AJKD

Original Investigation

Developing a Set of Core Outcomes for Trials in Hemodialysis: An International Delphi Survey

Nicole Evangelidis, BSocSc,[1,2] Allison Tong, PhD,[1,2] Braden Manns, MD, MSc,[3]
Brenda Hemmelgarn, MD, PhD,[3] David C. Wheeler, MD,[4] Peter Tugwell, MD,[5]
Sally Crowe, PGDip,[6] Tess Harris,[7] Wim Van Biesen, MD, PhD,[8]
Wolfgang C. Winkelmayer, MD, ScD,[9] Benedicte Sautenet, MD, PhD,[1,2,10,11]
Donal O'Donoghue, MB,[12] Helen Tam-Tham, MSc,[3] Sajeda Youssouf, MD,[12]
Sreedhar Mandayam, MD, PhD,[9] Angela Ju, BSc (Hons),[1,2]
Carmel Hawley, MBBS,[13,14,15] Carol Pollock, PhD,[16] David C. Harris, MD, MBBS,[17]
David W. Johnson, PhD,[13,14,15] Dena E. Rifkin, MD,[18,19] Francesca Tentori, MD, MS,[20]
John Agar, MBBS,[21] Kevan R. Polkinghorne, PhD,[22,23] Martin Gallagher, PhD,[24]
Peter G. Kerr, PhD,[22] Stephen P. McDonald, PhD,[25,26] Kirsten Howard, PhD,[1]
Martin Howell, PhD,[1,2] and Jonathan C. Craig, PhD,[1,2] on behalf of the Standardized
Outcomes in Nephrology–Hemodialysis (SONG-HD) Initiative*

Background: Survival and quality of life for patients on hemodialysis therapy remain poor despite substantial research efforts. Existing trials often report surrogate outcomes that may not be relevant to patients and clinicians. The aim of this project was to generate a consensus-based prioritized list of core outcomes for trials in hemodialysis.

Study Design: In a Delphi survey, participants rated the importance of outcomes using a 9-point Likert scale in round 1 and then re-rated outcomes in rounds 2 and 3 after reviewing other respondents' scores. For each outcome, the median, mean, and proportion rating as 7 to 9 (critically important) were calculated.

Setting & Participants: 1,181 participants (202 [17%] patients/caregivers, 979 health professionals) from 73 countries completed round 1, with 838 (71%) completing round 3.

Outcomes & Measurements: Outcomes included in the potential core outcome set met the following criteria for both patients/caregivers and health professionals: median score ≥ 8, mean score ≥ 7.5, proportion rating the outcome as critically important $\geq 75\%$, and median score in the forced ranking question < 10.

Results: Patients/caregivers rated 4 outcomes higher than health professionals: ability to travel, dialysis-free time, dialysis adequacy, and washed out after dialysis (mean differences of 0.9, 0.5, 0.3, and 0.2, respectively). Health professionals gave a higher rating for mortality, hospitalization, decrease in blood pressure, vascular access complications, depression, cardiovascular disease, target weight, infection, and potassium (mean differences of 1.0, 1.0, 1.0, 0.9, 0.9, 0.8, 0.7, 0.4, and 0.4, respectively).

From the [1]Sydney School of Public Health, University of Sydney; [2]Centre for Kidney Research, The Children's Hospital at Westmead, Sydney, NSW, Australia; [3]Departments of Medicine and Community Health Sciences; Libin Cardiovascular Institute and O'Brien Institute of Public Health, University of Calgary, Calgary, Canada; [4]Centre for Nephrology, University College London, London, United Kingdom; [5]Department of Medicine, University of Ottawa, Ottawa, Canada; [6]Crowe Associates, Oxon, United Kingdom; [7]PKD International, Geneva, Switzerland; [8]Renal Division, Ghent University Hospital, Ghent, Belgium; [9]Selzman Institute for Kidney Health, Section of Nephrology, Baylor College of Medicine, Houston, TX; [10]INSERM, U1153, Paris; [11]Department of Nephrology and Clinical Immunology, University Hospital of Tours, Tours, France; [12]Department of Renal Medicine, Salford Royal NHS Foundation Trust, Salford, United Kingdom; [13]Queensland School of Medicine, University of Queensland at Princess Alexandra Hospital; [14]Translational Research Institute; [15]Metro South and Ipswich Nephrology and Transplant Services (MINTS), Princess Alexandra Hospital, Brisbane; [16]Renal Division, Kolling Institute, Sydney; [17]Centre for Transplantation and Renal Research, The Westmead Institute for Medical Research, University of Sydney, Westmead, New South Wales, Australia; Divisions of [18]Nephrology and [19]Preventive Medicine, University of California, San Diego, CA; [20]Arbor Research Collaborative for Health, Ann Arbor, MI; [21]University Hospital Geelong, Geelong; [22]Monash Medical Centre and Monash University; [23]Department of Epidemiology & Preventative Medicine, Monash University, Clayton; [24]Concord Clinical School, The University of Sydney, Sydney; [25]Central Northern Adelaide Renal and Transplantation Service, Royal Adelaide Hospital; and [26]Faculty of Health Science, University of Adelaide, Adelaide, Australia.

*A list of the SONG-HD initiative investigators is available at www.songinitiative.org.

Because an author of this article was an editor for AJKD at the time the manuscript was submitted, the peer-review and decision-making processes were handled entirely by an Associate Editor (Amit X. Garg, MD, MA, FRCPC, PhD) who served as Acting Editor-in-Chief. Details of the journal's procedures for potential editor conflicts are given in the Information for Authors & Journal Policies.

Address correspondence to Nicole Evangelidis, BSocSc, Centre for Kidney Research, The Children's Hospital at Westmead, Westmead, NSW 2145, Sydney, Australia. E-mail: nicole.evangelidis@sydney.edu.au

Limitations: The Delphi survey was conducted online in English and excludes participants without access to a computer and internet connection.

Conclusions: Patients/caregivers gave higher priority to lifestyle-related outcomes than health professionals. The prioritized outcomes for both groups were vascular access problems, dialysis adequacy, fatigue, cardiovascular disease, and mortality. This process will inform a core outcome set that in turn will improve the relevance, efficiency, and comparability of trial evidence to facilitate treatment decisions.

