Scope and Consistency of Outcomes Reported in Randomized Trials Conducted in Adults Receiving Hemodialysis: A Systematic Review

Bénédicte Sautenet, Allison Tong, Gabrielle Williams, Brenda R. Hemmelgarn, Braden Manns, David C. Wheeler, Peter Tugwell, Wim van Biesen, Wolfgang C. Winkelmayer, Sally Crowe, Tess Harris, Nicole Evangelidis, Carmel M. Hawley, Carol Pollock, David W. Johnson, Kevan R. Polkinghorne, Kirsten Howard, Martin P. Gallagher, Peter G. Kerr, Stephen P. McDonald, Angela Ju, and Jonathan C. Craig

Background: Clinical trials are most informative for evidence-based decision making when they consistently measure and report outcomes of relevance to stakeholders. We aimed to assess the scope and consistency of outcomes reported in trials for hemodialysis.

Study Design: Systematic review.

Setting & Population: Adults requiring maintenance hemodialysis enrolled in clinical trials.

Selection Criteria: All Cochrane systematic reviews of interventions published by August 29, 2016, and the trials published and registered in ClinicalTrials.gov since January 2011.

Interventions: Any hemodialysis-related interventions.

Outcomes: Frequency and characteristics of the reported outcome domains and measures.

Results: From the 362 trials, we extracted and classified 10,713 outcome measures (a median of 21 [IQR, 10-39] per trial) into 81 different outcome domains, of which 42 (52%) were surrogate; 25 (31%), clinical; and 14 (17%), patient reported. The number of outcome measures reported significantly changed over time. The 5 most commonly reported domains were all surrogates: phosphate (125 [35%] trials), dialysis adequacy (120 [33%]), anemia (115 [32%]), inflammatory markers (114 [31%]), and calcium (109 [30%]). Mortality, cardiovascular diseases, and quality of life were reported very infrequently (73 [20%], 44 [12%], and 32 [9%], respectively).

Limitations: For feasibility, we included a sampling frame that included only trials identified in Cochrane systematic reviews or ClinicalTrials.gov.

Conclusions: Outcomes reported in clinical trials involving adults receiving hemodialysis are focused on surrogate outcomes, rather than clinical and patient-centered outcomes. There is also extreme multiplicity and heterogeneity at every level: domain, measure, metric, and time point. Estimates of the comparative effectiveness of available interventions are unreliable and improvements over time have been inconsistent.

Complete author and article information provided before references.

Correspondence to B. Sautenet (benedicte.sautenet@univ-tours.fr)

Am J Kidney Dis. 72(1): 62-74. Published online February 21, 2018.

doi: 10.1053/j.ajkd.2017.11.010

Worldwide, an estimated US $240 billion is invested annually into biomedical research, and about half this funding is from governmental or nonprofit sources.[1,2] It is estimated that 85% of the total investment is wasted due to problems with how research is prioritized, designed, reported, and disseminated.[1,3-7] A major cause of this postulated waste may be attributed to the way that outcomes are measured and reported. Many outcomes reported may have limited clinical or policy relevance, important domains may not be reported, and different measures may have been used. Collectively, these problems may make reliable judgments about the relative effects of interventions difficult and the clinical and policy relevance of published research findings substantially uncertain.[8]

The unreliability of trials introduced by the selective reporting of outcomes in favor of statistically significant outcomes (outcomes reporting bias) and the use of nonvalidated surrogates are well recognized.[9] The more fundamental problem of prioritizing research questions based on commercial or other considerations rather than the needs of patients, clinicians, and regulators and the focus on outcomes that may not be critically relevant to the end-users of research have also been highlighted.[9-12] Despite these major concerns, detailed empirical evaluations of the scope and variability of outcomes in randomized trials and at every level—domain, measurement, threshold, and time point—are sparse.

Hemodialysis is one of the most costly interventions used in health care and imposes an immense burden on health care systems and people.[13-15] An estimated 2 million people with end-stage kidney disease depend on hemodialysis globally; these patients have a mortality rate of 27 deaths per 100 patient-years in the 4 first months after dialysis therapy initiation, which exceeds that of the general population by more than 10-fold.[16-19] Hemodialysis is also an invasive and demanding treatment that requires patients to be attached to a machine for 10 to 30 hours per week, with devastating consequences on quality of life.[20-23] Although the research investment in hemodialysis has been substantial, the translation into improvements in patient outcomes has been modest at best.[24] The aims of this study were to describe the scope,

quality, and consistency of outcome domains and measures in hemodialysis trials to inform strategies to improve outcome selection and reporting and thereby increase the value of future trials to inform clinical decisions.

