Design of Clinical Trials in Acute Kidney Injury: Report from an NIDDK Workshop on Trial Methodology


Summary
Acute kidney injury (AKI) remains a complex clinical problem associated with significant short-term morbidity and mortality and lacking effective pharmacologic interventions. Patients with AKI experience longer-term risks for progressive chronic ESRD, which diminish patients’ health-related quality of life and create a larger burden on the healthcare system. Although experimental models have yielded numerous promising agents, translation into practice has been unsuccessful, possibly because of issues in clinical trial design, such as delayed drug administration, masking of therapeutic benefit by adverse events, and inadequate sample size. To address issues of clinical trial design, the National Institute of Diabetes and Digestive and Kidney Diseases sponsored a workshop titled “Clinical Trials in Acute Kidney Injury: Current Opportunities and Barriers” in December 2010. Workshop participants included representatives from academia, industry, and government agencies whose areas of expertise spanned basic science, clinical nephrology, critical care medicine, biostatistics, pharmacology, and drug development. This document summarizes the discussions of collaborative workgroups that addressed issues related to patient selection, study endpoints, the role of novel biomarkers, sample size and power calculations, and adverse events and pilot/feasibility studies in prevention and treatment of AKI. Companion articles outline the discussions of workgroups for model trials related to prevention or treatment of established AKI in different clinical settings, such as in patients with sepsis.

Patient Selection in AKI—Inclusion/Exclusion Criteria
Moderators: Dr. Adeera Levin and Dr. Sushrut S. Waikar

The appropriate enrollment criteria for clinical trials of a drug or device targeting prevention or treatment of acute kidney injury (AKI) depend on multiple factors, including phase of drug or device development, intended population, mechanism of action and toxicity potential of the intervention, and logistic considerations regarding recruitment and expected event rates.

To enrich event rates, studies of AKI prevention or treatment must identify and recruit high-risk individuals on the basis of demographic data, comorbidity assessment, and clinical measures. Key considerations are summarized in Table 1. Commonly identified variables in risk scores include baseline CKD, diabetes mellitus, age, congestive heart failure, and liver disease. Although there is a rationale for restricting studies to patients in the highest risk strata, that impulse must be balanced by the possibility of decreased therapeutic effect because those in the highest risk strata may also be biologic nonresponders. Inclusion of procedural variables in assessment of AKI risk, such as duration of cardiopulmonary bypass during cardiac surgery or the volume or type of iodinated contrast agent administration, can pose logistic issues because such data will not be available when patients are recruited before the procedure.

Pre-existing CKD is the most important risk factor for subsequent development of AKI (1). Underlying CKD potentially alters the natural history of AKI, its pathophysiologic mechanisms and drug targets, and the biologic response to interventions. Although the associations of elevated serum creatinine concentration (SCr) and reduced estimated GFR (eGFR) with AKI have been best characterized, proteinuria (with or without albuminuria) is also an important risk factor for development of AKI (2,3). Although inclusion of individuals with CKD may increase the rates of AKI in intervention or prevention trials, it may also diminish response rates. The inclusion of patients with moderate to advanced CKD may be counterproductive because of the potential for differing disease mechanisms or a decreased biologic response as well as diminished generalizability to patients with normal or near-normal baseline kidney function. Recruitment of patients with an eGFR of 15 to 45 ml/min per 1.73 m² or an eGFR of 46 to 60 ml/min per 1.73 m² in
the presence of another risk factor for AKI might provide a reasonable balance between risk and response. SCr is limited in its ability to identify early AKI (4). Although numerous novel biomarkers have been proposed for early identification of intrinsic AKI, to date no biomarker has been shown to possess sufficient predictive ability to be used as a primary enrollment criterion and should not supplant SCr for enrollment into trials exploring the effects of agents on established AKI. However, the use of biomarkers for risk stratification in AKI trials warrants further study because reproducible and qualified biomarkers may enhance trial design.

