INVITED COMMENTARY

Wise decisions on renal replacement therapy require knowledgeable patients and good self-report knowledge scales*

Fabienne Dobbels and Nathalie Duerinckx

Health Services and Nursing Research, Department of Public Health and Primary Care, University of Leuven, Leuven, Belgium

Correspondence
Fabienne Dobbels MSC, PhD, Health Services
and Nursing Research, Department of Public
Health and Primary Care, University of Leuven,
Kapucijnenvoer 35/4, 3000 Leuven, Belgium.
Tel.: +32 16 33 69 81;
fax: +32 16 33 69 70;
e-mail: fabienne.dobbels@med.kuleuven.be

Conflict of interests
The authors have declared no conflict of
interests.

Patients with advanced chronic kidney disease (CKD) face the difficult task to decide which form of renal replacement therapy (RRT) they would like to pursue. For clinicians, it is evident that home dialysis has economic and quality of life advantages relative to hospital hemodialysis and that living donor kidney transplantation not only avoids or shortens the pretransplant waiting time on dialysis, but also results in better post-transplant outcomes compared with deceased donor transplant. Shared decision-making, however, will only be possible if patients are fully informed of existing modalities. Unfortunately, multiple studies consistently show that the level of knowledge on available RRT options in renal patients is suboptimal [1–5]. This is concerning, as inadequate knowledge, alongside other patient-, clinician- or healthcare system-related factors which fall beyond the scope of this commentary [6,7],

might to a reasonable extent explain why some efficacious treatment modalities are still underutilized. A standardized, validated self-report questionnaire to quantify renal patients' knowledge on RRT might therefore represent a major asset for clinicians.

Ismail and colleagues present the carefully designed and psychometrically sound Rotterdam Renal Replacement Knowledge-Test (R3K-T) [8], of which its value surpasses the few existing self-report knowledge questionnaires [1–5] in many ways: (i) It addresses the whole spectrum of RRT simultaneously instead of focusing exclusively on selected modalities; (ii) it concentrates on knowledge needed for shared treatment decision-making rather than on information needed once patients have chosen a RRT modality, such as self-care knowledge; (iii) items allow for one correct answer only, while some scales use ordinal scaling ranging

from, for example, totally disagree to totally agree, creating confusion on which response is correct; and (iv) the authors adopted item response theory as "modern" psychometric methodology instead of classical test theory applied in existing scales almost exclusively.

In line with a recently published report [9], the R3K-T is exemplary of a patient-reported outcome measure that undoubtedly will redesign kidney health services. The benefits of incorporating this scale into clinical care cannot be ignored: Firstly, clinicians often assume that education is perfectly encoded in the patient's brain once provided. Yet, older age, poor health literacy, and the cognitive impairment often observed in renal patients might explain why many studies show that some information is not fully understood or registered [1–5]. Given these many influencing factors, it is very difficult, if not impossible for clinicians to grasp the patient's knowledge during short and sporadic clinical encounters. The R3K-T creates a window of opportunity to assess the effectiveness of education provided in an easy, but standardized way. Secondly, even if clinicians use checklists, without proper validation, it remains unclear if items adequate capture the patient's level of understanding. Moreover, clinicians need to complete this checklist for every patient, a task that understandably most likely will not have the highest priority in the light of competing medical responsibilities. The R3K-T only needs to be completed once by the patient to provide clinicians with an accurate knowledge estimate, reducing observer bias caused by subjective clinician's judgments. Thirdly, evidence shows that education for renal patients differs largely in timing, content, and intensity. Ideally, all patients should receive the same comprehensive educational package, necessitating financial and staffing resources. Most nephrologists, however, do not have access to multidisciplinary support for patients with advanced CKD not yet on RRT. By looking at the item scores, the R3K-T could allow for a more custom-made and perhaps less resource-intense approach as only patients with suboptimal scores need to be educated and on topics related to wrong responses only. Finally, the R3K-T can easily be incorporated in clinical practice improvement or research projects to objectify the effect of educational interventions. Many initiatives are currently ongoing to promote, for example living donor kidney transplantation. Although insufficient knowledge is not the only driving factor in patients' preferences for deceased vs. living donor kidney transplantation, it is certainly a critical and modifiable one, as evidence shows that misconceptions are common [5]. Although the availability of multiple translations might facilitate a smooth transition from bench to bedside, the authors should work on parallel versions of the R3K-T to avoid remembering the correct answer with repeat administration of the same questionnaire, without patients understanding why a given response is correct.

