UNDERSTANDING KEY CONCEPTS IN NEPHROLOGY RESEARCH 2024

A Collaboration of the International Society of Nephrology (ISN) and Can-SOLVE CKD Network

Topic A: Patient-Reported Outcomes in Clinical Trials

Topic B: Re-Examining Targets and Thresholds in Clinical Trials for CKD Complications: Enrolment Criteria, Treatment Targets, and Outcomes

Topic C: Precision in Study Design

Topic D: Hierarchical Composite Endpoints

This document was prepared by Can-SOLVE CKD in collaboration with the ISN for the patient partners invited to attend the ISN's International Consensus Meeting on Changing Paradigms of Studies in CKD, held in Vancouver, BC, Canada, November 22-23, 2024.

This document contains explanations of research concepts, layperson summaries of the preparatory readings suggested by the meeting workshop's leads, and a glossary of clinical research terms. It was developed to prepare the six patient partner attendees of the ISN's meeting to enable informed participation in the workshops and plenary discussions. Note that many of the clinician and trialist attendees of the meeting considered the document to be an excellent educational tool for patients, residents, and colleagues.

> "It helped me to understand some of the complex technical concepts better than I had ever been able to; our residents would love this ! "
>
> MD PhD Researcher.

Content Contributors:

Topic A: Patient-Reported Outcomes in Clinical Trials
Leads: Alison Jaure, Jane Schell, Matthias Rose, Kevin Weinfurt

Topic B: Re-Examining Targets and Thresholds in Clinical Trials for CKD Complications: Enrolment Criteria, Treatment Targets, and Outcomes
Leads: Adeera Levin, Morgan Grams, Jennifer Lees

Topic C: Precision in Study Design
Leads: Meg Jardine, Rathika Krishnasamy, Smeeta Sinha

Topic D: Hierarchical Composite Endpoints
Leads: Hiddo Lambers Heerspink, Dustin Little, Patrick Schloemer

Can-SOLVE CKD Staff: Melanie Talson, Cynthia MacDonald, Omosomi Enilama, Tara Ozbay, Isabela Varela

Patient Partners: Bill Wang, Charles Cook, Duane Sunwold, Jocelyn Jones, Kate Chong, Thomas Ng

ISN Staff: Sandrine Damster and Jo-Ann Donner

Table of Contents

A

Patient-Reported Outcomes in Clinical Trials

Clinical outcomes assessments (COA) are measures that describe or reflect how a patient feels, functions, or survives due to a medical condition or its treatment. The four types of clinical outcome assessments are:

Patient-reported outcomes (PROs): Patients self-report their symptoms or quality of life.

Clinician-reported outcomes (ClinRO): Clinicians assess and report patient condition.

Performance outcomes (PerfO): These are objective measurements of a patient's ability to perform specific tasks.

Observer-reported outcomes (ObsRO): Caregivers or others report on the patient's health.

In clinical care, a patient's voice or perspective is typically documented qualitatively by clinicians in patient charts or electronic health records. These entries provide a narrative account of the patient's experiences, but they are often subjective and can vary in detail and focus, depending on the specific conversation or interaction.

While valuable, this approach can make it harder to systematically track health trends over time or compare experiences across patient groups with similar conditions. By using patient-reported outcome measures (PROMs), clinicians and researchers can gather patients' views on their health in measurable, numerical forms. This data can then be easily integrated into treatment decisions and medical studies, offering a clearer view of patients' experiences with their conditions.

A

PROMs

Patient-reported outcome measures (PROMs) are standardized, validated survey tools that assess health outcomes reported by patients. These instruments can be used to assess concepts that are narrow (e.g., **pain intensity**) or broad (e.g., **health–related quality of life**). PROMS can be developed for generic and disease-specific use and consist of well-crafted questions that focus on a specific concept and provide a range of response options. Using PROMs is a good way to incorporate a patient's perspectives into clinical trial **endpoints** (targeted outcome of a clinical trial).