The implications of specific inclusion and exclusion criteria on feasibility of recruitment and on potential generalizability of study findings should also be assessed. Realistic estimates of potentially eligible populations and likely enrollment rates should be assessed during study planning using electronic health records and laboratory and administrative databases. The clinical setting (outpatient versus inpatient versus intensive care unit) and concomitant severity of comorbid illnesses should also be considered.

### Tenable Endpoints for AKI Trials

**Moderators: Dr. Ron Wald and Dr. Glenn M. Chertow**

Study endpoints must provide reliable signals for the efficacy of the intervention being tested. At the same time, endpoints must be measured accurately and reliably and be meaningful to patients. In the context of trials involving AKI, the choice of endpoints will differ depending on whether the intervention targets prevention or treatment of established AKI. In patients with established AKI, endpoint selection will depend on the timing relative to the course of disease (e.g., following an abnormal test result for a putative kidney injury biomarker versus “advanced” AKI that necessitates dialysis). Finally, although surrogate markers of efficacy could be acceptable in phase 2 studies, phase 3 trials must explore outcomes with clear clinical relevance—assessing whether an intervention changes how long or how well a person lives or feels.

Different endpoints must be considered for the following scenarios: (1) a phase 2 trial testing a novel intervention to prevent AKI after a predictable insult (e.g., cardiopulmonary bypass or abdominal aortic aneurysm repair in highly susceptible patients), (2) a phase 2 trial testing an intervention to improve outcomes in established AKI, and (3) a phase 3 trial testing an intervention to improve outcomes in established AKI. For example, for a phase 2 trial testing a novel intervention to prevent AKI after a predictable insult, it would be appropriate to use the Acute Kidney Injury Network (AKIN) definition of AKI (a rise in SCr of 50% or $\geq 0.3$ mg/dl), using the preprocedure SCr as the baseline (5). To enhance the specificity and clinical significance of this endpoint, it might be reasonable to require that the elevation in SCr be sustained for 48 hours, with serial determinations of SCr every 12 hours. Although novel biomarkers have been proposed for diagnosis of AKI, such biomarkers require more extensive replication and qualification before they can be used as study endpoints.

For interventions directed at patients with established AKI, endpoints would depend on the timing of implementation of the intervention. For an early intervention (e.g., after a small increase in SCr or positive result for a putative biomarker), a composite endpoint of death, provision of dialysis (or reaching prespecified criteria at which dialysis would typically be instituted), or a sustained loss of kidney function at a discrete time point (e.g., 28 or 60 days) would be meaningful. An alternative approach would be the use of a weighted composite endpoint incorporating death, provision of dialysis, and sustained loss of kidney function (i.e., change in SCr from baseline to day 28 or 60) in descending order of importance or rank (6). In trials testing interventions later in the course of AKI (e.g., AKIN stage 3 or after dialysis initiation), a composite endpoint of death and kidney “nonrecovery” (i.e., dialysis dependence at 28 or 60 days or longer) might be appropriate.

Phase 3 clinical trials testing an early intervention in established AKI must emphasize outcomes that have unequivocal relevance to patients. A composite endpoint of death, provision of dialysis, and sustained loss of kidney function could be supplemented by several secondary outcomes examining longer-term consequences of impaired kidney function, including survival, development and progression of CKD, adjudicated cardiovascular events, health-related quality of life (HRQOL), and neurocognitive function. Although time-to-event analyses would not be appropriate when evaluating short-term outcomes, they may be useful in longer-term analyses.

