

# Kidney disease trials for the 21st century: innovations in design and conduct

William G. Herrington<sup>1,2</sup>, Natalie Staplin<sup>1</sup> and Richard Haynes<sup>1,2</sup>

1. Medical Research Council Population Health Research Unit at the University of Oxford, Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, Oxford, UK

2. Oxford Kidney Unit, Oxford University Hospitals NHS Foundation Trust

E-mail: [richard.haynes@ndph.ox.ac.uk](mailto:richard.haynes@ndph.ox.ac.uk)

**Abstract** | Compared to other specialties, nephrology has reported relatively few clinical trials and most are too small to detect moderate treatment effects. Consequently, interventions that are commonly used by nephrologists have not been adequately tested and some may be ineffective or harmful. More randomized trials are urgently needed to address important clinical questions in patients with kidney disease. The use of robust surrogate markers may accelerate early phase drug development. However, scientific innovations in trial conduct developed by other specialties should also be adopted to improve trial quality and enable more, larger trials in kidney disease to be completed in the current era of burdensome regulation and escalating research costs. Examples of such innovations include utilizing routinely collected healthcare data and disease-specific registries to identify and invite potential trial participants, and for long-term follow-up; use of pre-screening to facilitate rapid recruitment of participants; use of pre-randomization run-in periods to improve participant adherence and assess responses to study interventions prior to randomization; and appropriate use of statistics to monitor studies and analyze their results. Nephrology is well positioned to harness such innovations due to its advanced use of electronic healthcare records and the development of disease-specific registries. Adopting a population approach and efficient trial conduct along with challenging unscientific regulation may increase the number of definitive clinical trials in nephrology and improve the care of current and future patients.

## Key points

- Nephrology has the potential to benefit from large streamlined trials similar to those that have led to advances in cardiology and diabetology

- As effective interventions are developed and population level risks fall, larger trial sample sizes are often needed to ensure sufficient statistical power to demonstrate the effects of new therapies
- When moderate effect sizes are anticipated, real-world evidence from association studies cannot provide a reliable estimate of the effect of an intervention; only “randomization” guarantees the elimination of moderate biases
- Potential surrogate outcomes in renal trials include change in albuminuria and estimated glomerular filtration rate slopes; however, no surrogate exists for safety and large trials with sufficiently long follow-up remain necessary
- Precision medicine approaches have the potential to reduce trial sample sizes but might exclude at risk groups who could potentially benefit from an intervention and can lead to time-consuming and costly recruitment procedures
- In this era of complex research governance and burdensome regulation, innovations in trial conduct to enable large-scale invitation, better adherence and low-cost follow-up may be more important than innovations in trial design

## **[H1] Introduction**

Randomized trials are an indispensable tool for those seeking to improve patient outcomes. Over the last four decades, several fields including cardiology have benefited from conducting many large streamlined trials. The central principle in the design and conduct of these trials is that only the information that is necessary to address the primary research question is recorded.<sup>1</sup> Such an approach enables large sample sizes and long follow-up to be feasible. Large streamlined trials have provided a reliable evidence base for thromboprophylaxis in atrial fibrillation, treatments for heart failure, and lowering of atherosclerotic risk.<sup>2</sup> Falling rates of vascular death may be in part the result of widespread adoption of the results of large randomized trials by the cardiology community.<sup>3,4</sup> The field of diabetology has also now been rewarded for embracing large cardiovascular safety studies, with new insights into reducing cardiovascular risk and the identification of renoprotective effects of sodium-glucose co-transporter-2 (SGLT-2) inhibitors<sup>5,6</sup> and anti-GLP-1 receptor agonists.<sup>7</sup>

The field of nephrology has conducted fewer trials than other medical specialities<sup>8,9</sup> to the detriment of patients. Furthermore, the majority of trials in patients with chronic kidney disease (CKD) and/or acute kidney injury (AKI) have been too small to provide reliable answers about the efficacy of the interventions under study. For this reason, the effects of many common practices in nephrology on patient outcomes, such as the use of phosphate binders to lower serum phosphate levels and consequently cardiovascular risk, are uncertain. In some cases, these practices could be harmful. The high individual<sup>10</sup> and societal burden<sup>11</sup> of kidney disease is likely to increase in the future as CKD becomes more prevalent owing to ageing of the global population and maturation of the current epidemic of type 2 diabetes mellitus (T2DM).

The international Standardised Outcomes in Nephrology (SONG) initiative surveyed patients with kidney disease, their carers and clinicians to identify the key health outcomes that need to be improved.<sup>12</sup> They identified clinical outcome priorities for various renal subpopulations, including kidney transplant recipients, patients on haemodialysis, patients on peritoneal dialysis and patients with polycystic kidney disease. **[Au: It isn't necessary to list all of the outcomes for all of the subpopulations but please could you give readers an indication of what types of outcomes are important. For example, are quality of life related outcomes particularly important for patients?]** These priorities now need to be addressed by designing, funding and conducting more high-quality, sufficiently large randomized trials.

Although often considered separately, the tasks of designing and conducting trials are intimately connected and should be based on scientific principles. In this Review, we explain how innovations in trial design and conduct could help to achieve the goal of conducting a greater number of larger renal trials. We discuss the need for randomized trials rather than real-world evidence in nephrology, why such trials need to be larger and how larger sample sizes can be achieved using cost-effective processes. We also explain how to ensure that bias is not introduced following randomization and describe developments in outcome ascertainment, appropriate choice of trial outcomes and the role of non-traditional trial designs. Finally, we

highlight the importance of challenging unscientific and burdensome regulation, which can distract from the primary trial objective and key determinants of quality data.

### **[H1] The need for randomized trials**

It has been argued that collecting sufficient information about various prognostic features in observational studies enables the use of statistical approaches (e.g. propensity-score matching [G]) to attempt to correct for differences between patients who are or are not prescribed a treatment and estimate the treatment effect. However, moderate or even large apparent treatment effects in such studies should not be used to guide clinical decision making because such analyses cannot guarantee elimination of moderate systematic biases. Despite technically proficient analysis, a high chance remains of incomplete adjustment for differences between patients who are prescribed medications and those who are not. Many examples exist in which the apparent treatment effects are implausibly large or are quantitatively or qualitatively different from the results of randomized trials.<sup>13,14</sup> For example, analysis of real-world data predicted that hormone replacement therapy (HRT) would reduce the risk of coronary disease, whereas the results of subsequent, large-scale randomized trials showed the opposite effect.<sup>15,16</sup> Observational studies and post-marketing surveillance that suggested an association between statin use and substantial risks of cataracts or cognitive decline have also been refuted by a comprehensive review of relevant randomized trials.<sup>17</sup>

Observational analyses using Mendelian randomization [G] (MR) methods may be less subject to confounding and reverse causality than classical “real-world” observational analyses. The MR approach harnesses the random assortment of genetic variation before conception to infer how life-long exposure to inherited differences in a risk factor relate to a disease.<sup>18</sup> If a naturally occurring variation in the gene that encodes a therapeutic target exists, MR may predict the on-target effects of a treatment.<sup>19</sup> However, MR requires several assumptions to hold, including the absence of confounding associations between the genetic instrument and the outcome, and cannot quantify any unintended off-target effects of an intervention.<sup>20</sup> MR experiments can therefore substantially strengthen the rationale

for performing randomized trials, but are not substitutes for these trials. Properly conducted, adequately powered trials are the only method that eliminates systematic biases<sup>21</sup> and measures all the clinical effects of an intervention. They therefore remain the gold-standard for identifying and quantifying the beneficial and adverse effects of interventions in clinical practice.

### **[H1] The need for large renal trials**

Clinicians might consider that evidence from large randomized trials might not apply to individual patients. However, one of the main reasons why trials need to be so large is because of the differences between patients. It is very unlikely that patients who are recruited into large randomized trials will be so different from a typical patient that the relative effect of treatment will differ in a meaningful way. Trial participants are, however, selected from specific populations and the absolute effect of an intervention (often quantified as the number needed to treat) should therefore generally not be extrapolated to individual patients. A more suitable approach to estimate the absolute benefits or harms of an intervention for an individual patient is to estimate their absolute risk of an event using an established risk score and then apply the trial-reported relative risk to estimate the effect of the intervention (Box 1).

Although medical interventions with large beneficial effects are occasionally identified,<sup>22,23</sup> most interventions only have small to moderate effects on cardiovascular and renal outcomes (i.e. a proportional reduction of up to ~25%), even when they have large effects on intermediate disease markers. For example, antihypertensive therapy and statin-based therapy can substantially reduce blood pressure and LDL-cholesterol levels, respectively, but their relative effects on cardiovascular disease outcomes are more moderate. A 10 mmHg reduction in systolic blood pressure<sup>24,25</sup> or a 1.0 mmol/l reduction in LDL-cholesterol level will reduce the risk of myocardial infarction by only ~20%.<sup>26</sup> Such small or moderate effects of interventions on clinical outcomes in renal disease may occur because multiple pathophysiological mechanisms are generally responsible for disease progression and/or complications whilst a single intervention is likely to only appreciably modify one mechanism.

Meta-analyses are a useful tool to estimate treatment effects for sample size calculations and their findings support the contention that effect sizes (when they exist) are usually moderate.<sup>27,28</sup> When estimating required trial sample sizes, consideration should be given to the fact that meta-analyses that include small trials often report larger effects than subsequent large definitive trials.<sup>29</sup> This difference may be a consequence of the generally short duration of early trials or the fact that small trials may tend to only be published if the results are positive.