SONG Outcomes

The **SONG Core Outcomes** in nephrology refer to a standardized set of key outcomes identified by the Standardized Outcomes in Nephrology (SONG) Initiative. The four main SONG Core Outcomes identified across kidney disease stages are mortality, kidney function, cardiovascular disease, and life participation. **Life participation** refers to the ability to participate in meaningful activities of daily life, including work, study, and social recreational activities. This is rarely assessed in clinical trials.

Psychometric Evaluations

To ensure that PROMS are reliable, valid, and responsive to change, **psychometric evaluations** can be done to assess factors such as the following:

Validity - the extent to which a tool measures what it is intended to measure
Reliability - a measurement tool's ability to produce similar results under consistent conditions
Acceptability - how suitable and user-friendly a tool is for participants
Completion Rate - the percentage of participants who fully complete a given test or survey
Internal Consistency - how well different items within a test measure the same concept
Generalizability - how well a test or measure works for different people and situations outside the original group studied

A

Additional Terms

Health–related quality of life (HRQOL) is a measure of how well a person can live and enjoy their life despite health issues. It looks at different areas like physical health, mental well-being, social connections, and daily activities.

Endpoint is a targeted outcome of a clinical trial that can help determine how well and how safe a new therapy or treatment is. Endpoints may also be used throughout a study to determine if a participant's risk of continuing to be in a study is too great. Endpoints for a clinical trial may include one or more **clinical outcome assessments and/or surrogate endpoints**.

Summary of Readings

Selewski DT, et al. Patient-Reported Outcomes in Glomerular Disease. Clin J Am Soc Nephrol. 2017 Jan 6;12(1):140-148.

This paper provides an overview of terminology described above such as **COA, PROMs, ClinRO, ObsRO, PerfO, and HRQOL**, and development methods for PROMs.

Some of the main messages of the paper are:

- A key aspect of capturing the patient voice in clinical trials is deciding what to measure.
- **PROMs** and other **COAs** that aim to measure a broad range of concepts reported in a summary score are generally not the best way to show how effective a treatment, drug, or intervention is in a clinical trial.
- Some instruments can contain questions assessing outcomes that could be indirectly related to the effects of the trial, unrelated to the trial itself (i.e. family support), or outcomes that can be affected by other factors other than the disease or trial.
- It is important to use specific **PROMs** that focus on core disease–related symptoms or functional effects of a disease, which may be affected by the clinical trial. For example, including aspects of well-being, like feelings of worry and hope, that are only loosely connected to how well a drug works can make it harder to measure the drug's effectiveness.

Summary of Readings Continued

Weinfurt KP. Constructing arguments for the interpretation and use of patient-reported outcome measures in research: an application of modern validity theory. Qual Life Res. 2021 Jun;30(6):1715-1722.

This paper describes how the definition and evaluation of **validity** has changed over time within the fields of psychology and education, and how these new methods can be used to evaluate PROMs in a research setting.

Instead of just checking off all possible validation methods (the checklist approach), researchers could now focus on specific, logical arguments to make sure the tool measures what it's supposed to. This new method, called "argument-based validity," helps researchers focus on important evidence and assumptions to support PROMs.

The argument-based approach to validity also involves starting with a clear idea of how scores will be used or understood, because scores can be interpreted in different ways. For example, a score computed from multiple items that ask about how far and how easily one can walk could be interpreted directly as an indicator of ease and extent of walking, or indirectly used to assess lower mobility or physical functioning if certain assumptions are held.

Using the argument-based validity approach, the authors suggest that researchers should hold certain assumptions to properly interpret and use **PROMs** in a clinical trial or study:

- **PROMs** should reflect all of what is being studied, not just part of it.
- The patient should understand the questions and the possible responses.
- The responses should not be majorly affected by other things, like cultural backgrounds or reading level.
- Responses are usually given a score, for example, "very likely" might equal 5. All responses are added and turned into a score. This score should make sense.
- The score should describe how patients feel or function in their daily lives.
- The scores should be "sensitive enough," meaning they are able to show differences in patients who receive a certain treatment over time

Summary of Readings Continued

Jaure A, et al. Validation of a Core Patient-Reported Outcome Measure for Life Participation in Kidney Transplant Recipients: the SONG Life Participation Instrument. Kidney Int Rep. 2023 Nov 3;9(1):87-95.