### Biomarkers/Surrogate Markers

**Moderators: Dr. Patrick T. Murray and Dr. Chirag R. Parikh**

A biomarker is defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” (7). Biomarkers can serve multiple roles in clinical trials: quantification of risk for disease occurrence or progression (prognostic biomarker), assessment of likelihood of response to treatment (predictive biomarker), assessment of whether a biologic response has occurred (pharmacodynamic biomarker), or assessment of an endpoint (surrogate biomarker).
What Are the Proper Biomarkers to Be Used in Interventional Studies in Patients with AKI? Demographic and clinical risk factors for the development of contrast agent–induced AKI and cardiac surgery–induced AKI are well defined. There is, however, less precision in characterizing milder episodes of AKI compared with ascertaining the need for acute renal replacement therapy. In both settings, impaired kidney function at baseline is the dominant risk factor. Although identification of other biomarkers may refine pre procedural risk prediction, the main limitation to the conduct of prevention trials in these settings will remain the need to recruit very large numbers of patients to provide the adequate number of outcomes (events) required to achieve adequate study power.

Currently, changes in SCr remain the most useful biomarker for case identification and enrollment in intervention trials (5,8,9). Although decreased urine output is an alternate functional biomarker, there is less evidence to establish the appropriateness of the urine output components of the Risk, Injury, Failure, Loss of function, and End-stage Renal Disease (RIFLE) and AKIN staging systems as trial entry criteria. Whichever functional criterion is used for AKI case definition, appropriate enrollment might be further improved by the use of urine microscopy and indices of tubular function, such as fractional excretion of sodium (FE\text{Na}) or urea combined with biomarkers of renal tubular damage (9,10) to differentiate patients with and without intrinsic kidney injury. Although this approach is conceptually attractive, data on use of novel AKI biomarkers for early AKI case identification in clinical trials are inadequate. In the only published trial to have been guided by urinary biomarker elevation, the biomarkers used were not sufficiently predictive to identify the desired study population (11). Real-time, point-of-care–measured GFR represents another tool that could improve case identification and staging or stratification at clinical trial enrollment, as well as serving as a trial endpoint, but such technology is not currently available.

Staging of AKI, using the RIFLE or AKIN systems, to assess the severity of injury may help demonstrate proof of concept in phase 2 studies of interventions to prevent or mitigate the severity of AKI (5,8,12,13). Novel biomarkers, if suitably robust in their discriminatory ability, may further enrich such assessments (14,15). Accordingly, decisions to proceed with development of AKI therapies could be guided by a combination of biochemical, physiologic (measures of GFR, tubular function, and urine output), and pathologic structural (tubular AKI damage biomarkers) markers of kidney damage. The combined use of case adjudication and biomarkers could be used to target specific AKI “phenotypes.” For example, prerenal azotemia cases could be characterized by low FE\text{Na}, absence of kidney damage markers, and improvement in kidney function with volume expansion or restoration of renal perfusion. Such cases might be appropriate for early hemodynamic interventions or renal vasodilators. In contrast, cases of intrinsic kidney damage (e.g., acute tubular necrosis) would have a high FE\text{Na} and positivity for damage markers and be less responsive or unresponsive to volume expansion.

Are There Valid Surrogate Markers to Be Used in Studies of AKI Therapies for Determining Severity and Prognosis? A wealth of information documents the independent association of AKI defined by SCr increments and adverse patient outcomes, including death (12–15). However, in the absence of trials demonstrating that prevention or treatment of AKI mitigates these outcomes, it has not been possible to definitively validate small increments in SCr or other biomarkers as surrogate endpoints for AKI clinical trials. In the past, successful improvement of an endpoint such as a 25% relative or 0.5 mg/dl absolute increment in SCr within 48 hours of radiocontrast exposure may have been sufficient to change clinical practice if an existing agent was associated with minimal risk (e.g., N-acetylcysteine or sodium bicarbonate) and definitively shown to be beneficial; however, such endpoints are not currently acceptable for regulatory approval of a new drug. Nonetheless, such endpoints may be acceptable for proof-of-concept studies, particularly when supplemented with supportive and mechanistic information derived from novel biomarkers. Biomarkers of AKI severity might also be included as components of clinical composite endpoints or as secondary or tertiary endpoints for later stage (phase 2b and 3) clinical trials of AKI therapies, in addition to the currently accepted endpoints of death and need for dialysis.