Editorial, p. 453

The enormous investment in biomedical research, particularly in randomized trials, may not have led to the improvements in health that were hoped for.[1-3] It has been estimated that 85% of the worldwide US $240 billion invested in research annually is wasted.[3] In nephrology, there has been substantial research investment into hemodialysis (HD), yet survival rates have not improved correspondingly over the past 40 years and quality of life remains poor, even compared with patients with many cancers.[4-7] This may be partly attributable to what outcomes are selected and reported in trials, a challenge well recognized across medical specialties.[1,8-10]

Surrogate end points are frequently used in clinical trials because of feasibility, in preference to outcomes that are directly relevant to patients and clinicians.[9,11,12] In HD, biochemical markers such as serum phosphorus, calcium, and parathyroid hormone levels, are commonly reported but are not strongly and consistently associated with mortality, cardiovascular disease (CVD), or quality of life.[13-17] Patients on HD therapy prioritize outcomes relevant to their well-being and lifestyle—fatigue, ability to travel, ability to work, sleep, and anxiety/stress[18,19]— all of which are largely absent as outcomes reported by HD trials. In addition, the large heterogeneity of outcome measures and potential for outcome reporting bias (in which trials selectively report results for outcomes that favor the intervention) undermine the reliability of trial evidence to inform clinicians and patients about the relative effects of interventions.[20]

Engaging all stakeholders in establishing a core outcome set, an agreed minimum set of standardized outcomes to be measured and reported in all trials for a particular clinical area,[21,22] can increase the relevance, efficiency, and reliability of trials. Initiatives to develop core outcomes are seen in rheumatology and oncology and have demonstrated improvements in consistent reporting of relevant outcomes.[10,23,24] As part of the international Standardized Outcomes in Nephrology−Hemodialysis (SONG-HD) initiative, this study aimed to generate a consensus-based prioritized list of outcome domains for people on HD therapy, which will be used to establish a core outcome set that reflects the shared priorities of patients, caregivers, and health professionals.

METHODS

Study Design

The Delphi method is a technique for achieving consensus among a panel of experts. This process involves sequential surveys, typically conducted over 3 rounds and answered anonymously, and gives equal influence to all who participate. It was first developed by the RAND Corporation in the 1950s[25] and has since been increasingly used as a valid approach to develop consensus-based core outcomes for clinical trials in various medical specialty areas.[10,22,26-28] The SONG-HD Delphi process is shown in Figure S1 (provided as online supplementary material).

Participant Selection and Recruitment

Stakeholders including patients, caregivers/family members, nephrologists, surgeons, nurses, social workers, psychologists, dieticians, pharmacists, policy makers, researchers, and industry with experience or interest in HD were invited to join the Delphi panel. Participants worldwide were eligible if they were older than 18 years and able to complete an online survey in English. All participants provided informed consent.

Using an opt-in snowballing sampling frame, we recruited patients/ caregivers through participating hospitals, patient/consumer organizations, and social media listed in Item S1. Health professionals were recruited via the investigators' networks and via e-mails and newsletters circulated by professional societies (Item S1). Participants registered their e-mail addresses on www.songinitiative.org prior to the survey launch. Ethics boards of the University of Sydney (2015/228), Baylor College of Medicine (H-37406), University of Calgary (REB15-0708), Monash Medical Centre (13082B), Salford Royal NHS (15/WM/0303), and Sydney West Area Health Service (HREC2009/6/4.15) approved this study.

Data Collection

Overview

The 34 outcome domains for the 3-round Delphi survey were identified from a systematic review of outcomes reported in trials in HD therapy, stakeholder interviews, and nominal group technique conducted with patients on HD therapy and caregivers.[29,30] The ordering of outcomes was randomized and included a plain language definition (Item S2). The survey was reviewed by the SONG Executive Committee and SONG-HD investigators and piloted among 10 patients. The Delphi survey was completed online via LimeSurvey during September to November 2015. The online survey administration minimizes data entry error, allows for wider dissemination, and is more efficient compared to a paper survey.

Round 1

Participants rated the importance of each of the 34 outcomes based on a 9-point Likert scale. A score of 7 to 9 indicated that the outcome was of "critical importance," 4 to 6 indicated "important but not critical," and 1 to 3 indicated "limited importance" according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) process.[31] An option of "unsure" was provided. Participants could enter comments about their choice of ranking for each outcome. In addition, participants could suggest new outcomes that were not included in the survey. Outcomes with a mean and median value < 7 for patients/caregivers and health professionals were not included in round 2.

Round 2

Participants reviewed the group scores and their own score for each outcome and re-rated the 29 remaining outcomes using the same 9-point Likert scale. Group scores were displayed in an interactive column graph that showed the distribution of scores for patients/caregivers, health professionals, and the total sample combined ([weighted] Fig S1). Instructions on how to read the graph were provided to ensure that participants were able to understand the results. For each outcome, an optional comments box allowed participants to explain reasons for their ratings. Outcomes with a mean and median value ≤ 7 for patients/caregivers and health professionals were excluded from round 3.

Round 3

Participants were asked to re-rate the 20 remaining outcomes using the same Likert scale in the previous rounds after viewing the scores and, in addition, de-identified comments (ie, free text responses from participants relating to reasons for their rankings or observations on the results for each outcome) from round 2. Comments were divided into 2 boxes: Patients and caregivers and Health professionals, with the ability to scroll down and read all comments. A free text box was provided for each outcome so participants could provide additional comments. In addition, participants completed a forced ranking question, using a drag-and-drop function, to rank outcomes relative to each other.

Data Analysis

We used SPSS (IBM; version 22.0) to calculate descriptive statistics. We calculated the median, mean, and proportion of participants (rating, 7-9) for each outcome. Scores were calculated separately for patients/caregivers and health professionals, with the difference in mean values considered significant at $P < 0.05$ based on the t test. For ranking scores, we calculated median and interquartile range (IQR) for each outcome to determine rank. Any analysis of the total sample was weighted equally between patients/caregivers and health professionals.

Definition of Consensus

Consensus was defined a priori based on the Outcome Measures in Rheumatology (OMERACT) definition using proportion scores. "Consensus in" is defined as ≥70% of participants scoring as 7 to 9 and <15% of participants scoring as 1 to 3.[29] However, because most participants rated all outcomes in round 3 as critically important with scores of 7 to 9, these criteria resulted in a list of 16 outcomes, which exceeded the recommended 3 to 5 outcomes for a core outcome set. Therefore, the definition and threshold for "consensus in" were revised to determine a maximum of 5 core outcomes to be considered for the core outcome set.

Outcomes from round 1 with a mean and median score ≥ 7 for patients/caregivers and health professionals were included in round 2. This was validated against the proportion of critically important scores (rated 7-9) for each outcome to ensure that important outcomes were not excluded.

Outcomes from round 2 with a mean and median score > 7 for patients/caregivers and health professionals were included in round 3. This was validated against the proportion of critically important scores (rated 7-9) for each outcome.

Outcomes included in the potential core outcome set met the following criteria for both patients/caregivers and health professionals: median score ≥ 8; mean score ≥ 7.5; proportion of participants rating the outcome "critically important" ≥ 75%, and median score < 10 in the forced ranking question.