The authors of this paper reviewed the few existing outcome measures (tools or methods) used to assess **life participation** in kidney transplant recipients and found that they were too costly, burdensome for patients to complete (lengthy, questionable relevance of content), heterogenous (different questions being asked), and indirect.

As a result, the authors decided to hold a consensus workshop attended by kidney transplant recipients, caregivers, and health professionals from across eight countries and used a rigorous methodological framework to develop an instrument called SONG-LP. **SONG-LP** is a short and simple tool to assess life participation dimensions of direct importance and relevance to kidney transplant recipients. It also has the potential to be used in other health populations.

The authors performed a psychometric evaluation on SONG-LP and found that this instrument had moderate to high reliability, a high completion rate, and good acceptability and internal consistency. It also showed validity, but it requires further validation in broader populations.

The authors ended the paper by highlighting some of the potential issues or limitations with the SONG-LP. First, many of the kidney transplant recipients reported high scores on SONG-LP so another evaluation done on a broader cross-section of patients, ranging from kidney transplant recipients who are clearly unwell to well, may need to be done to see whether it is an underlying problem with the SONG-LP measure. Second, there was a weak correlation (relationship) between life participation and clinical variables (kidney function and time since transplant) and more work needs to be done to better understand the patterns and relationships between clinical factors (e.g., infection, cardiovascular disease, diabetes, and other comorbidities) or medication side effects and life participation. And third, participants in the assessment were recruited from high-income English-speaking countries, and the assessment has not been validated cross-culturally, limiting its generalizability.

B

ISN | RESEARCH

Can-SOLVE CKD Network

Re-Examining Targets and Thresholds In Clinical Trials For CKD Complications: Enrolment Criteria, Treatment Targets and Outcomes

Conditions such as CKD are associated with multiple abnormal parameters measured clinically and biochemically. Many of these have been measured longitudinally (over time) and are associated with adverse patient outcomes. Clinical trials design has focused on restoring these parameters to a normal range, but this has often not improved patient outcomes, and sometimes the interventions have caused harm.

This may happen because some abnormal values are not necessarily harmful or require correction. In addition, variability in measurements and reliance on **surrogate outcomes** can lead to misleading results, as improvements in surrogate markers don't always translate to better outcomes like reduced mortality. Lastly, the diverse nature of CKD populations and incomplete understanding of key **biomarkers** can make clinical trials less reliable and risk excluding potentially effective treatments.

In clinical trials targeting CKD complications, treatment effectiveness is often measured using different approaches to define treatment goals. **Absolute targets** aim for a specific lab value, such as potassium levels below 5.0 mmol/L or phosphorus levels below 4.5 mg/dL. **Relative targets** focus on percentage changes, such as reducing phosphorus levels by 30% from baseline. Alternatively, some approaches prioritize general treatment strategies, such as prescribing phosphate binders without setting a strict numerical goal.

B

ISN | RESEARCH

Can-SOLVE CKD Network

Key Biomarkers in CKD

Parameter/Biomarker	Change in CKD	Explanation	Related Symptoms
Serum Creatinine		Waste product from muscle breakdown; high levels indicate reduced kidney function.	Fatigue, confusion, swelling due to waste buildup
Blood Urea Nitrogen (BUN)		Byproduct of protein breakdown; this builds up when kidneys can't filter effectively.	Nausea, vomiting, metallic taste, fatigue (**uremia** symptoms)
Estimated GFR (eGFR)		Measure of kidney filtration efficiency; lower values reflect declining function.	Often asymptomatic in early stages, fatigue or swelling in later stages
Serum Potassium		High levels can result from poor excretion; this affects heart and muscles.	Muscle weakness, irregular heartbeat (**hyperkalemia**)
Serum Phosphorus		Excess phosphorus can damage bones and lead to calcification of blood vessels.	Itchy skin, bone pain, increased risk of fractures
Serum Calcium		Low calcium is linked to bone weakening and poor regulation of phosphorus.	Muscle cramps, tingling in hands/feet, brittle bones