Patients with CKD have the highest risk for developing AKI. Therefore, it is important that the cutoff values and test performance of AKI biomarkers are appropriate for patients with underlying CKD. The characteristics of biomarkers used for detection of kidney injury need to be fit for purpose. The required cutoffs and desired test performances may differ among early-phase drug development in healthy persons, proof-of-concept studies in patients with comorbid conditions, and “definitive” (event-driven) AKI trials.

Power and Sample Size Issues

Moderators: Dr. Andrew D. Shaw and Dr. Alan S. Go

How Should Power Analyses for Interventional Studies of AKI Therapies versus AKI Prevention Strategies Be Informed? The statistical methods for calculating sample sizes given minimally acceptable power are well described and are driven by several factors, including the ratio of treatment group size, estimated effect size of the prevention or treatment strategy compared with the control approach, rate of the outcome of interest in the control group and its constancy over time, drop-out (and, where applicable, drop-in) rates in each group, and diagnostic clarity (i.e., error in the diagnosis of AKI itself) (16). The current challenge for the field of AKI prevention and treatment is the dearth of robust information in several of these areas.

What Are the Critical Data Needed to Inform Power/Sample Size Calculations for AKI Prevention or Treatment Trials Using Biomarker/Surrogate Markers? “Hard” Clinical Outcomes? Patient-Centered Outcomes? What Are the Effect Sizes That Are Likely to Have Clinical and Public Health Relevance but Also Allow for Feasible Recruitment? The event rate and biologically plausible effect sizes in both prevention and treatment trials depend on the clinical endpoints chosen. For example, event rates based on small changes in SCr (e.g., ≥0.3 mg/dl or ≥50% increase in SCr) will be substantially higher than rates of severe AKI necessitating dialysis or 28-day mortality.
Similarly, the plausible effect of any intervention will vary with the type of endpoint, and generally with larger effect sizes associated with surrogate compared with hard endpoints. Thus, the use of biomarkers or surrogate markers might allow for study designs that are shorter or require substantially fewer participants. These designs will be more appropriate for phase 2 trials, while more definitive phase 3 trials should be powered on “hard,” clinically meaningful endpoints.

For both AKI prevention and treatment trials, patient centered-outcomes (17) such as functional status and HRQOL are also important but should be considered secondary to objective events, such as death, substantial and sustained loss of kidney function, or other complications (e.g., adverse cardiovascular events), in sample size and power calculations. Given that functional status and HRQOL are typically measured using continuous scales, a prevention or treatment trial that is adequately powered to detect differences in event rates will probably be more than adequately powered for these secondary outcomes.

Determining a plausible effect size in prevention or treatment studies is difficult given the limited data available. Previous trials have generally used overly optimistic treatment effect assumptions, causing these studies to be grossly underpowered. Data related to hard clinical endpoints collected during phase 2 trials evaluating surrogate outcomes will help inform more accurate estimates of potential effect sizes. In the absence of better data, it is probably safest to adopt a conservative approach, remembering the “winner’s curse” effect (18) in which follow-up studies almost always have smaller effects than the initial positive study.

What Are the Optimal Data Sources That Could Be Used to Inform These Analyses for Different Target Populations (e.g., After Elective Surgery, Administration of Radiocontrast Agents, Sepsis, Trauma/Intensive Care Unit)? Although new national and international registries of patients who are at risk for or experience AKI would be ideal to construct, the cost and time needed are probably prohibitive for providing short-term insights needed for current trial planning. Multiple data sources could be leveraged to inform clinical trial planning for selected AKI prevention and treatment populations (Table 2) (19–34). Depending on the data source, insights into the likely proportion of eligible patients as well as preliminary outcome rates at different follow-up times would be available to assist in sample size and power calculations. Selected information about the range of event rates for presumed control groups may be available from pooled analyses of existing randomized trials involving target populations.