The burden of multiple treatments and the increased frequency of symptoms that may be attributed (rightly or wrongly) as an adverse effect, mean that high levels of non-adherence to treatment may also be common among patients with kidney disease, particularly those with more advanced disease. Non-adherence has a substantial eroding effect on study power, which is discussed further below. A combination of over-optimistic anticipated treatment effects and poorer than expected adherence may explain why kidney disease trials that reported negative results, such as the EVOLVE trial (discussed further below), might have missed potential benefits of the intervention.<sup>30,31</sup>

## ***[H2] Eligibility criteria and uncertainty***

The likelihood of obtaining a widely generalizable trial result is increased by designing trials with simple eligibility criteria and adopting the uncertainty principle. This principle asserts that trial entry is only ethically justified if both the clinician and patient are substantially uncertain as to which treatment option is optimum.<sup>32</sup> Exclusion criteria should therefore restrict entry to those with definite indications or definite contraindications for the test intervention. Use of simple inclusion criteria to identify an at-risk population can then facilitate large-scale invitation of potential participants. Once these individuals have been identified, application of the uncertainty principle can be applied following formal screening of the potential trial participant at a trial visit.

Many examples exist in which the nephrology community has adopted treatments before adequate evidence of efficacy or safety has become available.<sup>33</sup> For example, in the first clinical tests of erythropoiesis-stimulating agents (ESAs), the use of placebo was considered to be unnecessary, perhaps because the reduced

requirement for blood transfusions with ESAs was considered to be a major advance. The placebo-controlled TREAT trial investigated the efficacy of an ESA targeted to increase haemoglobin levels from 9 g/dl (90 g/l) to ~13 g/dl (130 g/l)<sup>35</sup>. At the time of its design, the use of ESAs had become so common that the inclusion of a placebo group was considered by some to be unethical. However, before the trial was completed, a meta-analysis of 9 trials in just over 5,000 patients demonstrated that normalisation of haemoglobin levels using ESAs was associated with increased risk of death from any cause.<sup>34</sup> The TREAT design was then scrutinized for the high haemoglobin target in the intervention group. Eventually, the trial provided reliable data that targeting a higher haemoglobin level with ESAs does not reduce the risk of cardiovascular disease in patients with diabetes and CKD, and may in fact cause harm.<sup>35</sup> When these data were published ESAs had already been available for more than two decades.

## ***[H2] Innovations to facilitate recruitment***

Association studies using data from electronic healthcare records do not provide reliable estimates of the effects of many medical interventions. Such data can, however, be invaluable to facilitate more efficient trial recruitment.

***[H3] Use of routinely collected healthcare data.*** In many developed countries, nephrology is advanced in its use of electronic healthcare records and disease-specific registries. In Europe, 36 countries now contribute data on patients receiving renal replacement therapy to the ERA-EDTA register,<sup>36</sup> the US Renal Data Service is well established, and in the UK, over 20,000 people are registered on the Renal Association's National Register of Rare Kidney Diseases (RADAR).<sup>37</sup> Large numbers of people with kidney disease, including those with rare conditions, can be efficiently identified using such data.<sup>38</sup> For example, co-ordinated collaboration of national Alport syndrome registries has been suggested as a mechanism to establish international trials in this rare condition.<sup>39</sup> These approaches are even more scalable for more common conditions, such as CKD and diabetes. For example, multiple retinopathy registries were leveraged to generate hundreds of thousands of invitations and recruit over 15,000 people with diabetes in the UK into the ASCEND placebo-controlled trial of aspirin and/or omega-3 fish oils for the

prevention of cardiovascular disease, which was conducted using mail-based methods.<sup>40,41</sup>

**[H3] Pre-screening.** Another time and cost-effective method to accelerate patient recruitment is pre-screening. This process involves large-scale invitation of potential participants from locally held clinic lists or databases to obtain provisional agreement to join a trial while other time-consuming aspects of the study — for example, acquiring the relevant approvals, setting up the drug supply and developing IT systems — are still being completed. Potentially eligible patients are provided with relevant information about the trial and asked if they would agree to attend a future study visit.

Pre-screening can be done at the clinic or more efficiently by mass-mailing invitations followed shortly by direct patient contact by their clinical team. This procedure enables local trial sites to determine whether sufficient numbers of potential participants exist to meet recruitment targets and to accelerate recruitment from the pool of interested patients once the study formally begins. Pre-screening also empowers patients by offering them an opportunity to volunteer to participate in trials rather than being selected by their clinical team. During Pre-screening, it is important to communicate that a patient's agreement to join the trial during the pre-screening phase is provisional, as the trial might not proceed in certain regions or sites for a number of reasons. For those co-ordinating the trial, pre-screening also enables early assessment of the likely composition of the future trial population (Box 2). The UK HARP-III trial of combined sacubitril and valsartan therapy versus irbesartan in patients with CKD found that sites that adopted pre-screening had twice the recruitment rate of those that relied on more traditional invitation approaches.<sup>42</sup>

### **[H1] Avoiding bias after randomization**

Several ways exist in which biases can be introduced into a trial following patient recruitment and randomization. These include incomplete follow-up, unnecessary unblinding, differential non-adherence and inappropriate statistical analysis.

### ***[H2] Innovations to maintain adherence***



Poor adherence to the study intervention in clinical trials can lead to underestimation of the effects of the intervention. High rates of drop-out (participants who are allocated to active treatment stopping that treatment) or drop-in (participants who are allocated to placebo starting open-label active treatment) quickly erodes the trial's statistical power. In general, the effect on statistical power of a single treatment drop-out or drop-in post-randomization is equivalent to that of recruiting two fewer patients to the trial.

The EVOLVE trial of cinacalcet versus placebo in patients with secondary hyperparathyroidism on haemodialysis provides an excellent case study of the effect of loss of adherence<sup>30</sup>. During the course of the trial, nephrologists became less uncertain about the benefits (and presumably more confident about the safety) of cinacalcet even though large-scale evidence had not been reported. Commercially available cinacalcet was therefore prescribed to 20% of participants who were allocated to the placebo group. This drop-in contributed to the collapse of study power for the primary assessment (a composite of death, myocardial infarction, hospitalization for unstable angina, heart failure or a peripheral vascular event) from 90% to 54%<sup>31</sup>. The EVOLVE trial was therefore unable to provide a reliable conclusion regarding its primary assessment, to the detriment of all patients on dialysis.

If open-label use of a study intervention in an individual patient is likely, that patient should not be included in the trial. Exclusion of patients with characteristics that would make drop-in highly likely should be considered in eligibility criteria. For example, in the SHARP trial of lowering LDL-cholesterol levels using statin therapy, patients with prior myocardial infarction were excluded as drop-in to open-label statin therapy among these participants was considered likely.<sup>43</sup> However, such criteria need to be combined with a wider appreciation of the benefits of embracing and retaining uncertainty until clear evidence of benefit and safety are demonstrated. To protect the safety and well-being of trial participants, and to reassure investigators, independent data monitoring committees regularly review unblinded trial data and will terminate a trial early if evidence exists of overwhelming efficacy or recommend modification of the trial design if safety concerns arise either for the overall study population or for specific patient subgroups.

Adherence to study procedures can also be improved by minimizing the impact of the study on the daily routine and clinical care of the patient and by making procedures for research teams simple and time efficient. For example, requiring baseline physical examinations is a drain on local investigators and rarely provides rewarding data. In CKD progression trials, the practice of asking participants for an extra confirmatory blood test about 30 days after they potentially meet an important percentage decline in an eGFR-based outcome is an unnecessary burden for participants and trial teams. Regular and relatively frequent follow-up eGFR measurements (e.g. 2-3 times a year) enable a definition of sustained change in eGFR to be applied without this burden, and without risk of differential ascertainment between groups owing to extra follow-up of selected participants.

**[H2] Pre-randomization run-in phase** *[Au: Unfortunately, there is not space to change this title to “Pre-randomization enrichment run-in phase” because our character limit for H2 subheadings is 39 characters including spaces. I suggest that you keep the current subheading.]*

Another aspect of trial design that can improve adherence is a pre-randomization run-in phase following the first study visit (Box 3). This approach enables early assessment of compliance to a study intervention before a participant is randomly assigned to a study group.<sup>44</sup> The run-in period can be placebo only to provide time for patients, investigators and clinicians to reconsider the patient's participation and enable tests to check eligibility to be conducted.

An active run-in period in which all participants receive the study intervention is an alternative approach that might be particularly helpful in maximizing adherence when a drug has common adverse effects that may limit tolerability.<sup>45</sup> Part of the pre-randomization run-in (referred to as an enrichment phase) of the SONAR trial of the endothelin A receptor antagonist atrasentan in patients with T2DM and CKD served this function as those who developed substantial fluid retention on atrasentan were not included in the randomized trial phase<sup>46</sup>. During the follow-up period, adherence to atrasentan was almost identical to that of placebo.

**[H2] Ensuring reliable outcome ascertainment**

As incomplete outcome ascertainment during follow-up could introduce systematic bias, all trial participants should continue to attend study visits even if they have stopped treatment or reached a study outcome. A key design feature to reduce systematic ascertainment bias is to implement double-blind design so that investigators and participants do not consciously or subconsciously modify their reporting of events owing to knowledge of treatment allocation. Open-label studies should be avoided whenever possible by using placebo or sham operations in the control groups.