B

ISN | RESEARCH

Can-SOLVE CKD Network

Key Biomarkers in CKD

Parameter/Biomarker	Change in CKD	Explanation	Related Symptoms
Parathyroid Hormone (PTH)		Overactive in CKD to balance calcium and phosphorus; worsens bone health.	Bone pain, weakness, fractures
Serum Albumin		Low levels of this protein can mean malnutrition, dehydration, or inflammation.	Swelling in legs/ankles, fatigue, poor wound healing
Hemoglobin		Red blood cells decrease due to reduced erythropoietin production by kidneys.	Fatigue, shortness of breath, pale skin (**anemia** symptoms)
Urinary Protein/Albumin		Indicates kidney damage as proteins leak through damaged filters.	Frothy urine, swelling in legs/feet
Bicarbonate		Low levels reflect metabolic acidosis, where the body becomes too acidic.	Fatigue, confusion, breathlessness

Key Biomarkers in CKD

Parameter/Biomarker	Change in CKD	Explanation	Related Symptoms
Serum Uric Acid		Excess uric acid can accumulate due to poor excretion by kidneys.	Joint pain, swelling, gout like symptoms
Cystatin C		This protein rises as kidney function worsens; confirms eGFR findings.	Often no direct symptoms; supports CKD diagnosis
Vitamin D (25- OH)		Low levels affect calcium absorption, worsening bone health.	Bone pain, fatigue, muscle weakness
Triglycerides		Elevated in CKD, increasing cardiovascular risk.	No direct symptoms, but linked to heart disease
LDL Cholesterol		“Bad” cholesterol often rises, further increasing heart disease risk.	Chest pain, shortness of breath (if severe **atherosclerosis** develops)

B

ISN | RESEARCH

Can-SOLVE CKD Network

Summary of Readings

Levin A, et al. Targets, trends, excesses, and deficiencies: refocusing clinical investigation to improve patient outcomes. Kidney Int. 2013 Jun;83(6):1001-9.

This paper talks about the challenges and failures of **clinical trials** in nephrology. The main problem is that many treatments don't work as expected, and sometimes even cause harm. The paper suggests that focusing too much on "normalizing" lab test results, like blood levels, might not help patients and could make things worse.

The paper also looks at why some clinical trials succeed, while others fail. It suggests that clinical trials need to better understand the disease, the patients, and how treatments work before they start. One big issue is that people with kidney disease are very different from each other, so one-size-fits-all treatments don't always work. For example, some patients might improve, while others get worse, even with the same treatment.

The authors suggest a few ideas for better clinical trials:

1. **Study the patients more carefully** to understand their different conditions.
2. **Use better tests (biomarkers)** to figure out which patients will benefit from which treatments.
3. **Create treatments that work with the body's natural processes** instead of trying to fix lab values that might not even be causing the problem.
4. **Reconsider how we design trials** – not everything needs to be compared to a "perfect" treatment. It might be better to test treatments in real-world settings.
5. **Include more patients** in long-term studies to gather more data.

The paper also points out that many trials failed because they focused on lab numbers, like blood pressure or kidney function, instead of real patient outcomes, such as survival or quality of life. The goal should be to help patients live better, not just to fix numbers.

In short, kidney disease is complex, and trial designs need to reflect that. More careful research and a focus on understanding the disease, rather than just hitting lab targets, might lead to better treatments and outcomes for patients.

Summary of Readings Continued

Riddle MC, et al. A1C Targets Should Be Personalized to Maximize Benefits While Limiting Risks. Diabetes Care. 2018;41(6), 1121–1124.

This article discusses new guidelines from the **American College of Physicians (ACP)** about how people with type 2 diabetes (T2D) should manage their blood sugar levels, specifically focusing on the **glycated hemoglobin (A1C) target**. The ACP suggests relaxing the goals for A1C control, recommending that for most patients, the target should be between 7% and 8%, and that for people over 80 or with serious health conditions, there should be no specific A1C target at all. The guidelines also propose reducing medication for people whose A1C is below 6.5%.

However, the authors of this article disagree with these guidelines. They argue that the ACP is not considering all the evidence. For example, in some large studies, more intensive treatment to keep A1C below 7% did not lead to harmful outcomes, and in fact, it helped prevent long-term complications like heart disease and damage to the eyes and kidneys. The authors believe that aiming for **A1C levels lower than 7%** is still a good goal for many people, especially if it can be done safely without causing low blood sugar (**hypoglycemia**) or weight gain.