What Are the Implications of Different Clinical Trial Designs and Follow-up Timeframes for Power and Sample Size for AKI Prevention versus Interventional Trial Approaches? The use of a traditional randomized trial design is appropriate for both prevention and treatment studies. Of note, however, follow-up for both prevention and treatment trials should have a time horizon that will allow assessment of sustained losses in kidney function along with other clinical events. This would require assessment of outcomes during the index hospitalization as well as at 60–90 days and later (e.g., 12 months or longer) to understand whether improved prevention and treatment strategies have long-term clinical benefits. In addition, adaptive study designs should be considered because they provide for adjustment in sample size based on prespecified interim assessments of endpoint rates (35). Such a design would facilitate more timely identification of strategies that have a higher likelihood of success, as well as allow early termination of approaches with low probability of success.

What Are the Advantages and Disadvantages of Incorporating Stratified Randomization Variables for Specific Patient Subgroups of Interest? Identifying subsets of patients who are at higher risk for adverse events is a key step because they may be likely to benefit from a prevention or treatment strategy. However, enrolling such patients does not necessarily guarantee a higher rate of preventable adverse events, just a higher overall outcome rate, because higher-risk patients may also exhibit diminished response to interventions. The relative effect on individual components of a composite endpoint must also be considered. For example, a preventive strategy may decrease the risk for AKI but have no other effects on the underlying rate of progression or other unrelated causes of death.

For prevention trials, targeted subgroups may include those with impaired kidney function or overt proteinuria, as well as patients with impaired kidney function admitted with decompensated heart failure. For treatment trials, there is a need to achieve standards regarding how to define AKI using existing and newer urine and serum biomarkers along with further epidemiologic studies that identify AKI subsets most likely to experience clinical outcomes of interest across various target populations.

How Should Stopping Guidelines Be Used in Different AKI Prevention and Interventional Trials? How Does This Influence Power and Sample Size? Controversy still exists regarding whether Bayesian versus frequentist statistical methods are optimal for planning and analyzing clinical trials. This controversy has important implications for clinical trial design and sample size estimates, especially if adaptive trial designs are implemented for a definitive outcome trial. Clinical trials should be planned to err on the side of longer follow-up (i.e., more conservative stopping rules) given experience from other disciplines about the disadvantages associated with premature stopping of studies due to apparent greater-than-expected benefit early in a trial (36).

Considerations regarding Adverse Events in Pilot and Feasibility Studies in AKI

Moderators: Dr. Sarah G. Faubel and Dr. John A. Kellum

Adverse events in clinical trials are context specific and generally fall into three categories: medication- or intervention-specific, patient-specific, and disease-specific.

Medication- or intervention-specific adverse events are due to the drug (or device) itself. The risk profile of an intervention may range from very low, with rare adverse events, to high, with frequent expected adverse events.

Patient-specific adverse events are related to the underlying diagnoses in the patient population being studied. It is important to understand the comorbid conditions of the study population in order to attribute adverse events to the intervention as opposed to expected adverse events related to the underlying health of study participants.
Disease-specific adverse events are related to the natural history of the specific disease being studied. The mortality attributable to the disease in the presence of AKI must also be considered. For example, patients admitted to the hospital with community-acquired pneumonia have a 90-day mortality rate of 5% that increases to 10% in the setting of AKI (37). Similarly, hospital mortality after routine cardiothoracic surgery is generally low, at less than 2%, but increases to 11% in the setting of AKI (38). When studies in patients with AKI are being planned, it is critical to accurately assess the natural history of the diseases in the population with AKI and to integrate this knowledge with potential side effects of the drug or intervention to be studied.

A clear understanding of patient- and disease-specific outcomes in the context of treatment and medication outcomes will facilitate improved clinical trial design as well as interpretation of adverse effects that may occur during the trial. Knowledge of expected outcomes is particularly important in pilot studies with small sample size, where it cannot be assumed that the control group will be of sufficient size to allow an accurate estimate of event rates. As with all interventions, those developed for prevention or treatment of AKI carry risk, with risk-to-benefit ratios varying across trials of prevention and treatment. Although risk of a particular intervention may be relatively constant, its potential benefit may vary as a function of disease severity. Patients without disease can benefit only if a treatment prevents disease. In general, acceptance of risk will increase as expectation of benefit increases, and the tolerance of risk will generally increase with severity of AKI.