RESULTS

Participant Characteristics

In total, 1,181 people from 73 countries participated in round 1 of the Delphi survey, including 202 (17%) patients/caregivers and 979 (83%) health professionals. Round 2 included 165 (17%) patients/caregivers and 784 (83%) health professionals from 63 countries. In the third and final round, 150 (18%) patients/caregivers and 688 (82%) health professionals participated. The full survey completion rate was 71%. Participant characteristics are provided in Tables 1 and 2.

In round 3, of the 150 patients/caregivers, 115 (77%) were aged 41 to 70 years and 76 (51%) were women. Patients/caregivers were from 14 and 11 countries in rounds 1 and 3, respectively. In round 3, the majority of patient/caregiver participants were from Australia (40 [27%]), Canada (37 [25%]), United Kingdom (25 [17%]), United States (19 [13%]), and New Zealand (18 [12%]). Among the 116 (77%) patients on HD therapy, 63 (42%) were on in-center HD therapy and 51 (34%) were on home HD therapy. Health professionals included 857 (51%) nephrologists, 386 (38%) nurses, 53 (5%) researchers, and 63 (6%) in other roles. Health professionals were from 72 and 62 countries in rounds 1 and 3, respectively.

Delphi Scores

Round 1

Values for mean, median, and proportion of participants rating the outcome as 7 to 9 (critical importance) for each of the 34 outcomes are shown in Table S1. The top 3 outcomes rated by patients/caregivers based on mean scores (1-9) were dialysis adequacy (7.5 ± 2.1 [standard deviation]), ability to travel (7.5 ± 1.9), and dialysis-free time (7.3 ± 1.8). The top 3 outcomes for health professionals were vascular access problems (8.1 ± 1.3), CVD (7.9 ± 1.3), and mortality (7.7 ± 1.6). Definitions of high rating outcomes are provided in Box 1.

The following outcomes were excluded from round 2 because they had a mean or median score < 7 (not of critical importance) among both patient/caregiver and health professional groups in round 1: nausea/vomiting, sexual function, restless legs syndrome, itching, and cramps. Less than 25% of participants

Table 1. Characteristics of Patients/Caregivers

Characteristic	Round 1 (n = 202)	Round 2 (n = 165)	Round 3 (n = 150)
Participant type			
Patient	168 (83.2)	138 (83.6)	127 (84.7)
Caregiver/family member	34 (16.8)	27 (16.4)	23 (15.3)
Sex			
Male	96 (47.5)	78 (47.3)	74 (49.3)
Female	106 (52.5)	87 (52.7)	76 (50.7)
Age group			
18-40 y	33 (16.4)	23 (13.9)	16 (10.7)
41-50 y	41 (20.3)	34 (20.6)	29 (19.3)
51-60	59 (29.2)	49 (29.7)	47 (31.3)
61-70 y	48 (23.8)	40 (24.2)	39 (26.0)
≥71 y	21 (10.4)	19 (11.5)	19 (12.7)
Marital status[a]			
Single	24 (11.9)	20 (12.1)	17 (11.3)
Partner/de facto[b]	14 (7.0)	11 (6.7)	9 (6.5)
Married	110 (54.5)	87 (52.7)	78 (52.0)
Divorced/separated/ widowed	41 (20.3)	34 (20.6)	34 (24.6)
No. of children[a]			
0	64 (31.7)	51 (30.9)	43 (28.7)
1-2	83 (41.1)	66 (40.0)	63 (42.0)
3-4	41 (20.3)	34 (20.6)	31 (20.7)
Employment status[a]			
Employed	74 (36.7)	55 (33.3)	46 (35.4)
Unemployed	37 (18.3)	28 (17.0)	25 (16.7)
Retired	67 (33.2)	60 (36.4)	58 (38.7)
Student	1 (0.5)	1 (0.6)	1 (0.8)
Education[a]			
Did not complete HS	28 (13.9)	21 (12.7)	18 (12.0)
HS graduate	29 (14.4)	22 (13.3)	22 (14.7)
Professional certificate	43 (21.3)	36 (21.8)	32 (21.3)
Undergraduate degree	59 (29.2)	47 (28.5)	43 (28.7)
Postgraduate degree	24 (11.9)	22 (13.3)	19 (12.7)
Current type of treatment[a]			
In-center HD	91 (45.0)	71 (43.0)	63 (42.0)
Home HD	70 (34.7)	57 (34.5)	51 (34.0)
Peritoneal dialysis	2 (1.0)	2 (1.2)	2 (1.3)
Transplantation	19 (9.4)	17 (10.3)	17 (11.3)
HD vintage[a]			
<1 y	26 (12.9)	18 (10.9)	17 (11.3)
1-5 y	95 (47.0)	76 (46.1)	69 (46.0)
6-10 y	27 (13.4)	25 (15.2)	22 (14.7)
11-15 y	20 (9.9)	17 (10.3)	14 (9.3)
>15 y	15 (7.4)	12 (7.3)	12 (8.0)
Country			
Canada	53 (26.2)	41 (24.8)	37 (24.7)
Australia	49 (24.3)	41 (24.8)	40 (26.7)
United Kingdom	35 (17.3)	29 (17.6)	25 (16.7)
United States	25 (12.4)	20 (12.1)	19 (12.7)
New Zealand	21 (10.4)	20 (12.1)	18 (12.0)
Other[c]	19 (9.5)	14 (8.4)	11 (7.3)

Note: Values are given as number (percentage).
Abbreviations: HD, hemodialysis; HS, high school.
[a]Percentages do not add up to 100 due to undisclosed responses (excluded).
[b]Couple living together.
[c]Other includes 9 countries: Romania, India, Spain, Czech Republic, Egypt, the Netherlands, Indonesia, Italy, and Philippines.

