Evaluation of Surrogate End Points for Progression to ESKD
Inker, Lesley A.; Heerspink, Hiddo L.

Published in:
American Journal of Kidney Diseases

DOI:
10.1053/j.ajkd.2018.08.017

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Evaluation of Surrogate End Points for Progression to ESKD: Necessary and Challenging

Lesley A. Inker and Hiddo L. Heerspink

Chronic kidney disease (CKD) is recognized as a global health problem and is associated with significant morbidity and mortality, in particular, progression to kidney failure. There is a paucity of therapies to slow kidney disease progression.1 In part, this is because CKD often progresses slowly, requiring randomized controlled trials (RCTs) of long duration, or restricting enrollment to patients with rapidly progressive or more severe forms of the disease. The use of surrogate end points can increase the number of patients with events, thereby allowing for clinical studies of shorter duration or the enrollment of patients in early stages of CKD for whom therapeutic benefits may be larger.2 However, multiple examples in other fields have demonstrated that the premature use of surrogate end points can result in a paradoxical situation in which the intervention shows beneficial effects on the surrogate, but neutral or even harmful effects on the clinical end point, reinforcing the importance of empirical validation of candidate surrogate end points.3,5 Thus, we welcome analyses such as those by Palmer et al6 in this issue of AJKD, which aimed to do just that.

The systematic review by Palmer et al6 evaluated the correlation between the treatment effects of antihypertensive medications on surrogate end points and on end-stage kidney disease (ESKD) in 22 RCTs evaluated using a Bayesian meta-analysis to calculate correlations between treatment effects. They specifically evaluated changes in albuminuria and proteinuria or doubling of serum creatinine level as possible surrogate end points. The results showed low or no correlation between treatment effects on any of the surrogate end points and those on ESKD (range, −0.41 to 0.66 across the various outcomes), with very wide Bayesian credible intervals for all. This led the authors to conclude that there was insufficient information to enable confident use of these markers to guide clinical or regulatory decision making.

The major strength of this systematic review was the use of a Bayesian analysis that estimated the relationship between treatment effects on the surrogate and clinical end points across different randomized trials. However, there are several limitations to this analysis that in our view preclude a conclusion for or against the use of these surrogate end points for RCTs in CKD. First, analyses for each of the outcomes were underpowered. For example, the narrowest credible interval around the R² was −0.14 to 0.73 for doubling of serum creatinine level, and the analysis of regression of albuminuria was performed using only 2 RCTs. Even the analysis for which the most data were available, doubling of serum creatinine level, included only 15 trials. Access to the individual patient data could have allowed for harmonization of definitions across all RCTs, thereby increasing the power of the analysis. Inclusion of previously published RCTs with relevant data would have further increased the statistical power.7-12

Another issue is that the metric used to evaluate the associations, R², is sensitive to the range of the data. The authors included only trials of blood pressure–lowering agents, which also limited the range of the data and therefore the usefulness of the R² metric.

Third, doubling of serum creatinine level has been previously validated and is now accepted as a clinical end point for regulatory purposes. In most studies, it is not practical to just use ESKD; as such, we suggest that doubling of creatinine level could have been included as part of the clinical end point to increase precision of the reference measurement versus treatment effect on albuminuria/proteinuria.

The objective of the study was to evaluate the correlation between treatment effects on candidate surrogates for CKD progression and those on ESKD, and the authors were not able to provide definitive information for that topic. However, what we learn from this article is that evaluation of surrogate end points is complex and there is no easy way to undertake this process. Three types of evidence have traditionally been used to support the validity of a candidate surrogate end point: (1) biological plausibility, (2) epidemiologic association between the candidate surrogate and subsequent development of the clinical end points, and (3) trial-level association between treatment effect on the candidate surrogate end point and treatment effect on the clinical end point.13 The first 2 are necessary but not sufficient criteria because they cannot indicate whether treatment effects on the surrogate end point can be used to make reliable conclusions about those on the clinical end point, which is the critical objective for a surrogate end point. A more direct assessment of the validity of a candidate surrogate end point is provided by the third criterion (ie, trial-level associations), which describes the prediction of treatment effects on the clinical end point using treatment effects on the surrogate end point14 (Fig 1). It is these trial-level associations that have become the focus of the literature on validation of surrogates.
To adequately determine trial-level associations, a large number of RCTs are required across a heterogeneous group of populations and interventions; achieving this requires harmonizing of definitions and thus often individual patient data. Note that the interventions should be restricted to those for which the surrogate end point being evaluated is hypothesized to have biological plausibility as a surrogate end point. Second, the clinical end point used as a reference should be meaningful; that is, short trials with insufficient statistical power to demonstrate a treatment effect on the clinical end point are not useful for comparison to the surrogate. Third, interventions should account for the correlation among the sampling errors on surrogate and clinical end points. Finally, because the goal of the trial-level analyses is to apply these results to future RCTs, it is informative to include a description of a full range of metrics, including the slope and intercept regression line, reporting the association between the treatment effects on the surrogate and clinical end points and confidence and prediction intervals around the regression line.

Investigators, including our group, have been actively involved in addressing how to best evaluate surrogate end points. Two reports evaluated the doubling of serum creatinine level or halving of glomerular filtration rate (GFR) and found that they are valid surrogates. Doubling of serum creatinine level is now accepted for use by regulatory agencies as part of a composite clinical end point. Lesser GFR declines (30% and 40%) are now also accepted end points, although there are caveats with their use, based in particular on the presence, direction, and magnitude of an immediate change in GFR after the initiation of treatment (commonly called acute effect). These caveats are important lessons indicating that even validated surrogates are not always the optimal end points in all circumstances. Large sets of trials, together with simulations, were able to provide these critical lessons. Two separate meta-analyses evaluated changes in proteinuria/albuminuria as a surrogate end point and resulted in different conclusions. The different conclusions likely resulted from differences in the inclusion of trials and analytical methodology, but did not allow for a definitive message regarding the use of change in albuminuria. Importantly, one of the meta-analyses included more than 30 RCTs yet was not sufficiently powered, emphasizing the importance of including a large number of diverse RCTs for conclusive results.

To further address the evaluation of candidate surrogate end points, in March 2018, the National Kidney Foundation (NKF), US Food and Drug Administration, and European Medicines Agency cosponsored a scientific workshop, “Change in Albuminuria and GFR as Endpoints for Clinical Trials in Early Stages of Chronic Kidney Disease” to evaluate surrogate end points for trials of kidney disease progression. For this workshop, we performed an individual patient meta-analysis of more than 45 heterogeneous RCTs using the methodologies described to evaluate the validity of using early changes in albuminuria and GFR as surrogate end points for trials of CKD progression, as well as to provide some initial guidance on when they should optimally be used. Publications are underway.

In conclusion, we agree with Palmer et al that surrogate end points are critical to our field and to advancing therapies in CKD, and that confidence in the results is required before they can be used in clinical or regulatory decision making. The study by Palmer et al shows that the validation of end points is not easy and the process should not be taken lightly. To deliver the most accurate answers, it is likely that rigorous analyses cannot be performed well with summary data from RCTs, but instead requires integration of individual patient data from multiple large-scale RCTs using a uniform analytical method. The nephrology community has been generous in participating in large-scale collaborations that have allowed for such analyses. We are confident that together, we will be able to identify validated surrogates and identify when they should be used for the ultimate goal of advancing therapies to halt the progression of CKD.
null