Methods

Selection Criteria

We searched the Cochrane Database of Systematic Reviews to identify all systematic reviews that included trials involving prevalent patients on maintenance hemodialysis therapy up to August 29, 2016, without language restriction. Search terms are given in Item S1. From each systematic review, we obtained the full text of randomized controlled trials from the list of included studies. To obtain a contemporary sample, we also searched ClinicalTrials.gov (from January 2011 to August 29, 2016) for all published trials that enrolled prevalent patients treated by maintenance hemodialysis and had analyzed published results from the trials. We excluded trials that did not include adults (defined as those aged ≥18 years).

Data Extraction

For each trial, we extracted the following trial characteristics: first author, year of publication, participating countries, sample size, mean age of participants, study duration, intervention type, primary outcome, and all outcomes/outcome measures. An outcome measure was defined as any measurement or event reported separately for all trial arms. All levels of specification of the outcome measures were extracted if reported: domain (eg, kidney function), specific measurement (eg, estimation of glomerular filtration rate by the MDRD [Modification of Diet in Renal Disease] Study equation), method of aggregation (percentage change), specific metric (between the start and end of the study period),[25,26] and time frame in relation to the commencement of the trial and the measure of the outcome.

Analysis

Outcome measures from all included trials were grouped into outcome domains by 2 reviewers (B.S. and G.W.) independently and discrepancies were discussed to reach agreement. The list of outcome domains was reviewed and agreed on by 4 reviewers (A.T., B.S., G.W., and J.C.C.).

Reviewer B.S. grouped all outcome domains into 3 categories: surrogate (biochemical or physiologic outcomes that may or may not be validated), clinical (medical outcomes based on clinician assessment or diagnosis), and patient reported (outcomes reported by patients usually relating to quality of life or symptoms), using standard definitions.[27,28] The classification was agreed on by 2 reviewers (A.T. and J.C.C.).

The number of trials that reported each outcome domain was then calculated. The primary outcome, if specified, was identified and we noted whether multiple primary outcomes were reported in the same trial. For feasibility reasons, we were unable to evaluate the different outcome measures for all 81 outcome domains, but limited our detailed measure-specific analysis to 2 selected frequently reported outcome domains in each category (6 in total), including the measurement, aggregation, metric, and timing reported. We performed 3 sensitivity analyses: excluding trials of less than 3 months in duration, trials of 20 or fewer patients, and trials published after 2010. We compared characteristics of the outcome domains, primary outcomes, and outcome measures by intervals (1977-1980, then 5-year intervals for 1981-2015) with binomial regression, analysis of variance, and χ^2 tests. We performed statistical analyses using R, version 3.2.3 (R Foundation for Statistical Computing) to evaluate change over time.

Results

Trial Characteristics

We identified a total of 362 trials involving 42,081 participants (Fig 1). Trial characteristics are provided in Table 1. The trials were conducted in 49 countries, including the United States (90 [25%] trials), Italy (31 [9%]), Germany (21 [6%]), Canada (18 [5%]), and Japan

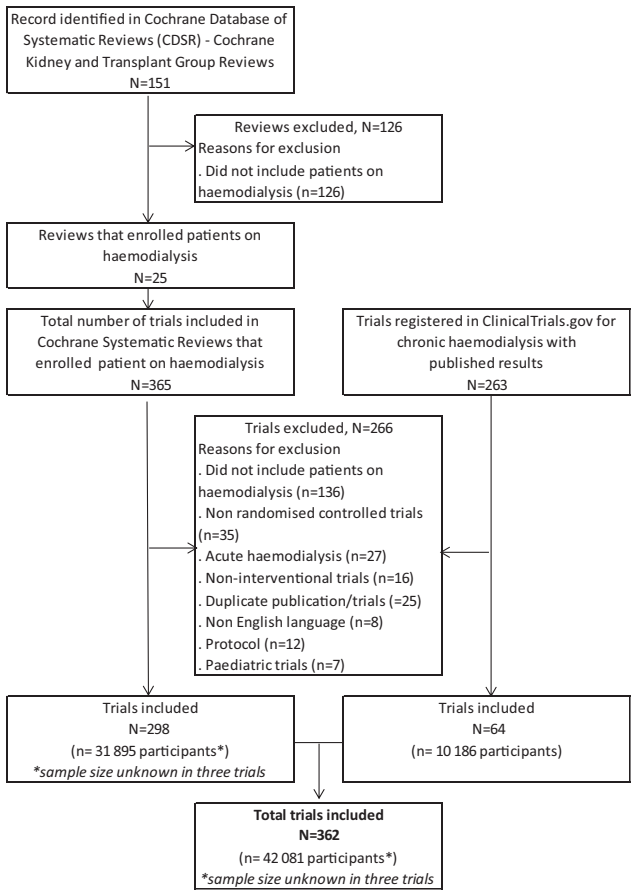

Figure 1. Search results.