Table 2. Characteristics of Health Professionals

Characteristic	Round 1 (n = 979)	Round 2 (n = 784)	Round 3 (n = 688)
Participant type[a]			
Nephrologist	483 (46.9)	450 (57.4)	401 (58.3)
Nurse	386 (37.5)	277 (35.3)	233 (33.9)
Researcher	53 (5.2)	50 (6.4)	47 (6.8)
Nephrology trainee	44 (4.3)	36 (4.3)	28 (3.8)
Policy maker	17 (1.7)	16 (1.9)	16 (2.2)
Industry	13 (1.3)	13 (1.6)	12 (1.6)
Dietician	11 (1.1)	9 (1.1)	9 (1.2)
Social worker	7 (0.7)	7 (0.8)	7 (1.0)
Pharmacist	4 (0.4)	3 (0.4)	3 (0.4)
Psychologist	3 (0.3)	1 (0.1)	1 (0.1)
Surgeon	2 (0.2)	2 (0.2)	2 (0.3)
Other	6 (0.6)	3 (0.4)	3 (0.4)
Sex			
Male	447 (45.7)	362 (46.2)	318 (46.2)
Female	532 (54.3)	422 (53.8)	370 (53.8)
Age group			
18-40 y	435 (44.4)	317 (40.4)	268 (39.0)
41-50 y	262 (26.8)	220 (28.1)	197 (28.6)
51-60 y	207 (21.1)	181 (23.1)	164 (23.8)
61-70 y	63 (6.4)	58 (7.4)	52 (7.6)
≥71 y	12 (1.2)	8 (1.0)	7 (1.0)
Experience in HD			
≤10 y	406 (41.5)	297 (37.9)	252 (36.6)
11-20 y	308 (31.5)	254 (32.4)	228 (33.1)
≥21 y	265 (27.1)	117 (29.7)	63 (9.1)
No. of HD trials as investigator			
0	439 (44.8)	336 (42.9)	287 (41.7)
1-5	364 (37.2)	299 (38.1)	266 (38.7)
6-10	93 (9.5)	80 (10.2)	72 (10.5)
≥11	83 (8.5)	69 (8.8)	63 (9.2)
Other roles			
Government, policy	119 (10.9)	93 (10.6)	87 (11.2)
Guidelines	391 (35.8)	314 (35.6)	278 (35.6)
Funding	73 (6.7)	66 (7.5)	60 (7.7)
Other	509 (46.6)	408 (46.3)	355 (45.5)
Country			
Australia	133 (13.6)	120 (15.3)	108 (15.7)
Saudi Arabia	131 (13.4)	83 (10.6)	66 (9.6)
Spain	122 (12.5)	98 (12.5)	85 (12.4)
Romania	101 (10.3)	73 (9.3)	59 (8.6)
Canada	58 (5.9)	48 (6.1)	42 (6.1)
Portugal	56 (5.7)	48 (6.1)	44 (6.4)
United Kingdom	49 (5.0)	45 (5.7)	41 (6.0)
United States	35 (3.6)	30 (3.8)	29 (4.2)
New Zealand	30 (3.1)	26 (3.3)	23 (3.3)
Turkey	21 (2.1)	13 (1.7)	12 (1.7)
Other[b]	243 (24.7)	213 (25.5)	179 (26.0)

Note: Values are given as number (percentage).

Abbreviation: HD, hemodialysis.

[a]Some have multiple roles.

[b]Other includes 63 countries (in descending order of number of participants): Philippines, India, Belgium, Germany, Italy, Netherlands, Poland, Hungary, Argentina, Egypt, Greece, Sweden, Thailand, France, China, Croatia, Lithuania, Russian Federation, Singapore, Bangladesh, Brazil, Chile, Israel, Nigeria, Syria, Uruguay, Colombia, Czech Republic, Ireland (Republic), Japan, Kuwait, Malaysia, Mexico, Pakistan, Serbia, Slovakia, Slovenia, South Africa, Switzerland, Armenia, Austria, Belarus, Bolivia, Bosnia Herzegovina, Cameroon, Chad, Denmark, El Salvador, Indonesia, Iran, South Korea, Lebanon, Macedonia, Malawi, Morocco, Niger, Oman, Paraguay, Peru, Senegal, Venezuela, Vietnam, and Yemen.

Box 1. High-rating Outcome Domains and Definitions

Vascular access problems

Problems with fistula, graft, or catheter required for dialysis (eg, access infections, bleeding, bruising, pain, discomfort, clotting)

Death/mortality

No. of people on hemodialysis therapy who have died, risk for death, how long the patient will live

Cardiovascular disease

Disease of the heart and blood vessels (eg, heart attack, stroke, blockage of blood vessels)

Dialysis adequacy

How well the dialysis cleans the blood, clearance, Kt/V

Fatigue/energy

Feeling tired with no energy for weeks, for most of the time

Ability to travel

To go away for holiday/vacation, event, visiting family, work

Dialysis-free time

Time (hours/days) not doing dialysis

suggested new outcomes in round 1 (Table S2). These outcomes were not considered for inclusion in round 2 for the following reasons: the outcome could not be measured in a clinical trial for the majority of adult patients on HD therapy, the outcome was too broad conceptually or ambiguously defined, and the outcome was described as an intervention.

Round 2

Round 2 included 29 outcomes (Table S3). The top 3 outcomes for patients/caregivers were dialysis adequacy (7.7 ± 1.8), ability to travel (7.6 ± 1.9), and dialysis-free time (7.5 ± 1.7). The top 3 outcomes rated by health professionals were vascular access problems (8.4 ± 1.0), CVD (8.2 ± 1.1), and mortality (8.2 ± 1.2).

Outcomes that had a mean and median score ≤ 7 with <70% of the sample rating the outcome as 7 to 9 (critical importance) were excluded from round 3: anxiety/stress, food enjoyment, calcium, parathyroid hormone level, cognition, sleep, bone health, financial impact, and phosphate.

Round 3

Round 3 included 20 outcomes (Table S4). The top 3 outcomes for patients/caregivers were dialysis adequacy (7.9 ± 1.8), ability to travel (7.7 ± 1.7), and vascular access problems (7.7 ± 2.0). The top 3 outcomes for health professionals were vascular access problems (8.6 ± 0.9), CVD (8.4 ± 1.0), and mortality (8.3 ± 1.1).

Outcomes that met at least 2 of the following criteria for consensus within both stakeholder groups

(median ≥ 8, mean ≥ 7.5, proportion $\geq 75\%$, and median rank < 10; Table 3) were CVD, mortality, dialysis adequacy, fatigue, and vascular access problems. All participant comments for each outcome are provided in Item S3.

Changes in Scores From Rounds 1 to 3 Within Stakeholder Groups

As shown in Fig 1, patient/caregiver mean scores increased between rounds 1 and 3 for the following 7 outcomes: vascular access problems (mean score difference, 0.9; $P < 0.001$), CVD (mean difference, 0.7; $P = 0.002$), infection/immunity (mean difference, 0.7; $P = 0.004$), decrease in blood pressure (mean difference, 0.6; $P = 0.02$), mobility (mean difference, 0.6; $P = 0.02$), target weight (mean difference, 0.6; $P = 0.02$), and washed out after dialysis (mean difference, 0.5; $P = 0.01$).

For health professionals, mean scores increased for 16 outcomes between rounds 1 and 3 (Fig 2): mortality (mean difference, 0.6; $P < 0.001$), CVD (mean difference, 0.5; $P < 0.001$), vascular access problems (mean difference, 0.4; $P < 0.001$), ability to travel (mean difference, 0.4; $P < 0.001$), dialysis-free time (mean difference, 0.4; $P < 0.001$), dialysis adequacy (mean difference, 0.4; $P < 0.001$), washed out after dialysis (mean difference, 0.3; $P < 0.001$), ability to work (mean difference, 0.3; $P < 0.001$), infection/immunity (mean difference, 0.3; $P < 0.001$), decrease in blood pressure (mean difference, 0.3; $P < 0.001$), hospitalization (mean difference, 0.3; $P < 0.001$), fatigue (mean difference, 0.3; $P < 0.001$), impact on family/friends (mean difference, 0.3; $P < 0.001$), mobility (mean difference, 0.2; $P < 0.001$), pain (mean difference, 0.2; $P = 0.007$), and blood pressure (mean difference, 0.2; $P = 0.02$). No mean scores decreased across the 3 rounds of the survey for either patients/caregivers or health professionals.