The article also talks about how new treatments for diabetes, such as certain medications that don't cause hypoglycemia, could help improve patient outcomes, especially for those at high risk of heart disease or kidney problems.

In conclusion, the authors of this article believe that a more personalized approach to managing diabetes is needed, one that takes into account each patient's specific health, preferences, and goals, rather than applying a one-size-fits-all guideline. They suggest that the guidelines from the **American Diabetes Association (ADA)**, which recommend keeping A1C below 7% for most people, are more balanced and better supported by evidence.

Summary of Readings Continued

Canney M, et al. Regional Variation in Hemoglobin Distribution Among Individuals With CKD: the ISN International Network of CKD Cohorts. Kidney Int Rep. 2023; 8: 2056–2067.

This article examines differences in **hemoglobin** levels in CKD patients, and argues that these differences should lead to changes in how healthcare is provided.

The researchers propose that differences in hemoglobin vary geographically, as individuals living in Sub-Saharan Africa and South/Central Asia have lower hemoglobin levels and higher levels of **anemia** in general. Additionally, there are sex-based differences in hemoglobin levels, as women on average tend to have lower levels of hemoglobin compared to men. The researchers point out that these differences are largely ignored in a CKD setting, as there are frequent assumptions that all individuals living with CKD have similar hemoglobin levels.

This assumption is harmful, as it creates knowledge gaps, and decreases healthcare providers' abilities to individualize research approaches when assessing anemia in CKD patients.

The authors found that women and individuals living in Sub-Saharan Africa and South/Central Asia do have lower hemoglobin levels on average compared to others. Despite the fact that a women's hemoglobin levels are on average lower, this gap between men and women decreases as we compare individuals whose CKD is more progressed.

Additionally, their findings show that individuals who have CKD **and** another chronic disease, such as diabetes, polycystic kidney disease, or hypertension, have higher levels of hemoglobin in general.

In conclusion, the researchers state that anemia that occurs in the presence of CKD varies significantly around the world. There are many reasons for these geographic and sex-based differences, including: prevalence of disease, genetic factors and diet.

Precision in Study Design

When it comes to **precision medicine trials**, several trial design options means treatments can be tailored to specific patient groups.
Basket trials test the effectiveness of a single drug on a variety of diseases, all with a shared genetic mutation.
Umbrella trials explore different treatments for multiple subtypes of a single disease.
Platform trials allow multiple treatments to be tested simultaneously across multiple diseases.
Adaptive trials adjust protocols based on interim data to improve patient outcomes and reduce trial costs.

The key **benefits** of these trial designs include cost-effectiveness, increased efficiency, and the ability to test multiple treatments across multiple diseases or conditions. Some of the **challenges** include the need for specialized logistical and operational support, coordination between multiple sites, and managing data sharing.

- **Inclusion and exclusion criteria** play a crucial role in determining who can participate and who cannot.
- **Exclusion criteria** in particular, when too restrictive, can result in a trial population that does not accurately reflect the diversity of the broader patient population, thereby limiting the generalizability of the findings.
- Another challenge in trial design is **outcome selection**, as the outcomes used to measure the effectiveness of an intervention need to be carefully chosen to reflect what is most meaningful to patients.
- The selection of **surrogate endpoints**, such as biomarker levels or early disease indicators, can be useful, but they don't always correlate with tangible improvements in hard outcomes like mortality or quality of life.

Summary of Readings

[The Adaptive Platform Trials Coalition. Adaptive platform trials: definition, design, conduct and reporting considerations. Nat Rev Drug Discov. 2019;18:797–807.]

This paper talks about **Adaptive Platform Trials (APTs).** APTs are a type of clinical trial designed to test multiple treatments at once, adjusting which treatments are given based on how well they're working during the trial. APTs are different from traditional clinical trials because they can change while they're running, depending on how the treatments perform.

Key Ideas:

1. **Design and Simulation:**
   - A lot of planning is needed before starting an APT. Researchers have to think carefully about the rules for how treatments are assigned, how missing data is handled, and the chances of false positives (wrongly thinking a treatment works) or false negatives (missing a good treatment).
   - They use simulations to test how the design will work in different situations before starting the trial.