The contrasting findings of open-label<sup>47</sup> and double-blind<sup>48</sup> renal denervation trials in patients with resistant hypertension illustrate the major distortion of results that can be caused by differential biases. In the SYMPLICITY-HTN trials, renal denervation versus standard care seemed to result in large reductions in clinic-measured systolic blood pressure (-32 mmHg) when an open-label trial design<sup>47</sup> was used, whereas smaller reductions were observed in the open-label studies that employed automated ambulatory blood pressure measurements.<sup>49</sup> However, no significant reductions in clinic or ambulatory systolic blood pressure with renal denervation were apparent in the subsequent large confirmatory SYMPLICITY-HTN3 trial that used a sham procedure to blind both observers and participants to the intervention.<sup>48</sup> Open-label designs may provide reliable results when there is systematic (i.e. unbiased) and complete follow-up for clinical outcomes that are not subjective (e.g. all-cause mortality) and perhaps when outcomes are ascertained and adjudicated by individuals who are blinded to treatment allocation. However, the SYMPLICITY-HTN trials illustrate that cautious interpretation of open-label trials is still required, as small observer biases and/or changes in participant behaviour which differ between study groups can distort results.

The benefits of streamlining study visits by collecting minimal data are described above. Using electronic case report forms completed with real-time data checks rather than relying on retrospective checks is an efficient method to ensure the completeness and internal validity of data. Electronic systems also enable coordinating centres to monitor key aspects of the trial that are critical to its quality, such as completeness of follow-up.<sup>50</sup> In sufficiently large trials, central statistical analysis of data from all sites to detect missing data, outlying sites or fraud<sup>51</sup> may be

more efficient and is certainly much more cost-effective than traditional frequent ‘on-site monitoring’ in which a member of the trial coordinating team visits a study site to manually cross-check data. Central statistical monitoring can be used to identify sites with potential problems and trigger on-site monitoring visits for these sites.

## ***[H2] Use of registry data for follow-up***

Although simple observational associations using “real-world” data should rarely be used to assess moderate treatment effects,<sup>14</sup> routinely collected healthcare data may provide a valuable resource to streamline trial follow-up. Linkage of trial cohorts to full-coverage national mortality, cancer or dialysis registries can provide the sole means of very cost-effective follow-up both within the study period and in the long-term following trial completion.<sup>52</sup> Alternatively, such data can be used to supplement direct participant follow-up to maximize the completeness of follow-up data.<sup>41</sup>

Trial publications often provide detailed definitions of clinical outcomes as applied by adjudication panels. These panels are tasked with verifying reported outcomes using source documents such as hospital notes. However, increasing evidence indicates that in double-blind cardiovascular outcomes trials, relative risks that are estimated from adjudicated outcome data are similar to those that are estimated when pre-adjudication outcomes are used.<sup>53,54</sup> These findings indicate that for certain cardiovascular outcomes, reports from patients and investigators are sufficiently reliable and any misclassification (e.g. as a result of imprecise diagnosis or a small number of missed events) is usually non-differential with respect to study intervention allocation. Consequently, use of adjudicated data has little effect on relative effects when the trial is large, randomized and blinded.<sup>1,17</sup>

Depending on the research question, systematic confirmation of some outcomes is important (e.g. determining a single underlying cause of death using pre-specified rules or differentiating ischaemic and haemorrhagic strokes). However, verification of outcomes that are reliably reportable by patients (e.g. kidney transplantation or receipt of dialysis) is a trial process that could be streamlined in certain situations. Examples now exist of cardiovascular trials embedded within registries that only use routinely collected healthcare data for participant recruitment and follow-up.<sup>55</sup>

## ***[H2] Appropriate selection of outcomes***

Primary outcomes should be both clinically important and likely to be modified by the effect of the study intervention. Because of the time required for many patients with CKD to progress to ESRD (i.e. require dialysis or transplantation), biochemical surrogate outcomes have been proposed that might enable trials to be smaller and shorter. To be considered a suitable surrogate, these outcomes should correlate with and fully capture the effect of treatment on clinical outcomes.<sup>56</sup> If a surrogate outcome meets these criteria, a test of the treatment effect on this outcome will be a valid test of the treatment effect on the clinical end point. Although some biomarkers do fulfil these criteria (e.g. a substantial decrease in eGFR<sup>57</sup>), many proposed surrogate outcomes are not true surrogates.

The use of albuminuria as a potential surrogate outcome has been the subject of much debate.<sup>58,59</sup> A meta-analysis provided some support for change in albuminuria as a surrogate outcome (particularly in patients with high levels of albuminuria), but predicted that treatments would need to reduce albuminuria by 30% to have a high likelihood of a significant effect on a clinical renal disease progression end point.<sup>60</sup> As albuminuria is highly variable within individuals, apparent changes may not represent treatment responses.<sup>61</sup> Moreover an increase in albuminuria does not preclude a treatment response as disease progression (or even regression) might occur despite such an apparent response. As uncertainty exists regarding the validity of albuminuria as a surrogate outcome, the recommendation for post-marketing trials to verify the results of trials that use this outcome is important and appropriate.<sup>62</sup> Moreover, we do not fully understand whether change in albuminuria is a requirement for renoprotection with all interventions. The apparent beneficial effects of SGLT2 inhibitors on renal progression in patients with T2DM may not require the existence of albuminuria and therefore might be independent of effects on albuminuria.<sup>5,6</sup>

The slope of change in eGFR over time has also been proposed as a surrogate renal progression outcome. As change in eGFR slope correlates with established eGFR-based surrogate outcomes (e.g. doubling of creatinine levels), it may be a more appropriate outcome than change in albuminuria. Analysis of eGFR slopes

requires use of appropriate statistical methods to avoid introducing bias<sup>63,64</sup> (discussed further below).

Importantly, no surrogate for safety exists and smaller, shorter trials may not have sufficient statistical power to detect plausible harms of treatment. Reassuringly, several of the ongoing trials of hypoxia inducible factor stabilizers (i.e. prolyl hydroxylase domain enzyme inhibitors) include cardiovascular outcomes in addition to outcomes based on change in haemoglobin levels.<sup>65</sup> When the intervention is relevant and trial conduct methods allow, collection of SONG-defined core outcomes and outcome measures should also be designed into new renal trials.<sup>12</sup>

When combining end points, outcomes should be of broadly similar clinical importance and likely to be affected in the same direction. For example, trials of antithrombotic therapy need to consider ischaemic and haemorrhagic stroke separately to enable the net benefit in different types of participant to be accurately estimated, whereas a trial of blood pressure-lowering treatment might reasonably combine these subtypes because the epidemiological associations of blood pressure with ischaemic and haemorrhagic stroke are qualitatively similar.

All-cause (or total) mortality is one of the least sensitive end points and is rarely an appropriate outcome because an effective treatment is likely to have a moderate beneficial effect on only a few causes of death and, therefore, only a small effect on total mortality. This point is particularly pertinent in the case of renal trials as no single cause of death dominates mortality in renal populations. By contrast, at least 80% of deaths in patients with heart failure and reduced ejection fraction are cardiovascular.<sup>66</sup> Variation in causes of death between populations also means that an effect on all-cause mortality in one trial population might not be generalizable to a different population.

## ***[H2] Appropriate statistical analysis***

Even after proper randomization and complete follow-up for relevant outcomes, bias can be introduced by inappropriate statistical analysis. Excluding patients who do not adhere to their study intervention is surprisingly common in analyses of trials in kidney disease. A systematic review of renal trials reported that more than half did

not include all of the randomly assigned participants in their analyses.<sup>67</sup> Such exclusion will introduce bias if the prognosis of those who are excluded from one study group differs from that of those who are excluded from another group. Such bias is highly likely because adverse effects are a common cause of non-adherence and are often more pronounced in patients who have more severe illness. The findings of the Coronary Drug Project provide an example of the effect of such bias<sup>68</sup>. In this study, patients who took  $\geq 80\%$  of their allocated clofibrate during the 5-year follow-up period had lower mortality at 5 years than those with poorer adherence (15.0% versus 24.6% respectively,  $P=0.0001$ ). However, the difference in mortality between patients who did or did not take  $\geq 80\%$  of their allocated placebo was even more striking (15.1% versus 28.3%,  $p<0.00001$ ). This finding demonstrates that on-treatment analyses (i.e. those that censor participants when they stop taking the study intervention and include lag-censoring approaches) can seriously distort the balance of randomization and should not be encouraged. Instead, the primary analyses of trials should use the intention-to-treat principle — that is, they should compare the outcome among all those who were originally allocated to the treatment group (regardless of whether they ever took the treatment or for how long) with the outcome among all those who were allocated to the comparator group.

When presented with an overall positive or negative trial result, clinicians frequently want to know in which subgroup(s) of participants the treatment is particularly beneficial or harmful. Looking at the treatment effect in the subgroup of interest in isolation is not appropriate because trials are rarely powered to assess effects in subgroups directly. A more appropriate question is whether the treatment effect observed in a subgroup differs significantly from the overall treatment effect. This question can be addressed by using statistical tests to determine whether or not the effect in a subgroup differs from the effects in those who are not in that subgroup (i.e. **statistical tests for heterogeneity [G]** or trend). In the absence of evidence of heterogeneity, the most appropriate approach is to generalize the observed relative effects for the entire trial cohort.<sup>21</sup> When a nominally significant statistical test for heterogeneity is identified, consideration should be given to the size of the subgroup (as random error is more common with small numbers of outcomes); how many subgroups were pre-specified to be assessed (as the probability of falsely finding

heterogeneity is increased by multiple testing of hypotheses); the effects on related outcomes (as effects on fatal outcomes are usually concordant with corresponding outcomes that did not lead to death); and the biological plausibility of any modification of treatment effect (ideally stated a priori in the data analysis plan).