Admittedly, clinicians are often reluctant to use self-report questionnaires. Many believe that patient-reported outcomes are highly subjective and can never be reliably quantified by self-report. Unfortunately, many questionnaires with a questionable development process and validation process continue to make it into the literature, further strengthening this widespread misconception. To fully appreciate the work of Ismail and other test developers, one should understand, however, that "hard" psychometric science underpins good questionnaire development and validation. A brief introduction might prevent clinicians from feeling overwhelmed by test theory and their respective impact on instrument development, scoring, and interpretation [10–12].

Theoretically, an infinite pool of items exists targeting a concept of interest. In the R3K-Y, this is knowledge on dialysis, transplantation, and living donation. Given that it would be impossible to present all possible items to patients, the challenge is then to select those items that best approach the patients' true knowledge level, that is, with the lowest measurement error possible. Classical test theory provides a reliability coefficient reflecting an estimate of the level of precision with which a score on a questionnaire (e.g., the R3K-T) reflects a person's true, but unknown score (e.g., knowledge). For each question, the proportion of patients responding correctly can be calculated, with lower values suggesting a higher item difficulty and higher values suggesting that items are easier. Yet, the proportions obtained will largely depend on the abilities of the population under investigation, hence difficulty scores will be test dependent. Moreover, measurement bias might also occur when people with different background characteristics (e.g., gender, race, country of origin) *with the same level of knowledge* have a different probability of given a certain response, also referred to as differential item functioning (DIF). Ismail and colleagues, for instance excluded five items that showed DIF and would have provided distorted knowledge estimates for the different research groups. Please note that DIF analysis is not the same as known group differences, as DIF analyses examine item response across groups at *equivalent* levels of knowledge, while the later expects different underlying knowledge levels in, for example patients with lower and higher education. While classical test theory does not allow to disentangle item difficulty from a person's true ability, IRT applies sophisticated nonlinear mathematical models, generating an estimate of the probability of a correct response on a question as a function of the characteristics of the item (such as difficulty, ability to discriminate between low, and high knowledge) and the true ability of patients. This results in better calibrated items and hence a more accurate estimate of a

patients' knowledge than just summing up individual item scores.

To summarize, clinicians should be encouraged to learn more about the principles behind classical and item response theory. Reading easy-to-follow brief publications [10–12] will help to appreciate the value of self-report questionnaires and to critically separate chaff from wheat. Just like physicians would never use unreliable diagnostic tests, so should every professional be able to critically select good self-report questionnaires. Recent reports forecast that self-report questionnaires not only will be increasingly embedded in clinical care, but will also be incorporated in public accountability evaluations of health services and professionals [9]. Ismail and coworkers provided compelling evidence that the R3K-T is psychometrically sound and hence can trustworthily be incorporated in research and clinical care.

Funding

References

1. Finkelstein FO, Story K, Firanek C, *et al.* Perceived knowledge among patients cared for by nephrologists about kidney disease and end-stage disease therapies. *Kidney Int* 2008; **74**: 1178.
2. Devins M, Binik YM, Mandin H, *et al.* The kidney disease questionnaire: a test for measuring patient knowledge about end-stage renal disease. *J Clin Epidemiol* 1990; **43**: 297.
3. Wright JA, Wallston KA, Easy TA, *et al.* Development and results of a kidney disease knowledge survey given to patients with CKD. *Am J Kidney Dis* 2011; **57**: 387.
4. Urstad KH, Andersen MH, Oyen O, *et al.* Patients' level of knowledge measured five days after kidney transplantation. *Clin Transplant* 2011; **25**: 646.
5. Stothers L, Gourlay WA, Liu L. Attitudes and predictive factors for live kidney donation: a comparison of live kidney donors versus nondonors. *Kidney Int* 2005; **67**: 1105.
6. Morton RL, Tong A, Howard K, *et al.* The views of patients and carers in treatment decision making for chronic disease: a systematic review and thematic synthesis of qualitative studies. *BMJ* 2010; **340**: c112.
7. Harwood L, Clark AM. Understanding pre-dialysis modality decision-making. A meta-synthesis of qualitative studies. *Int J Nurs Stud* 2013; **50**: 109.
8. Ismail S, Timmerman L, Timman R, *et al.* A psychometric analysis of the renal replacement knowledge-test using item response theory. *Transplant Int* 2013; **26**: 1164.
9. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013; **346**: f167.
10. Thomas ML. The value of item response theory in clinical assessment: a review. *Assessment* 2011; **18**: 291.
11. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21th century. *Med Care* 2000; **38**(Suppl. 9): II28.
12. Teresi JA, Ramirez M, Jones RN, *et al.* Modifying measures based on differential item functioning (DIF) impact analysis. *J Aging Health* 2012; **24**: 1044.