Differences Between Stakeholder Groups

Differences in ratings between stakeholder groups are shown in Fig 3. Based on the difference in mean scores in round 3, four outcomes were rated higher by patients/caregivers: ability to travel (mean difference, 0.9; $P < 0.001$), dialysis-free time (mean difference, 0.5; $P < 0.001$), dialysis adequacy (mean difference, 0.3; $P = 0.05$), and washed out after dialysis (mean difference, 0.2; $P = 0.05$).

Health professionals rated 11 outcomes higher compared with patients/caregivers: mortality (mean difference, 1.0; $P < 0.001$), hospitalization (mean difference, 1.0; $P < 0.001$), decrease in blood pressure (mean difference, 1.0; $P < 0.001$), vascular access problems (mean difference, 0.9; $P < 0.001$), depression (mean difference, 0.9; $P < 0.001$), CVD (mean difference, 0.8; $P < 0.001$), target weight

Table 3. Inclusion Criteria for the Core Outcome Set Based on Median, Mean, Proportion, and Median Rank Scores

		Criterion: Median Score ≥ 8		Criterion: Mean Score ≥ 7.5		Criterion: ≥75% Give Rating of 7-9[a]		Criterion: Median Rank of 1-9[b]	
		Pts	HPs	Pts	HPs	Pts	HPs	Pts	HPs
Core outcome set: outcomes appear consistently across all measures for both groups	Vascular access problems	9.0	9.0	7.7	8.6	82%	97%	7.0	4.0
	Death/mortality	9.0	9.0	7.3	8.3	73%	94%	8.5	3.0
	Cardiovascular disease	8.0	9.0	7.6	8.4	77%	95%	7.0	3.0
	Dialysis adequacy	9.0	8.0	7.9	7.6	90%	81%	8.0	11.0
	Fatigue	8.0	8.0	7.6	7.6	82%	86%	8.0	11.0
Outcomes do not appear consistently across all measures for both groups	Ability to work	8.0	8.0	7.3	7.7	74%	85%	9.0	13.0
	Decrease in blood pressure	8.0	8.0	6.9	7.9	68%	89%	11.0	9.0
	Ability to travel	8.5	7.0	7.7	6.8	75%	57%	10.0	17.0
	Dialysis-free time	8.0	7.0	7.6	7.1	74%	67%	9.0	15.0
	Infection/immunity	7.0	8.0	7.2	7.6	74%	86%	11.0	9.0
	Anemia	8.0	7.0	7.4	7.3	77%	76%	11.0	10.0
	Mobility	8.0	7.0	7.3	7.2	76%	71%	11.0	14.0
	Blood pressure	8.0	7.0	7.4	7.3	76%	81%	10.0	9.0
	Washed out after dialysis	8.0	7.0	7.5	7.2	80%	78%	10.0	13.0
	Hospitalization	7.0	8.0	6.6	7.6	65%	86%	13.0	8.0
	Impact on family/friends	7.0	7.0	7.1	7.2	66%	72%	11.0	15.0
	Depression	7.0	7.0	6.4	7.2	61%	79%	13.0	12.0
	Target weight	7.0	7.0	6.6	7.2	58%	78%	13.0	11.0
	Potassium	7.0	7.0	6.7	7.1	63%	69%	13.0	11.0
	Pain	7.0	7.0	6.9	7.2	66%	75%	13.0	12.0

Note: Grey shading indicates that the outcome met the criteria. That is, outcome rates consistently high, according to inclusion criteria, for both patients/caregivers and health professionals. The outcome must appear consistently (at least twice) for each of the criteria (median ≥ 8, mean ≥ 7.5, proportion ≥ 75%, and median rank < 10).

Abbreviations: HPs, health care professionals; Pts, patients/caregivers.

[a]Score of 7 to 9 indicates critical importance.

[b]As given in the forced rank question.

(mean difference, 0.7; *P* < 0.001), infection/immunity (mean difference, 0.4; *P* = 0.002), potassium level (mean difference, 0.4; *P* = 0.02), ability to work (mean difference, 0.3; *P* = 0.008), and pain (mean difference, 0.3; *P* = 0.04).

Forced Ranking Scores

Results of the forced ranking question are shown in Table S5. The top outcomes ranked by patients/caregivers were CVD (median rank score, 7; IQR, 4-14), vascular access problems (median rank score, 7; IQR, 3-13), dialysis adequacy (median rank score, 8; IQR, 4-11), and fatigue (median rank score, 8; IQR, 4-11). The top outcomes ranked by health professionals were CVD (median rank score, 3; IQR, 2-7), mortality (median rank score, 3; IQR, 1-9), and vascular access problems (median rank score, 4; IQR, 2-7).
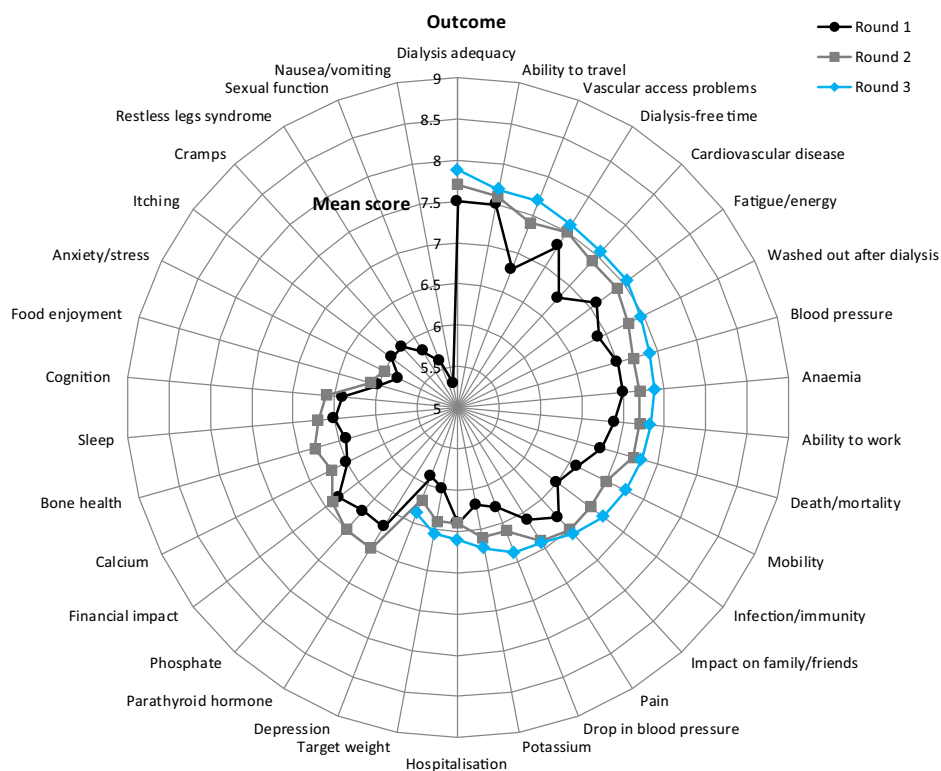
DISCUSSION