The use of responder analyses based on different post-randomization responses to an intermediate marker of effect can also introduce bias during analysis. Such responses are only measured in the active treatment group so no comparable placebo groups exist for the responders and non-responders (Figure 1).<sup>69</sup> Reliable, pre-specified subgroup analysis by different levels of responder status requires such status to be determined in all participants pre-randomization. This determination can be achieved using a pre-randomization run-in period during which all potential participants receive the study intervention. Such a design was used in the Heart Protection Study, which found no good evidence for variation in the beneficial effect of simvastatin versus placebo on major vascular events in patients with different relative changes in LDL-cholesterol levels during the active run-in phase.<sup>70</sup>

Increases in creatinine levels after starting ACE inhibitor treatment identify a high risk group that requires close monitoring; the risks and benefits of ACE inhibitor prescribing in this group should carefully considered.<sup>71</sup> A post hoc analysis used the pre-randomization active run-in phase of the ADVANCE trial, which assessed the effects of perindopril-indapamide versus placebo in 11,000 people with T2DM, to explore this hypothesis<sup>72</sup>. Consistent with the findings of observational studies,<sup>71</sup> the post hoc analysis showed that participants who experienced an increase in creatinine levels during the run-in period were at increased risk of the composite macrovascular and microvascular outcome. However, the relative beneficial effects of perindopril-indapamide versus placebo on this outcome were similar irrespective of the degree of change in creatinine on starting the therapy (i.e. there was no evidence of heterogeneity).<sup>72</sup> These randomized analyses therefore suggest that change in creatinine level is not a reliable marker of response (or harm) when starting perindopril-indapamide therapy.

Appropriate statistical methods also need to be employed for trials that use eGFR-based outcomes. Simply comparing eGFR values in different groups at final follow-



up can be biased if baseline imbalances in eGFR exist. To avoid this bias use of eGFR change scores (i.e. subtracting the follow-up value from the baseline) has been suggested, but this approach does not adequately address any baseline imbalance because of regression to the mean. Therefore, use of regression-based methods that fit straight lines relating follow-up values to baseline is recommended.<sup>73</sup>

Particular care needs to be taken not to introduce bias into analyses of rate of decline in eGFR in kidney disease trials. An under-appreciated problem is that standard statistical methods (such as **linear mixed models [G]**) assume that the rate of decline in eGFR does not differ between patients who drop-out during the course of the study and those who have complete follow-up. These methods can give biased estimates of eGFR slope when patients are censored for reasons related to their underlying rate of eGFR decline (e.g. death or progression to ESRD).<sup>63</sup> This bias is particularly problematic for trials in which the rates of these censoring events differ between treatment groups, as standard methods will no longer reliably estimate the effect of treatment on decline in eGFR. Approaches have been developed that can reduce this bias, such as **shared parameter models [G]** that jointly model eGFR slopes and the censoring events.<sup>64</sup>

### **[H1] Non-traditional trial designs**

Trial methodology has been developed to adapt to different questions that need to be addressed in healthcare systems and in response to new concepts, such as precision medicine. These non-traditional trial designs are discussed below.

#### ***[H2] Cluster randomized trials***

Traditional parallel group trials involve randomization at the level of individual participants (**Figure 2**). Cluster randomized trials randomly allocate groups of participants (for example, all those at a particular kidney unit) rather than individuals (**Figure 3**). Such designs are increasingly used to evaluate service delivery or policy interventions. The established infrastructure of dialysis centres linked to electronic healthcare record systems or registries is particularly well-suited to efficient delivery of such trials.<sup>74</sup>

When using a cluster design, sample sizes need to be larger than those that are required for a similarly powered traditional parallel-group trial because intracluster correlation reduces study power, particularly if clusters are large and few in number.<sup>75</sup> In addition, the randomized groups will only be balanced if either the vast majority or a random sample of eligible people within the cluster are included in the trial.<sup>76</sup> Blinded identification and recruitment of individuals within each cluster is therefore needed to avoid selection bias.<sup>77</sup>

In clinical trials, a meaningful difference between allocated study groups needs to be established and sustained. In cluster randomized trials, there may be less contact between the participants and the research team than in a conventional trial and individual patients might be enrolled without being invited and given the opportunity to choose whether to participate. For certain interventions, this lack of patient contact and sense of personal contribution to the research effort could adversely affect adherence. A key lesson is provided by the TiME trial, which used innovative methods to centrally link up and randomly assign nearly 270 US dialysis centres to long dialysis sessions ( $\geq 4.25$  hours) versus usual care<sup>78</sup>. The trial was stopped early at one year because the mean difference in dialysis duration between the two clusters was only 9 minutes. The investigators hypothesized that involving the participants in the trial design and a pilot study might have led to protocol modification that could have improved adherence to the intervention.

Stepped wedge cluster randomized trials are a variation on the cluster design in which clusters are randomly selected to start the study intervention at regular intervals. Initially no cluster but eventually all clusters receive the intervention. This design means that each cluster provides outcome data during the study intervention and control periods. Analyses are more complicated than for traditional parallel group and cluster randomized trial designs as they need to account for any temporal changes in event rates that might bias assessments of treatment effects, but the ethical advantage of this design is that no group is denied the intervention.<sup>79</sup> Such an approach was successfully used to assess the effects of sequentially implementing a multifaceted AKI intervention (including e-alerts, a care bundle, and an education program) across 5 UK hospitals.<sup>80</sup> This trial, which included 20,000 people who were hospitalized with AKI, found that the intervention did not reduce

AKI-related mortality but did reduce length of hospital stay and led to improvements in markers of quality of care.

## ***[H2] Enrichment designs***

By designing and conducting trials that are sufficiently large to identify modest treatment effects, cardiology has established a wide evidence base with a pragmatic population-based approach.<sup>81</sup> The generalizability of the trial findings to individual patients is assessed once the trials are completed using pre-specified subgroups analyses. A remarkable finding in cardiology is how infrequently there is evidence that the relative effects of treatments vary by common characteristics once sufficient numbers of people with different characteristics have been studied. For example, the aggregated worldwide evidence on statin-based therapy has found similar effects on atherosclerotic risk per 1 mmol/L reduction in LDL-cholesterol in almost every subgroup tested<sup>82</sup> with perhaps the exception of those on dialysis.<sup>26</sup>

In contrast to cardiology, oncology trials have shown that mutations in certain oncogenes do modify the effect of targeted treatments, with some treatments only licensed for use in cancers with specific mutations.<sup>81</sup> These observations and the promise of large-scale genotyping have led to precision medicine initiatives.

Several post hoc analyses of trial data have led to the hypothesis that polymorphisms in the *ACE* gene may modify the benefits of renin-angiotensin system inhibition in proteinuric diabetic and non-diabetic kidney diseases.<sup>83</sup> Kidney disease trial populations have therefore been postulated to include groups of people who may differ in their responses to study interventions.<sup>84</sup> To address this hypothesis, renal trials incorporating an enrichment phase into an active run-in have emerged. The SONAR trial pre-specified that its primary assessment would only include participants who experienced a  $\geq 30\%$  reduction in albuminuria during the active atrasentan run-in phase (termed “responders”).<sup>85</sup> However, despite large differences in the apparent albuminuria response, the effect of treatment on the primary outcome (doubling of serum creatinine or ESRD) was similar among the responders (HR 0.65; 95% CI 0.49-0.88) and non-responders (HR 0.75; 95% CI 0.55-1.03), with no evidence of heterogeneity ( $p=0.41$ ).<sup>46</sup> It is likely that the inherent variability in albuminuria meant that it was not possible to reliably distinguish

“response” from “non-response”. Moreover, the exclusion of large numbers of participants who apparently “fail to respond” during the run-in could increase trial costs substantially and may have other disadvantages, which are discussed further below.

## ***[H2] Adaptive design***

Steering committees who oversee traditional parallel-group trials may review participant characteristics and event rates during the trial and use this information to re-assess sample size calculation assumptions before completion of randomization. Expansion of the sample size may be required if event rates are lower than the original predictions. Adaptive trial design is a further development of this approach in which adaptations are pre-specified and design modifications are enforced if certain results emerge. These adaptations may include changing the sample size based on observed event rates, responding to a formal interim analysis by dropping underperforming or toxic doses or enriching for groups or subgroups of participants in whom responses appear to be larger. Such trials require transparent reporting and their interpretation can be complex.<sup>86,87</sup>

Adaptive design may accelerate early phase trials aimed at establishing an optimum dose. The importance of dose-finding and tolerability studies in specific renal populations has been highlighted by lessons in previous trials. For example, the ASCEND trial in patients with diabetic kidney disease selected a high dose of avosentan (25-50 mg) based on a previous pharmacodynamic study in patients with T2DM and relatively preserved kidney function in whom adverse effects were infrequent. However, use of this dose resulted in an excess of fluid overload, heart failure and other cardiovascular outcomes in the study population of ~1,400 patients with T2DM and CKD stages 3-4, which led to early trial termination<sup>88</sup>. The investigators raised the question of whether lower doses of avosentan might have had similar anti-albuminuric effects to those seen in the trial without the adverse effects.

## ***[H2] Master protocol trials***

An innovation from oncology trials is the simultaneous testing of more than one intervention or disease within a trial network using a master protocol (Figure 4).