2. **Review Before Launch:**
   - **Regulatory Review:** Before the trial starts, experts review the trial design to make sure it makes sense. They look at things like how treatments are chosen for patients and whether the study rules are clear.
   - **Ethical Review:** The trial needs to be explained in simple terms so patients can understand it and give their consent. The ethics committee ensures that risks are clear and that the trial is fair.

3. **Oversight During the Trial:**
   - Once the trial starts, it's monitored carefully. The team checks the data as it comes in to make sure everything is going as expected. If something unusual happens, they do more simulations to understand it.
   - The trial must follow strict rules to protect patients and ensure that the results are reliable.

Summary of Readings Continued

The Adaptive Platform Trials Coalition - Continued

4. **Reporting Results:**
    - After the trial, the results must be shared carefully. APTs may report data in a way that protects the trial's integrity, meaning they might not share all the raw data right away.

5. **Embedding APTs in Practice:**
    - APTs can be used in real-world clinical settings, where doctors use the findings of the trial to make decisions about patient care. This approach could help speed up the use of new treatments.

6. **Funding and Sponsorship:**
    - APTs don't fit into traditional funding models because they are ongoing and can test multiple treatments at once. They can be funded by government grants, non-profits, or companies interested in testing several treatments.

7. **When to Use APTs:**
    - APTs are especially useful when you want to test many treatments at the same time, especially if there are limited patients or you want to quickly add new treatments to the study.

In conclusion, APTs are a powerful tool for testing multiple treatments, especially for rare diseases or when trying to improve existing treatments. However, they require a lot of planning and careful management to make sure they are fair, ethical, and provide reliable results. As they become more common, the rules and best practices for running these trials will keep improving.

Summary of Readings Continued

Herrington WG, et al. Kidney disease trials for the 21st century: innovations in design and conduct. Nat Rev Nephrol. 2020 Mar;16(3):173-185.

This article calls for a major overhaul of **how kidney disease clinical trials are conducted**, emphasizing the importance of new technologies, personalized treatments, and focusing on outcomes that matter most to patients (e.g., **quality of life** and **kidney function**).

The authors suggest that **including patients more directly in trial design** and ensuring trials are faster and more cost effective should be a primary goal of the field of research in kidney disease. Traditional study designs are often slow and expensive and oftentimes do not reflect the real needs of patients.

Kidney disease is a leading cause of illness and death worldwide. However, despite many treatments being tested in clinical trials, kidney disease trials often take a long time and are costly. The authors suggest using new technologies such as phones, wearables, and electronic health records to collect more accurate and real time data.

This article also emphasizes **patient-centred approaches to clinical trials**. Researchers should ask patients what matters most to them, such as feeling better, avoiding dialysis, and focusing on their outcomes, as well as personalized medicine.

Kidney disease affects everyone differently. Therefore, the authors call for examining how different patients respond to treatments, including looking at a patient's age, genetics, lifestyle, and how advanced their CKD is.

Survival rates and "hard" outcomes are not the only variables that are important to consider when assessing a patient's outcomes. Quality of life, symptom relief, and maintaining kidney function are all equally as important to many patients.

Summary of Readings Continued

Woodcock J, et al. Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. N Engl J Med. 2017 Jul 6;377(1):62-70.

The **master protocol** is a new and different way to study therapies and diseases. Instead of researchers conducting independent trials for each therapy, by using a master protocol, researchers can **share overlapping information among the trials**.

For example, they could use a common screening platform to identify all trials for which a patient is eligible. This coordinated screening is at the heart of a master protocol and represents one of its chief advantages — **more efficient use of patients and resources**.

In independent trials, each separate trial does its own data collection and testing, with overlapping information being gathered for multiple trials but not shared among them. One innovative design feature unique to the master-protocol setting is the shared use of control patients among trials of the same biomarker or disease.

By taking advantage of **coordinated data collection across multiple trials**, master protocols can enhance other research initiatives. For example, having a master protocol in place for a rare disease can facilitate the collection of case histories of patients seen in the participating clinical practices, providing a data source for future externally controlled trials.