The highest priority outcomes shared among patients/caregivers and health professionals were vascular access problems, dialysis adequacy, fatigue, CVD, and mortality. Overall, most of these reflect common and high-impact outcomes in the context of HD, which have remained as major challenges in providing care for patients on HD therapy. Frequently

reported biochemical outcomes in HD trials, such as phosphate, calcium, and parathyroid hormone levels, were consistently rated to be of lower importance by both stakeholder groups.

Although there was convergence in ratings between patients/caregivers and health professionals across the 3 rounds of the Delphi, our findings also highlight some mismatches. Outcomes relating to lifestyle (ability to travel and dialysis-free time) and well-being (washed out after dialysis) rated higher among patients/caregivers compared with health professionals. Research has consistently shown that patients and caregivers prioritize lifestyle-related outcomes over biochemical end points.[18,19,30] Ability to travel, fatigue/energy, and dialysis-free time are important outcomes for patients and caregivers who want to retain normality in their day-to-day lives and maintain a reasonable quality of life.[19,30] In a recent discrete choice study, patients were willing to trade 23 months of life expectancy with home-based dialysis in exchange for increased ability to travel.[32] Biochemical end points are less important to patients and caregivers because they are considered intangible or imperceptible.[30]

Dialysis adequacy was also rated higher by patients/caregivers. Participant comments (Item S3) and

Figure 1. Mean scores of patients/caregivers in rounds 1 to 3. Ordered by round 3 scores. Round 1 (n = 202); round 2 (n = 165); round 3 (n = 150). Mean scores for rounds 2 and 3 are not available for outcomes that were excluded in those rounds.

discussions from the recent SONG-HD consensus workshop on establishing core outcomes in HD[33] indicate that participants conceptualized the term "dialysis adequacy" as a broad quality-of-life outcome (ie, dialysis that is adequate for enabling patients to feel well) rather than quantification of urea kinetics, which may explain this apparent divergence.

The rating of CVD by patients increased significantly between rounds 1 and 3. Based on comments provided by patients shown in Item S3, the reasons were largely due to gaining an understanding of the importance of CVD through reading other participant's comments, and a few patients had cardiovascular events during the survey period (Item S3). Similarly, studies have shown that patients with chronic kidney disease have limited knowledge about their risk for CVD.[30,34]

Outcomes rated higher by health professionals (with a mean difference > 0.5 in ratings) were mortality, hospitalization, decrease in blood pressure, vascular access problems, depression, CVD, and target weight. This perhaps reflects their perceived primary clinical role in preventing and managing these outcomes, which are common among HD patients, and awareness of their impact on patients' lives, whereas high-priority outcomes for patients

such as ability to travel and dialysis-free time may be seen by health professionals as impractical to measure.

Uremic symptoms such as itching and nausea/vomiting, as well as restless legs syndrome and cramps, were rated relatively lower by patients/caregivers in round 1 and were excluded from subsequent rounds. Instead, patients/caregivers gave higher priority to broader outcomes related to lifestyle impact and overall well-being (dialysis-free time, ability to travel, and fatigue) rather than specific symptoms. However, many patient-reported outcome measures are designed to assess symptoms and few robust and well-validated instruments measure specific patient-important outcomes that have an arguably more pervasive and long-term impact on lifestyle and well-being.[35-37] Furthermore, patients may not report lifestyle problems (eg, ability to work) that are not articulated in clinical settings because patients may perceive these to be beyond their clinician's control.

The SONG-HD Delphi survey had a large sample size (5 times greater than most similar Delphi surveys to establish core outcomes)[23,38,39] and achieved a broad engagement of a range of stakeholders, including patients/caregivers, multidisciplinary health care providers, policy makers, and industry, across 73
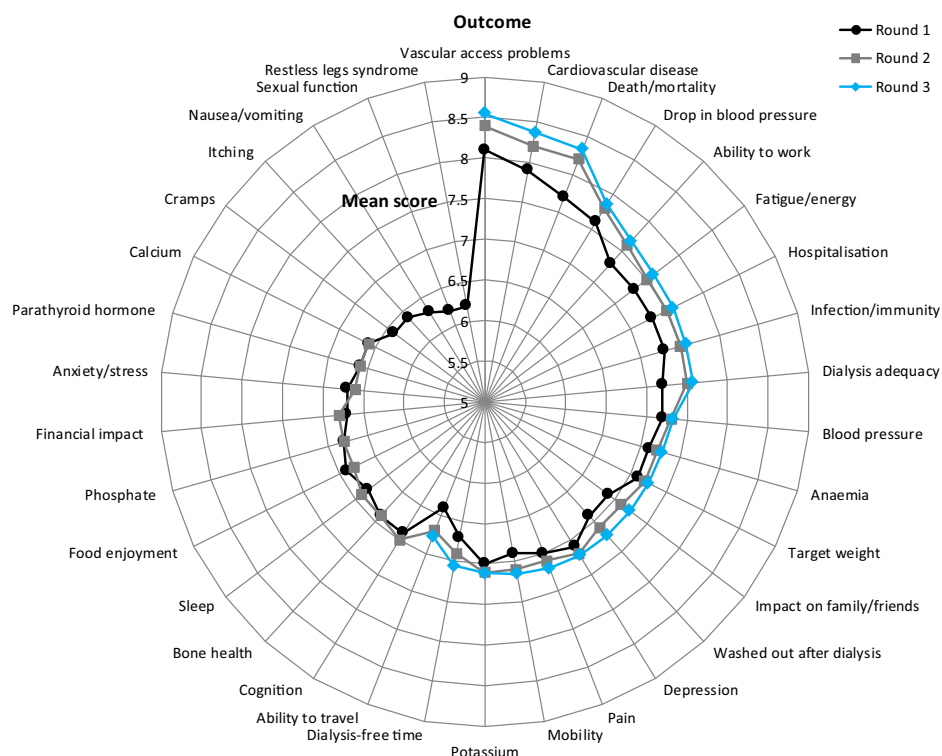
Figure 2. Mean scores of health professionals in rounds 1 to 3. Ordered by round 3 scores. Round 1 (n = 979); round 2 (n = 784); round 3 (n = 688). Mean scores for rounds 2 and 3 are not available for outcomes that were excluded in those rounds.

countries and retained a high response rate of 71% by round 3. The process provided a transparent and systematic way to gain consensus on the importance of outcomes to consider for a core outcome set in HD. However, there are some potential limitations. The Delphi survey was conducted online and excludes participants without access to a computer and internet connection. The survey was only available in English to retain consistency of meaning and for feasibility, although some participants submitted open-text responses in Spanish, which were translated. We also acknowledge that detailed analysis of the open-text responses was not conducted and is beyond the scope of the current study. Given the design of the Delphi survey, we recognize that the results are potentially biased toward participants who are English speaking with access to a computer and internet connection and who have the ability to use a computer.