Basket trials refer to the approach of testing a single intervention on multiple diseases or disease subtypes using a single protocol. The term umbrella trial has been used to describe testing multiple interventions on a single disease in a single protocol.<sup>89</sup> For example, in oncology an intervention may be selected based on the molecular phenotype of the tumour<sup>90</sup>.

A platform trial is a variation on the theme of an umbrella trial in which multiple different interventions or doses are tested versus a shared placebo or control group.<sup>91</sup> Such trials usually use a surrogate marker of effect and interventions with poor responses are dropped and replaced (i.e. adaptive designs are incorporated into the platform). Conversely, interventions with a good response are taken forward into a more definitive test with a hard clinical outcome, using the early phase as a vanguard (i.e. an enrichment design is incorporated into the platform).<sup>92</sup> This approach is more flexible than a traditional 1:1 parallel design as new drugs can be added to the platform rapidly.<sup>93</sup> Platform trials hold the potential to accelerate the early phase of drug development and the concept is being developed for application to renal trials by the International Society of Nephrology Advancing Clinical Trials (ISN-ACT) initiative.<sup>92</sup>

A potential disadvantage of any trial design that drops participants who fail to show a substantial response to a surrogate marker in the early phase (despite good adherence), is that such patients might not be studied in definitive trials. When surrogate markers are imperfect (as is common in renal disease),<sup>46</sup> this prejudice leaves a group of unstudied patients who are at risk of disease progression and could potentially benefit from the effect of the intervention (Table 1). A more pragmatic population-based approach would be to design a simpler, larger trial with an active run-in to assess responder status (analogous to HPS<sup>70</sup> and ADVANCE<sup>72</sup>) and sufficient power to identify a more modest treatment effect. A subgroup analysis by different levels of pre-randomization responder status can then be pre-specified once all types of patient at risk have been included and studied.

### **[H1] Unscientific and burdensome regulation**

One of the justifications for the necessity of platform trial designs is the high cost of trials. Unscientific and overly burdensome regulation of randomized trials has

caused increased organizational and operational complexity (Table 2). This approach is leading to an unsustainable increase in the cost of trials to the detriment of patients. Complex and often unscientific research governance guidance and regulation (and its interpretation), including the International Conference on Harmonisation's Good Clinical Practice (ICH-GCP) guidelines, are considered to be partly responsible.<sup>1,94</sup> For example, the workload generated by demanding pharmacovigilance requirements may contribute to reluctance of clinicians to participate in trials. Those running trials are also burdened by undue focus on complying with poorly conceived rules and regulations rather than innovating trial design and conduct to recruit larger numbers of participants.

To address these major concerns, the MoreTrials campaign has joined forces with major funders to develop new more straightforward and scientific trial guidelines.<sup>95</sup> These guidelines will acknowledge the strengths of randomization whilst still remaining focused on the fundamentals of ensuring no material impact on the rights, safety and well-being of trial participants or on the reliability of results, which in turn affects the safety of future patients. If adopted, simpler regulation and guidance could remove one of the major challenges that is faced by those who are trying to conduct trials in nephrology.<sup>33</sup> The work of the MoreTrials campaign complements that of the Clinical Trial Transformation Initiative,<sup>96</sup> a multi-stakeholder partnership with the US FDA which aims to develop and drive adoption of practices that will increase the quality and efficiency of clinical trials. An FDA initiative with the American Society of Nephrology called the Kidney Health Initiative is also aiming to forge innovation in kidney care, including the development of pragmatic trials.<sup>97</sup>

## **[H1] Conclusions**

Randomization and adequately large sample sizes are vital to ensure that trials provide a reliable test of an intervention and produce results that could potentially change clinical practice. The adoption of such trials and a population-based approach has led to many advances in cardiology. Nephrology lags behind cardiology and many other specialties in performing high-quality, large-scale definitive clinical trials. In this era of complex research governance and burdensome regulation, innovations in trial conduct to enable large-scale invitation, better

adherence and low-cost follow-up may be more important than innovations in trial design. Nephrology is well positioned to harness such innovations due to its advanced use of electronic healthcare records and the development of disease-specific registries. Through more widespread adoption of the uncertainty principle, innovations in trial conduct and challenging burdensome and unscientific regulation, the nephrology community could benefit from more reliable answers to relevant clinical questions. Such answers could lead to improvements in the care of current and future patients with kidney disease.

### **Acknowledgements**

The Medical Research Council Population Health Research Unit at the University of Oxford is funded by the UK Medical Research Council. WGH is funded by a Medical Research Council-Kidney Research UK Professor David Kerr Clinician Scientist Award.

### **Author contributions**

All authors researched the data for the article, contributed to discussions of the content, wrote the text and reviewed or edited the manuscript before submission.

### **Competing interests statement**

WGH, RH and NS report grants to the Renal Studies Group at the University of Oxford from Novartis, Boehringer Ingelheim, and Pfizer. CTSU has a staff policy of not accepting honoraria or other payments from the pharmaceutical industry, except for the reimbursement of costs to participate in scientific meetings ([www.ctsu.ox.ac.uk](http://www.ctsu.ox.ac.uk)).

**Table 1: Comparison of population and precision medicine based approaches to randomized trials**

	<b>Population approach</b>	<b>Precision medicine approach</b>
<b>Rationale</b>	<ul style="list-style-type: none"> <li>Differences in relative treatment effects are uncommon and may not be clinically important</li> </ul>	<ul style="list-style-type: none"> <li>Differences in relative treatment effects may exist</li> </ul>
<b>Design</b>	<ul style="list-style-type: none"> <li>Recruit a population at high risk of the primary outcome</li> <li>Pre-specify subgroup analyses by characteristics which may modify treatment effects</li> </ul>	<ul style="list-style-type: none"> <li>Recruit a population that has been demonstrated to respond to an intermediate marker or have a particular characteristic (e.g. a genotype)</li> </ul>
<b>Examples</b>	<ul style="list-style-type: none"> <li>Traditional large, simple parallel-group trials</li> </ul>	<ul style="list-style-type: none"> <li>Parallel-group trials selecting for specific characteristics</li> <li>Parallel-group trials with pre-randomization enrichment phases</li> <li>Platform trials</li> </ul>
<b>Advantages</b>	<ul style="list-style-type: none"> <li>No groups are excluded (a single definitive trial may suffice)</li> <li>Simple recruitment process</li> <li>Widely generalizable results</li> <li>Opportunity to streamline trial design</li> </ul>	<ul style="list-style-type: none"> <li>Larger than anticipated treatment effects could potentially enable smaller or shorter trials</li> </ul>
<b>Disadvantages</b>	<ul style="list-style-type: none"> <li>May require a larger trial population</li> </ul>	<ul style="list-style-type: none"> <li>Excludes at-risk groups who may benefit from the study intervention</li> <li>Complicated, time-consuming and costly recruitment procedures</li> <li>A positive trial may lead to complicated indication statements on drug labels</li> <li>May require a second definitive trial to study patients who are presumed to be modest responders</li> </ul>



**Table 2: Examples of unscientific and burdensome practices and alternative approaches**

<b>Problematic practice</b>	<b>Justification</b>	<b>Problem</b>	<b>Alternative approach</b>
Decentralized (i.e. local) institutional review board and/or ethics committee review	Specific local ethical or legal issues might exist	Large trials may have >100 sites. Obtaining approval for the same study at each individual site is very time-consuming and expensive without offering additional protection of the rights and well-being of participants	Adoption of a rigorous country-level institutional review board or ethics committee review (as is well developed in the UK) could substantially reduce study start-up and amendment time.
Requirement for 100% source data verification	Confirmation of participant-reported medical history from medical records is required to confirm phenotypic information	Very time-consuming and expensive and generates large numbers of data queries that disincentivise the trial site staff. Resolution of such queries is unlikely to improve the reliability of the trial result.	Rely on the participant as a source of data; any errors will mostly be small and will not affect the reliability of the treatment effect estimate, provided the trial is double-blind and sufficiently large.
Expedited reporting of SUSARs	Desire to rapidly identify adverse effects of treatment	Only report SUSARs for participants assigned active therapy so no control arm.  Unable to reliably identify harms of treatment except when event is very rare in general population and strongly associated with drug exposure (e.g. Stevens-Johnson syndrome).	Provide regular unblinded analyses of all adverse event data, in confidence, to an independent Data Monitoring Committee.
Re-consent of participants for any new safety information	Participant must be informed of all new information (no matter how significant or relevant)	Very expensive and time-consuming for site staff.	Appropriate risk assessment of new information. Important information can be conveyed without requiring re-consent.

Frequent on-site monitoring visits	Required to manage source data verification	A major contributor to the large cost of many trials.	Focus on processes that impact trial quality, for example the consent process.
------------------------------------	---	---	--

SUSAR, suspected unexpected serious adverse reaction.

**Figure 1 | Avoiding bias in responder analyses. a** | Responder analyses based on post-randomization responses to an intermediate marker of treatment effect can introduce bias during the analysis. As the response is only measured in participants who are randomly assigned to active treatment, no randomized control group of similar participants is available for comparison. **b** | This bias can be avoided using an active-treatment run-in period during which the response to treatment is defined in all participants before randomization. Responders in the active treatment group can then be compared to responders in the appropriate control group. R, randomization.