Prevention Trials. The most common settings in which prevention of AKI are studied are exposure to radiocontrast agents and cardiothoracic surgery. In both these settings the incidence rate of AKI in the general population is low. For example, AKI necessitating dialysis occurs in approximately 1% of patients after cardiac surgery (39). Thus, AKI prevention trials must accurately assess the side effect profile of medications and adverse outcomes of interventions because only a low threshold for adverse effects would be tolerated as a result of the low risk for adverse outcomes in general. This is another reason that selecting a population at increased risk is important.

Treatment Trials. Tolerance of risk for interventions and medications once AKI has developed is higher than for prevention trials. Mortality risk increases with AKI, especially in the setting of critical illness. Thus, the potential benefit of effectively managing AKI may outweigh the risks associated with certain medications and interventions. Assessment of severity of AKI is important for assessing risk-to-benefit ratios, given the increased risk for disease-specific adverse events with greater severity of AKI.

Interventions may also have the greatest benefit when administered early, at a stage at which AKI has not yet manifested or in patients whose disease severity is still low. Risk stratification and “fit for purpose” qualification of early biomarkers for detection of AKI are needed. The utility of current staging systems (e.g., AKIN) that rely on SCr may be limited, especially when early intervention is desired, because SCr is a lagging indicator. Therefore, stratification strategies and emerging biomarkers with sufficient sensitivity and specificity need to be developed.

When trials are designed for prevention or treatment, it is important to understand the mortality risk attributable to AKI in particular settings. Once the effect of AKI on mortality is estimated, the risk and benefit of a particular intervention may then be assessed in the appropriate context.

Pilot and Feasibility Studies. Inclusion criteria for pilot and feasibility studies for treatment of AKI should focus...
on distinct populations. A homogenous population with relatively few underlying comorbid conditions and a well defined course of AKI will facilitate adequate assessment of potential beneficial and adverse effects of a particular medication or intervention. Thus, patients with a clearly identified cause of AKI would be more suitable for pilot and feasibility studies than patients with multifactorial AKI. **Intervention and Definitive Studies.** In contrast to pilot and feasibility studies, inclusion criteria for intervention and definitive studies of treatment of AKI should be as broad as possible. For example, AKI in critically ill patients is generally multifactorial. Definitive trials focusing on a relatively “pure” cause of AKI would therefore have limited generalizability. Patients with certain conditions, such as urinary tract obstruction, acute interstitial nephritis, GN, and vascular conditions should be excluded because specific therapies for these disorders already exist and these conditions are sufficiently well defined to permit specific treatments. Similarly, patients with various conditions influencing renal perfusion (hepatorenal and cardiorenal syndromes) might be excluded from some trials. However, these conditions may also cause parenchymal injury and may also coexist with other conditions directly affecting the kidney.

Justification for broad inclusion of patients in definitive trials of AKI is supported by analogy to acute lung injury (40). Acute lung injury is defined by specific quantifiable criteria that do not distinguish among the over 60 causes associated with its diagnosis (40). Use of a uniform definition of acute lung injury in clinical trials has helped establish remarkable reductions in mortality over the past decade, largely associated with low tidal volume mechanical ventilation and conservative fluid management (41). By analogy, applying an intervention using broad inclusion criteria for clinical trials of AKI might result in similar advances in the overall care of patients with AKI.

**Acknowledgment**

The views expressed in this report should not be taken to represent the views of the US Food and Drug Administration.

**Disclosures**

None.

**References**


21. Coca SG, King JT Jr, Rosenthal RA, Perkal MF, Parikh CR; The duration of postoperative acute kidney injury is an additional


Published online ahead of print. Publication date available at www.cjasn.org.

This article contains supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.12791211/-/DCSupplemental.