We included participants from 73 countries and acknowledge that the majority of participants were from high-income countries. We recognize that these outcomes may not be shared by those who did not participate in the study or by individuals or groups within the study. A study of this type must necessarily accept the views of the majority of participants as representing the consensus position and will not represent the view of all individuals.

Although definitions were provided for each outcome domain, we acknowledge the inevitable interaction between some outcome domains, and participants may have interpreted the outcomes differently. We provided participant comments to make explicit how the outcomes were conceptualized, and based on these data, we did not note any apparent differences that may have explained variation in prioritization between patients/caregivers and health professionals. Results from the consensus workshop further support this.[33]

Currently, there is no core outcome set in chronic kidney disease research. The Delphi survey results will be used to establish a core outcome set to be reported in all trials in HD, which is expected to increase the quality and relevance of research. This has been done successfully by the OMERACT initiative, which has improved the reporting and relevance of outcomes in rheumatology trials.[10,24] Recent analyses have shown that a higher percentage of trials are now incorporating the OMERACT core outcome set in rheumatology trials.[24]

Results of this study were discussed at a recent SONG-HD consensus workshop. To effectively implement a core outcome set in chronic kidney disease, patient/caregiver and health professional participants suggested that outcome measures should be feasible and valid and take patient priorities into account.[33] When
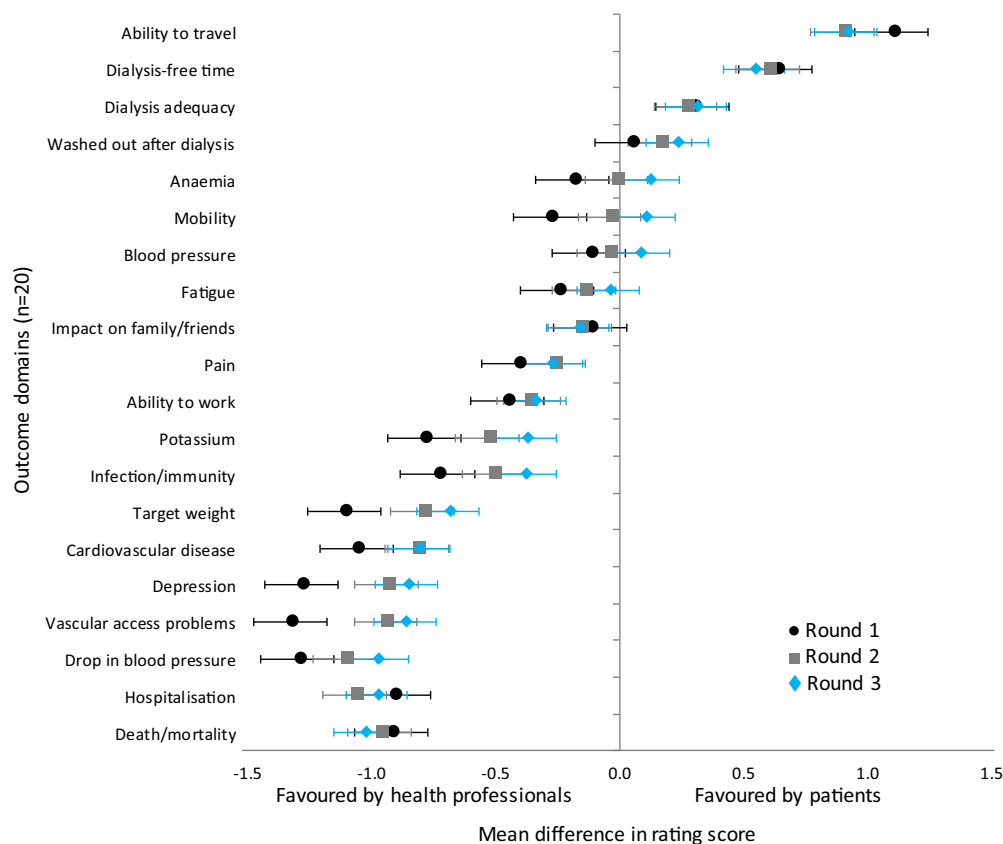
Figure 3. Difference in mean scores between patients/caregivers and health professionals (rounds 1-3). Ordered by difference in mean scores between patients/caregivers and health professionals. Error bars refer to standard deviations.

the core outcome domains have been established, which will be largely informed by results of this Delphi survey, further work will focus on identifying core outcome measures for the outcome domains.

In summary, the international SONG-HD Delphi study identified a prioritized set of outcome domains for trials in HD based on consensus among key stakeholder groups. The top prioritized outcomes by both patients/caregivers and health professionals were vascular access problems, dialysis adequacy, fatigue, CVD, and mortality. Patients/caregivers place the highest value on outcomes that will enable them to maintain their day-to-day well-being and lifestyle. The findings will directly inform the development of a core outcome set to be used in HD trials. Ultimately, this will strengthen the relevance and reliability of trial evidence to support shared decision making for people dependent on HD therapy in order to improve treatment outcomes.

ACKNOWLEDGEMENTS

Contributions: Research idea and study design: all authors; data acquisition: NE, AT, BM, BH, DCW, PT, SC, TH, WvB, WCW, HT-T, SY, SM, DCH, JCC; data analysis/interpretation: all authors; statistical analysis: NE, BS, JCC. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved. NE takes responsibility that this study has been reported honestly, accurately, and transparently; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Peer Review: Evaluated by 2 external peer reviewers and an Acting Editor-in-Chief.

SUPPLEMENTARY MATERIAL

Table S1: Round 1 outcome scores of patients/caregivers and health professionals.

Table S2: Suggested outcomes by participants in round 1.

Table S3: Round 2 outcome scores of patients/caregivers and health professionals.

Table S4: Round 3 outcome scores of patients/caregivers and health professionals.

Table S5: Median rank scores, forced ranking question (round 3).

Figure S1: SONG-HD Delphi process.

Item S1: Collaborating organizations.

Item S2: Survey outcome definitions.