**Figure 2 | Parallel group randomized trial designs. a** | Parallel group. Individuals are screened and may enter a pre-randomization run-in phase. They are then randomly assigned to the test intervention, often including a placebo group. Comparison of different doses of an active treatment versus placebo requires additional study groups. The simplicity of a 2-group design makes analysis and interpretation of results relatively straightforward. **b** | Factorial. A factorial design incorporating multiple randomizations can be added to a traditional parallel group design so multiple interventions can be tested in a single trial (a 2x2 factorial design is depicted). The advantage of this design is that more than one primary question is tested within a single protocol. However, participants need to take two or more test drugs and an adverse effect of one drug may lead to discontinuation of both drugs.

**Figure 3 | Cluster randomized trial designs. a** | Cluster. Groups of participants (for example all those at a participating kidney unit) rather than individuals are randomly allocated to study interventions. This design has the advantage of simplified recruitment procedures. However, sample sizes need to be larger than those that are required for a similarly powered traditional parallel group trial. A reduced sense of participation resulting from reduced participant contact by the research team may adversely affect adherence to the study intervention. **b** | Step wedge cluster. Individuals, or groups of individuals in a combined step wedge cluster design, are randomly selected to start the study intervention at regular intervals. Initially no cluster but eventually all clusters receive the intervention. Step-wedge cluster designs are particularly suited to evaluate service delivery and policy interventions during their implementation (for example, placebo and drug A could be replaced by control and policy intervention). The disadvantage of this design is that statistical

consideration is needed to account for the clustered nature of the design and the confounding effect of time.

**Figure 4 | Master protocol trial designs.** Use of a master protocol trial design enables the testing of more than one intervention in a single disease or the testing of a single intervention in more than one disease in a single trial. These trial designs are largely untested in nephrology, therefore, their value in this field remains unclear. **[Au: suggested addition OK?] a | Basket.** **[Au: Please add a few sentences to describe this trial design.] b | Umbrella.** When there is a strong rationale that interventions require a certain disease phenotype to be effective (e.g. tumours expressing a certain mutation), participants are randomly assigned to differing disease-directed treatments based on disease phenotyping prior to randomization. This design has the advantages and disadvantages of a precision medicine approach. **c | Platform.** A platform design provides the opportunity to assess multiple interventions (or doses) versus a shared placebo or control group. Such trials usually use a surrogate marker of effect. Interventions with poor responses are dropped and replaced (i.e. adaptive designs are incorporated into the platform). Conversely safe interventions with a good response are taken forward into a definitive outcome trial. This design is more flexible than a traditional 1:1 parallel design as new drugs can be added to the platform rapidly.

### ***Box 1: Estimating the absolute benefit of a treatment for an individual patient***

#### **Step 1**

Estimate the absolute risk of the event using an established risk score

#### **Step 2**

Identify the relative risk reduction with the intervention using the best-available randomized trial data or preferably data from a meta-analysis of large trials

#### **Step 3**

Apply the trial or meta-analysis relative risk estimate to the absolute risk estimate

#### **Step 4**

Discuss the results with the patient

### ***Example***

What is the vascular benefit of reducing LDL-cholesterol by 1 mmol/l with a statin in this patient?

- *The QRISK score estimates a 10-year risk of major vascular events for the patient of 10%*
- *The Cholesterol Treatment Trialists' collaboration meta-analysis reports a ~20% relative reduction in the risk of major vascular events per 1 mmol/l lower LDL-C<sup>22</sup>*
- *$10 \times 0.8 = 8\%$  10-year risk of major vascular event*
- *Taking a statin would reduce the patient's 10-year risk of a major vascular event from 10% to 8%*

## **Box 2 | Potential benefits of pre-screening**

### **For patients**

- Opportunity to volunteer empowers patients to join trials
- Provides extra time for considering and discussing participation

### **For local sites**

- Enables determination of whether sufficient numbers of potential participants exist locally to meet recruitment targets
- Reduces the number of wasted screening visits
- Enables advanced planning of a large number of screening visits
- Disentangles the tasks of identification and invitation of potential participants from the task of seeing participants at study visits

### **For trialists**

- Provides an early assessment of the likely composition of the future trial population
- Helps identify which exclusion criteria might most commonly apply
- Provides an early assessment of the engagement and size of a site
- Compresses the recruitment period

### **Box 3: Potential benefits of pre-randomization run-in periods**

#### **For patients**

- Provides extra time for considering and discussing participation

#### **For trialists**

- Provides an opportunity to exclude participants who are likely to be non-adherent to the study intervention before randomization, which helps to maintain study power
- Provides local clinicians with opportunities to review results from screening visits and confirm (or refute) uncertainty
- Provides time to establish standard of care when necessary
- Can be used to assess early effects on intermediate markers either to assess level of response prior to randomization for subgroup analyses (population approach) or to enrich the trial population for high responders (precision medicine approach)

#### **References**

- 1 Collins, R. Back to the future: the urgent need to re-introduce streamlined trials. *European Heart Journal Supplements* **20** (Supplement C), C14–C17 (2018).
- 2 European Society of Cardiology Clinical Practice Guidelines. <https://www.escardio.org/Guidelines/Clinical-Practice-Guidelines> (accessed 17th March 2019).
- 3 Nabel, E. G. & Braunwald, E. A tale of coronary artery disease and myocardial infarction. *N Engl J Med* **366**, 54–63, doi:10.1056/NEJMra1112570 (2012).
- 4 Herrington, W., Lacey, B., Sherliker, P., Armitage, J. & Lewington, S. Epidemiology of Atherosclerosis and the Potential to Reduce the Global Burden of Atherothrombotic Disease. *Circ Res* **118**, 535–546, doi:10.1161/CIRCRESAHA.115.307611 (2016).
- 5 Wanner, C. *et al.* Empagliflozin and Progression of Kidney Disease in Type 2 Diabetes. *N Engl J Med* **375**, 323–334, doi:10.1056/NEJMoa1515920 (2016).
- 6 Perkovic, V. *et al.* Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS Program randomised clinical trials. *Lancet Diabetes Endocrinol* **6**, 691–704, doi:10.1016/S2213-8587(18)30141-4 (2018).
- 7 Zelniker, T. A. *et al.* Comparison of the Effects of Glucagon-Like Peptide Receptor Agonists and Sodium-Glucose Co-Transporter 2 Inhibitors for Prevention of Major Adverse Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis of Cardiovascular Outcomes Trials. *Circulation*, doi:10.1161/CIRCULATIONAHA.118.038868 (2019).
- 8 Strippoli, G. F., Craig, J. C. & Schena, F. P. The number, quality, and coverage of randomized controlled trials in nephrology. *J Am Soc Nephrol* **15**, 411–419 (2004).

- 9 Inrig, J. K. *et al.* The landscape of clinical trials in nephrology: a systematic review of Clinicaltrials.gov. *Am J Kidney Dis* **63**, 771-780, doi:10.1053/j.ajkd.2013.10.043 (2014).
- 10 Tonelli, M. *et al.* Comparison of the Complexity of Patients Seen by Different Medical Subspecialists in a Universal Health Care System. *JAMA Netw Open* **1**, e184852, doi:10.1001/jamanetworkopen.2018.4852 (2018).
- 11 Kerr, M., Bray, B., Medcalf, J., O'Donoghue, D. J. & Matthews, B. Estimating the financial cost of chronic kidney disease to the NHS in England. *Nephrol Dial Transplant* **27 Suppl 3**, iii73-80, doi:10.1093/ndt/gfs269 (2012).
- 12 Standardised outcomes in nephrology (SONG). <http://songinitiative.org/> (accessed 13th July 2019).
- 13 Deeks, J. J. *et al.* Evaluating non-randomised intervention studies. *Health Technol Assess* **7**, iii-x, 1-173 (2003).
- 14 Gerstein, H. C., McMurray, J. & Holman, R. R. Real-world studies no substitute for RCTs in establishing efficacy. *Lancet* **393**, 210-211, doi:10.1016/S0140-6736(18)32840-X (2019).
- 15 Grady, D. *et al.* Hormone therapy to prevent disease and prolong life in postmenopausal women. *Ann Intern Med* **117**, 1016-1037 (1992).
- 16 Rossouw, J. E. *et al.* Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA : the journal of the American Medical Association* **288**, 321-333 (2002).
- 17 Collins, R. *et al.* Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* **388**, 2532-2561, doi:10.1016/S0140-6736(16)31357-5 (2016).
- 18 Bennett, D. A. & Holmes, M. V. Mendelian randomisation in cardiovascular research: an introduction for clinicians. *Heart* **103**, 1400-1407, doi:10.1136/heartjnl-2016-310605 (2017).
- 19 Sun, L. *et al.* Causal associations of blood lipids with risk of ischemic stroke and intracerebral hemorrhage in Chinese adults. *Nat Med*, doi:10.1038/s41591-019-0366-x (2019).
- 20 Holmes, M. V. Human Genetics and Drug Development. *N Engl J Med* **380**, 1076-1079, doi:10.1056/NEJMe1901565 (2019).
- 21 Collins, R. & MacMahon, S. Reliable assessment of the effects of treatment on mortality and major morbidity, I: clinical trials. *Lancet* **357**, 373-380, doi:10.1016/S0140-6736(00)03651-5 (2001).
- 22 Pitt, B. *et al.* The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone Evaluation Study Investigators. *N Engl J Med* **341**, 709-717, doi:10.1056/NEJM199909023411001 (1999).
- 23 Perkovic, V. *et al.* Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. *N Engl J Med* **380**, 2295-2306, doi:10.1056/NEJMoa1811744 (2019).
- 24 Ettehad, D. *et al.* Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet* **387**, 957-967, doi:10.1016/S0140-6736(15)01225-8 (2016).
- 25 Xie, X. *et al.* Effects of intensive blood pressure lowering on cardiovascular and renal outcomes: updated systematic review and meta-analysis. *Lancet* **387**, 435-443, doi:10.1016/S0140-6736(15)00805-3 (2016).
- 26 Herrington W, Emberson J, Mihaylova B, et al. How are the effects of reducing LDL cholesterol with a statin-based regime influenced by renal function? Meta-analysis of individual data from 28 randomised trials. *Lancet Diabetes Endocrinol* **4**, 829-839 (2016).



