-van Holthe J. Exclusively breastfed overweight infants are at the same risk of childhood overweight as formula fed overweight infants. *Archives of Disease in Childhood.* 2015;100(10):932-7.