Item S3: Participant comments Delphi survey (rounds 1-3).

Note: The supplementary material accompanying this article (http://dx.doi.org/10.1053/j.ajkd.2016.11.029) is available at www.ajkd.org

REFERENCES

1. Chalmers I, Bracken M, Djulbegovic B, Garattini S, Grant J, Gulmezoglu A. How to increase value and reduce waste when research priorities are set. *Lancet*. 2014;383:156-165.

2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374(9683):86-89.

3. Macleod MR, Michie S, Roberts I, et al. Biomedical research: increasing value, reducing waste. *Lancet*. 2014;383(9912):101-104.

4. Bell CM, Chapman RH, Stone PW, Sandberg EA, Neumann PJ. An off-the-shelf help list: a comprehensive catalog of preference scores from published cost-utility analyses. *Med Decis Making*. 2001;21(4):288-294.

5. Nissenson A. Improving outcomes for ESRD patients: shifting the quality paradigm. *Clin J Am Soc Nephrol*. 2014;9:430-434.

6. Wong G, Howard K, Chapman J, Pollock C, Chadban S, Salkeld G. How do people with chronic kidney disease value cancer-related quality of life? *Nephrology*. 2012;17:32-41.

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

8. Gargon E, Gurung B, Medley N, Altman D, Blazeby J, Clarke M. Choosing important health outcomes for comparative effectiveness research: a systematic review. *PLoS One*. 2014;9:e99111.

9. Ghandi G. Patient-important outcomes in registered diabetes trials. *JAMA*. 2008;299:2543-2549.

10. Boers M, Kirwan J, Wells G, Beaton D, Gossec L, d'Agostino M. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol*. 2014;67:745-753.

11. Svensson S, Menkes D, Lexchin J. Surrogate outcomes in clinical trials: a cautionary tale. *JAMA Intern Med*. 2013;173.612-612.

12. Yudkin J, Lipska K, Montori V. The idolatory of the surrogate. *BMJ*. 2011;343:d7995.

13. Halimi JM, Sautenet B, Gatault P, Roland M, Giraudeau B. Renal endpoints in renal and cardiovascular randomized clinical trials: time for a consensus? *Fundam Clin Pharmacol*. 2012;26(6):771-782.

14. EVOLVE Trial Investigators. Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. *N Engl J Med*. 2012;367(26):2482-2494.

15. Palmer SC, Hayen A, Macaskill P, et al. Serum levels of phosphorus, parathyroid hormone, and calcium and risks of death and cardiovascular disease in individuals with chronic kidney disease: a systematic review and meta-analysis. *JAMA*. 2011;305(11):1119-1127.

16. Palmer SC, Teixeira-Pinto A, Saglimbene V, et al. Association of drug effects on serum parathyroid hormone, phosphorus, and calcium levels with mortality in CKD: a meta-analysis. *Am J Kidney Dis*. 2015;66(6):962-971.

17. Perkovic V, Neal B. Trials in kidney disease–time to EVOLVE. *N Engl J Med*. 2012;367(26):2541-2542.

18. Morton R, Tong A, Howard K, Snelling P, Webster A. The views of patients and carers in treatment decision making for chronic kidney disease: systematic review and thematic synthesis of qualitative studies. *BMJ*. 2010;340:c112.

19. Morton R, Tong A, Webster A, Snelling P, Howard K. Characteristics of dialysis important to patients and family caregivers: a mixed methods approach. *Nephrol Dial Transplant*. 2011;26:4038-4046.

20. Saini P, Loke Y, Gamble C, Altman D, Williamson P, Kirkham J. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. *BMJ*. 2014;349:g6501.

21. Clark M. Standardising outcomes for clinical trials and systematic reviews. *Trials*. 2007;8:39.

22. Prinsen C, Vohra S, Rose M, King-Jones S, Ishaque S, Bhaloo Z. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a 'core outcome set.' *Trials*. 2014;15:247.

23. Gorst SL, Gargon E, Clarke M, Blazeby JM, Altman DG, Williamson PR. Choosing important health outcomes for comparative effectiveness research: an updated review and user survey. *PloS One*. 2016;11(1):e0146444.

24. Kirkham J, Boers M, Tugwell P, Clarke M, Williamson P. Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. *Trials*. 2013;14:324.

25. Dalkey N, Helmer O. An experimental application of the DELPHI method to the use of experts. *Manage Sci*. 1963;9(3):458-467.

26. Chiarotto A, Terwee C, Deyo R, Boers M, Lin C, Buchbinder R. A core outcome set for clinical trials on non-specific low back pain: study protocol for the development of a core domain set. *Trials*. 2014;15:511.

27. Harman N, Bruce I, Callery P, Tierney S, Sharif M, O'Brien K. MOMENT–Management of Otitis Media with Effusion in Cleft Palate: protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials*. 2013;14:70.

28. Waters A, Tudur Smith C, Young B, Jones T. The CONSENSUS study: protocol for a mixed methods study to establish which outcomes should be included in a core outcome set for oropharyngeal cancer. *Trials*. 2014;15:168.

29. Tong A, Manns B, Hemmelgarn B, et al. Standardised Outcomes in Nephrology - Haemodialysis (SONG-HD): study protocol for establishing a core outcome set in haemodialysis. *Trials*. 2015;16(1):364.

30. 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):444-454.

31. Schunemann H, Brozek J, Oxman A. *GRADE Handbook for Grading Quality of Evidence and Strength of Recommendation.* 2009. http://gdt.guidelinedevelopment.org/app/handbook/handbook.html. Accessed February 2017.

32. Morton RL, Snelling P, Webster AC, et al. Dialysis modality preference of patients with CKD and family caregivers: a discrete-choice study. *Am J Kidney Dis*. 2012;60(1):102-111.

33. Tong A, Manns B, Hemmelgarn B, et al. Establishing core outcome domains in hemodialysis: report of the Standardized Outcomes in Nephrology-Hemodialysis (SONG-HD) Consensus Workshop. *Am J Kidney Dis*. 2017;69(1):97-107.

34. Lopez-Vargas PA, Tong A, Howell M, et al. Patient awareness and beliefs about the risk factors and comorbidities associated with chronic kidney disease - a mixed-methods study. *Nephrology (Carlton)*. 2017;22(5):374-381.

35. 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):1605-1614.

36. 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):1033-1046.

37. Purnell TS, Auguste P, Crews DC, et al. Comparison of life participation activities among adults treated by hemodialysis, peritoneal dialysis, and kidney transplantation: a systematic review. *Am J Kidney Dis*. 2013;62(5):953-973.

38. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J*. 2015;24(6):1127-1142.

39. Harman N, Bruce I, Kirkham J, et al. The importance of integration of stakeholder views in core outcome set development: otitis media with effusion in children with cleft palate. *PloS One*. 2015;10(6):e0129514.