- 27 Jafar, T. H. *et al.* Angiotensin-converting enzyme inhibitors and progression of nondiabetic renal disease. A meta-analysis of patient-level data. *Ann Intern Med* **135**, 73-87, doi:200107170-00007 [pii] (2001).
- 28 Webster, A. C., Playford, E. G., Higgins, G., Chapman, J. R. & Craig, J. C. Interleukin 2 receptor antagonists for renal transplant recipients: a meta-analysis of randomized trials. *Transplantation* **77**, 166-176, doi:10.1097/01.TP.0000109643.32659.C4 (2004).
- 29 Egger, M. & Smith, G. D. Misleading meta-analysis. *BMJ* **310**, 752-754 (1995).
- 30 Chertow, G. M. *et al.* Effect of cinacalcet on cardiovascular disease in patients undergoing dialysis. *N Engl J Med* **367**, 2482-2494, doi:10.1056/NEJMoa1205624 (2012).
- 31 Parfrey, P. S. *et al.* Lessons Learned from EVOLVE for Planning of Future Randomized Trials in Patients on Dialysis. *Clin J Am Soc Nephrol* **11**, 539-546, doi:10.2215/CJN.06370615 (2016).
- 32 Haynes, R., Landray, M.J., Herrington, W.G., Baigent, C. in *Oxford Textbook of Clinical Nephrology (4th edition)* (2015).
- 33 Baigent, C. *et al.* Challenges in conducting clinical trials in nephrology: conclusions from a Kidney Disease-Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney international* **92**, 297-305, doi:10.1016/j.kint.2017.04.019 (2017).
- 34 Phrommintikul, A., Haas, S. J., Elsik, M. & Krum, H. Mortality and target haemoglobin concentrations in anaemic patients with chronic kidney disease treated with erythropoietin: a meta-analysis. *Lancet* **369**, 381-388, doi:10.1016/S0140-6736(07)60194-9 (2007).
- 35 Pfeffer, M. A. *et al.* A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease. *N Engl J Med* **361**, 2019-2032, doi:NEJMoa0907845 [pii] 10.1056/NEJMoa0907845 (2009).
- 36 European Renal Registry Annual Report 2016. <https://www.era-edta-reg.org/files/annualreports/pdf/AnnRep2016.pdf> (accessed 12th July 2019).
- 37 *The National Registry of Rare Kidney Disease (RaDaR)*. <https://www.renalreg.org/projects/the-national-registry-of-rare-kidney-disease-radar/> (accessed 12th July 2019)
- 38 Moxey-Mims, M. M. *et al.* Glomerular Diseases: Registries and Clinical Trials. *Clin J Am Soc Nephrol* **11**, 2234-2243, doi:10.2215/CJN.00540116 (2016).
- 39 Miner, J. H. *et al.* The 2014 International Workshop on Alport Syndrome. *Kidney Int* **86**, 679-684, doi:10.1038/ki.2014.229 (2014).
- 40 Bowman, L. *et al.* ASCEND: A Study of Cardiovascular Events iN Diabetes: Characteristics of a randomized trial of aspirin and of omega-3 fatty acid supplementation in 15,480 people with diabetes. *Am Heart J* **198**, 135-144, doi:10.1016/j.ahj.2017.12.006 (2018).
- 41 Group, A. S. C. *et al.* Effects of Aspirin for Primary Prevention in Persons with Diabetes Mellitus. *N Engl J Med* **379**, 1529-1539, doi:10.1056/NEJMoa1804988 (2018).
- 42 Haynes, R. *et al.* Effects of Sacubitril/Valsartan Versus Irbesartan in Patients With Chronic Kidney Disease. *Circulation* **138**, 1505-1514, doi:10.1161/CIRCULATIONAHA.118.034818 (2018).
- 43 Sharp Collaborative Group. Study of Heart and Renal Protection (SHARP): randomized trial to assess the effects of lowering low-density lipoprotein cholesterol among 9,438 patients with chronic kidney disease. *Am Heart J* **160**, 785-794.e710, doi:S0002-8703(10)00759-3 [pii]

10.1016/j.ahj.2010.08.012 (2010).

44 Lang, J. M., Buring, J. E., Rosner, B., Cook, N. & Hennekens, C. H. Estimating the effect of the run-in on the power of the Physicians' Health Study. *Stat Med* **10**, 1585-1593 (1991).

45 Landray, M. J. *et al.* Effects of extended-release niacin with laropiprant in high-risk patients. *N Engl J Med* **371**, 203-212, doi:10.1056/NEJMoa1300955 (2014).

46 Heerspink, H. J. L. *et al.* Atrasentan and renal events in patients with type 2 diabetes and chronic kidney disease (SONAR): a double-blind, randomised, placebo-controlled trial. *Lancet* **393**, 1937-1947, doi:10.1016/S0140-6736(19)30772-X (2019).

47 Symplicity, H. T. N. I. *et al.* Renal sympathetic denervation in patients with treatment-resistant hypertension (The Symplicity HTN-2 Trial): a randomised controlled trial. *Lancet* **376**, 1903-1909, doi:10.1016/S0140-6736(10)62039-9 (2010).

48 Bhatt, D. L. *et al.* A controlled trial of renal denervation for resistant hypertension. *N Engl J Med* **370**, 1393-1401, doi:10.1056/NEJMoa1402670 (2014).

49 Howard, J. P. *et al.* Quantifying the 3 Biases That Lead to Unintentional Overestimation of the Blood Pressure-Lowering Effect of Renal Denervation. *Circ Cardiovasc Qual Outcomes* **9**, 14-22, doi:10.1161/CIRCOUTCOMES.115.002533 (2016).

50 Meeker-O'Connell, A. *et al.* Enhancing clinical evidence by proactively building quality into clinical trials. *Clinical trials* **13**, 439-444, doi:10.1177/1740774516643491 (2016).

51 Pogue, J. M., Devereaux, P. J., Thorlund, K. & Yusuf, S. Central statistical monitoring: detecting fraud in clinical trials. *Clin Trials* **10**, 225-235, doi:10.1177/1740774512469312 (2013).

52 Group, C. S. C. Campath, calcineurin inhibitor reduction, and chronic allograft nephropathy (the 3C Study) - results of a randomized controlled clinical trial. *Am J Transplant* **18**, 1424-1434, doi:10.1111/ajt.14619 (2018).

53 Ndounga Diakou, L. A. *et al.* Comparison of central adjudication of outcomes and onsite outcome assessment on treatment effect estimates. *Cochrane Database Syst Rev* **3**, MR000043, doi:10.1002/14651858.MR000043.pub2 (2016).

54 Hlatky, M. A. *et al.* Use of Medicare data to identify coronary heart disease outcomes in the Women's Health Initiative. *Circ Cardiovasc Qual Outcomes* **7**, 157-162, doi:10.1161/CIRCOUTCOMES.113.000373 (2014).

55 Frobert, O. *et al.* Thrombus aspiration during ST-segment elevation myocardial infarction. *N Engl J Med* **369**, 1587-1597, doi:10.1056/NEJMoa1308789 (2013).

56 Prentice, R. L. Surrogate endpoints in clinical trials: Definition and operational criteria. *Statistics in medicine* **8**, 431-440, doi:10.1002/sim.4780080407 (1989).

57 Levey, A. S. *et al.* GFR Decline as an End Point for Clinical Trials in CKD: A Scientific Workshop Sponsored by the National Kidney Foundation and the US Food and Drug Administration. *American Journal of Kidney Diseases* **64**, 821-835, doi:<https://doi.org/10.1053/j.ajkd.2014.07.030> (2014).

58 Fried, L. F. & Lewis, J. Albuminuria is Not an Appropriate Therapeutic Target in Patients with CKD: The Con View. *Clinical Journal of the American Society of Nephrology* **10**, 1089-1093, doi:10.2215/cjn.10681014 (2015).

59 Lambers Heerspink, H. J. & Gansevoort, R. T. Albuminuria Is an Appropriate Therapeutic Target in Patients with CKD: The Pro View. *Clinical Journal of the American Society of Nephrology* **10**, 1079-1088, doi:10.2215/cjn.11511114 (2015).

60 Heerspink, H. J. L. *et al.* Change in albuminuria as a surrogate endpoint for progression of kidney disease: a meta-analysis of treatment effects in randomised clinical trials. *The Lancet Diabetes & Endocrinology* **7**, 128-139, doi:10.1016/s2213-8587(18)30314-0 (2019).

- 61 Smith, M. *et al.* Change in Albuminuria and Risk of Renal and Cardiovascular Outcomes: Natural Variation Should Be Taken into Account. *Kidney Int Rep* **3**, 939-949, doi:10.1016/j.ekir.2018.04.004 (2018).
- 62 Thompson, A. *et al.* Proteinuria Reduction as a Surrogate End Point in Trials of IgA Nephropathy. *Clinical Journal of the American Society of Nephrology* **14**, 469-481, doi:10.2215/cjn.08600718 (2019).
- 63 Schluchter, M. D., Greene, T. & Beck, G. J. Analysis of change in the presence of informative censoring: application to a longitudinal clinical trial of progressive renal disease. *Statistics in medicine* **20**, 989-1007, doi:10.1002/sim.720 (2001).
- 64 Vonesh, E. F., Greene, T. & Schluchter, M. D. Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in medicine* **25**, 143-163, doi:10.1002/sim.2249 (2006).
- 65 Kular, D. & Macdougall, I. C. HIF stabilizers in the management of renal anemia: from bench to bedside to pediatrics. *Pediatr Nephrol* **34**, 365-378, doi:10.1007/s00467-017-3849-3 (2019).
- 66 McMurray, J. J. *et al.* Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N Engl J Med* **371**, 993-1004, doi:10.1056/NEJMoa1409077 (2014).
- 67 Deo, A., Schmid, C. H., Earley, A., Lau, J. & Uhlig, K. Loss to analysis in randomized controlled trials in CKD. *Am J Kidney Dis* **58**, 349-355, doi:10.1053/j.ajkd.2011.04.023 (2011).
- 68 Coronary Drug Project Research, G. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *N Engl J Med* **303**, 1038-1041, doi:10.1056/NEJM198010303031804 (1980).
- 69 Eijkelkamp, W. B. *et al.* Albuminuria is a target for renoprotective therapy independent from blood pressure in patients with type 2 diabetic nephropathy: post hoc analysis from the Reduction of Endpoints in NIDDM with the Angiotensin II Antagonist Losartan (RENAAL) trial. *J Am Soc Nephrol* **18**, 1540-1546, doi:10.1681/ASN.2006050445 (2007).
- 70 Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* **360**, 7-22, doi:S0140-6736(02)09327-3 [pii] 10.1016/S0140-6736(02)09327-3 (2002).
- 71 Schmidt, M. *et al.* Serum creatinine elevation after renin-angiotensin system blockade and long term cardiorenal risks: cohort study. *BMJ* **356**, j791, doi:10.1136/bmj.j791 (2017).
- 72 Ohkuma, T. *et al.* Acute Increases in Serum Creatinine After Starting Angiotensin-Converting Enzyme Inhibitor-Based Therapy and Effects of its Continuation on Major Clinical Outcomes in Type 2 Diabetes Mellitus. *Hypertension* **73**, 84-91, doi:10.1161/HYPERTENSIONAHA.118.12060 (2019).
- 73 Vickers, A. J. & Altman, D. G. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ* **323**, 1123-1124, doi:10.1136/bmj.323.7321.1123 (2001).
- 74 Dember, L. M. *et al.* Pragmatic Trials in Maintenance Dialysis: Perspectives from the Kidney Health Initiative. *J Am Soc Nephrol* **27**, 2955-2963, doi:10.1681/ASN.2016030340 (2016).
- 75 Hemming, K., Eldridge, S., Forbes, G., Weijer, C. & Taljaard, M. How to design efficient cluster randomised trials. *BMJ* **358**, j3064, doi:10.1136/bmj.j3064 (2017).
- 76 Torgerson, D. J. Contamination in trials: is cluster randomisation the answer? *BMJ* **322**, 355-357, doi:10.1136/bmj.322.7282.355 (2001).

- 77 Eldridge, S., Kerry, S. & Torgerson, D. J. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? *BMJ* **339**, b4006, doi:10.1136/bmj.b4006 (2009).
- 78 Dember, L. M. *et al.* The TiME Trial: A Fully Embedded, Cluster-Randomized, Pragmatic Trial of Hemodialysis Session Duration. *J Am Soc Nephrol* **30**, 890-903, doi:10.1681/ASN.2018090945 (2019).
- 79 Hemming, K., Haines, T. P., Chilton, P. J., Girling, A. J. & Lilford, R. J. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* **350**, h391, doi:10.1136/bmj.h391 (2015).
- 80 Selby, N. M. *et al.* An Organizational-Level Program of Intervention for AKI: A Pragmatic Stepped Wedge Cluster Randomized Trial. *Journal of the American Society of Nephrology* **30**, 505-515, doi:10.1681/asn.2018090886 (2019).
- 81 Feldman, A. M. *et al.* Precision Medicine for Heart Failure: Lessons From Oncology. *Circ Heart Fail* **10**, doi:10.1161/CIRCHEARTFAILURE.117.004202 (2017).
- 82 Baigent, C. *et al.* Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. *Lancet* **376**, 1670-1681, doi:S0140-6736(10)61350-5 [pii] 10.1016/S0140-6736(10)61350-5 (2010).
- 83 Ruggenenti, P., Bettinaglio, P., Pinares, F. & Remuzzi, G. Angiotensin converting enzyme insertion/deletion polymorphism and renoprotection in diabetic and nondiabetic nephropathies. *Clin J Am Soc Nephrol* **3**, 1511-1525, doi:10.2215/CJN.04140907 (2008).
- 84 de Zeeuw, D., Heerspink, H. J. L., Jardine, M. & Perkovic, V. Renal trials in diabetes need a platform: time for a global approach? *Lancet Diabetes Endocrinol* **6**, 356-358, doi:10.1016/S2213-8587(17)30263-2 (2018).
- 85 Heerspink, H. J. L. *et al.* Rationale and protocol of the Study Of diabetic Nephropathy with AtRasentan (SONAR) trial: A clinical trial design novel to diabetic nephropathy. *Diabetes Obes Metab* **20**, 1369-1376, doi:10.1111/dom.13245 (2018).
- 86 Park, J. J., Thorlund, K. & Mills, E. J. Critical concepts in adaptive clinical trials. *Clin Epidemiol* **10**, 343-351, doi:10.2147/CLEP.S156708 (2018).
- 87 Thorlund, K., Haggstrom, J., Park, J. J. & Mills, E. J. Key design considerations for adaptive clinical trials: a primer for clinicians. *BMJ* **360**, k698, doi:10.1136/bmj.k698 (2018).
- 88 Mann, J. F. *et al.* Avosentan for overt diabetic nephropathy. *J Am Soc Nephrol* **21**, 527-535, doi:10.1681/ASN.2009060593 (2010).
- 89 Hirakawa, A., Asano, J., Sato, H. & Teramukai, S. Master protocol trials in oncology: Review and new trial designs. *Contemp Clin Trials Commun* **12**, 1-8, doi:10.1016/j.conctc.2018.08.009 (2018).
- 90 Renfro, L. A. & Sargent, D. J. Statistical controversies in clinical research: basket trials, umbrella trials, and other master protocols: a review and examples. *Ann Oncol* **28**, 34-43, doi:10.1093/annonc/mdw413 (2017).
- 91 Saville, B. R. & Berry, S. M. Efficiencies of platform clinical trials: A vision of the future. *Clin Trials* **13**, 358-366, doi:10.1177/1740774515626362 (2016).
- 92 Heerspink, H. J. L. & Perkovic, V. Trial Design Innovations to Accelerate Therapeutic Advances in Chronic Kidney Disease: Moving from Single Trials to an Ongoing Platform. *Clin J Am Soc Nephrol* **13**, 946-948, doi:10.2215/CJN.01290118 (2018).
- 93 Butler, C. C. *et al.* A trial like ALIC(4)E: why design a platform, response-adaptive, open, randomised controlled trial of antivirals for influenza-like illness? *ERJ Open Res* **4**, doi:10.1183/23120541.00046-2018 (2018).
- 94 Landray, M. Improving public health by improving clinical trial guidelines and their application. *European Heart Journal* (2017).
- 95 Moretrials (to do more trials better). <https://moretrials.net/major-funders-establish-initiative-develop-new-trial-guidelines/> (accessed 17th March 2019).

96 Clinical Trial Transformation Initiative. Who We Are. <https://www.ctti-clinicaltrials.org/who-we-are> (accessed 13th July 2019).

97 Linde, P. G. *et al.* Overcoming Barriers in Kidney Health-Forging a Platform for Innovation. *J Am Soc Nephrol* 27, 1902-1910, doi:10.1681/ASN.2015090976 (2016).

## **Glossary terms**

### **Propensity-score matching**

A statistical matching technique employed in observational studies to reduce the potential for bias by factors that predict use of an intervention rather than the effect of the intervention. Elimination of such biases can only be ensured by randomization in a trial.

### **Mendelian randomization (MR) methods**

Methods of analyzing observational study data in which allele variation in genes encoding risk factors or biomarkers are used to infer whether a risk factor may be a cause of disease. Such design requires assumptions but may control for biases (e.g. reverse causation or confounding) that are inherent in classical observational studies.

### **Statistical tests for heterogeneity**

In trials, this term refers to a statistical test of whether effects in a particular subgroup of participants differ significantly from the overall effect (e.g. a chi-squared test).

### **Linear mixed models**

Extension of simple linear models to include random effects (e.g. patient specific intercepts and/or slopes) to account for dependence between data points (such as multiple eGFR measurements from same patient).

### **Shared parameter models**

Statistical method that jointly models longitudinal data (using linear mixed models) and a censoring event (using a survival model) to enable unbiased estimation of a longitudinal outcome (e.g. rate of decline in eGFR) in the presence of a non-ignorable drop-out mechanism (i.e. patients are dropped from analyses for reasons related to their decline, such as end-stage kidney disease or death).