Master protocols come in different sizes and shapes but share many commonalities. All require increased planning efforts and coordination to satisfy the objectives of different stakeholders. Innovative design elements help ensure that maximum information is obtained from the research effort, and the infrastructure required for implementation increases data quality and trial efficiencies, compared with those in stand-alone trials. If designed correctly, master protocols can last many years, even decades, with innovations from the laboratory translating quickly to clinical evaluation. As the targets for new drugs become more and more precise, there is no alternative but to move forward with these coordinated research efforts.

D

ISN | RESEARCH

Can-SOLVE CKD Network

Composite Endpoints

Clinical trials aiming to develop new treatments for CKD have traditionally used **composite endpoints** (a combination of multiple outcomes) consisting of clinical events such as **kidney failure**, initiation of **kidney function replacement therapy** (dialysis or transplantation), and a sustained large (e.g., ≥50%) decrease in **GFR**. As many of these clinical events typically occur late in the disease course, we end up with large trials in which most participants do not contribute clinical events.

For those who do experience clinical events within the trial period, **time-to-event analyses** only capture the first event, overlooking any subsequent serious events that may occur later in the trial, and are usually done without considering how severe or clinically significant that event is. For instance, a participant might reach a composite endpoint after a 50% decline in GFR, while a more significant event like needing dialysis could happen later but isn't counted in the main analysis. This can lead to misleading results about treatment effectiveness.

GFR Slopes

To create smaller, more efficient trials, the rate of change of GFR over time (often referred to as "GFR slope") has been introduced and demonstrated to be a valid **surrogate endpoint** (indirect measurement) for CKD progression. The GFR slope provides an estimate of the effect of a treatment or intervention in all participants, regardless of if they experience a clinical event rapidly, slowly, or not at all. However, these GFR over time analyses can only capture gradual changes in kidney function rather than specific clinical events, which can make it challenging to interpret outcomes for patients with varying disease progression rates.

D

ISN | RESEARCH

Can-SOLVE CKD Network

Hierarchical Composite Endpoints

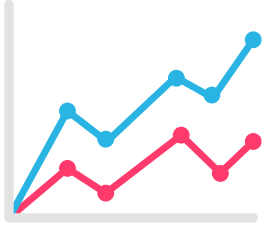
Hierarchical composite endpoints (HCE) has been introduced as a way to combine both clinical events and continuous GFR slopes, while ranking all components according to clinical severity. This approach lets researchers evaluate the effect of a treatment by comparing individual patient outcomes based on severity rather than just timing. HCEs also have the potential to include both adverse ("worsening") and favorable ("improvement") outcomes.

Example: If a CKD patient first experiences kidney failure, and later cardiovascular death, the more severe event—death—takes priority in the analysis. And if another CKD patient doesn't experience any clinical events during the trial, instead of their data being excluded, their GFR slopes would be analyzed and ranked last in terms of clinical severity.

Win Statistics

Win statistics which compare the number of wins, losses, or ties between the treatment groups can be used to analyze HCEs. First, the components of the HCE are ranked according to clinical importance, and each participant is analyzed according to their most clinically significant event, regardless of the timing. Then, each patient assigned to the treatment group is individually compared with each patient in the control group.

Each comparison results in a "win," "loss," or "tie" for the treatment group. The total number of wins, losses, and ties are used to calculate the win odds, which describes the odds of a better outcome for a patient in the treatment group compared with the control group. A win odds greater than 1 corresponds to a treatment benefit, indicating that treated patients are more likely to have a favorable outcome, compared with control patients.



D

ISN | RESEARCH

Can-SOLVE CKD Network

Summary of Readings

Heerspink HJL, et al. Development and Validation of a New Hierarchical Composite End Point for Clinical Trials of Kidney Disease Progression. JASN. 2023 Dec;34(12):2025-2038.

The authors develop and validate a kidney **HCE** using **win statistics**. The authors used data from several clinical trials that assessed the safety and efficacy of several drugs with a **composite endpoint** of kidney failure or death and **compared** them to the kidney HCE they developed. Their analysis showed that the win odds gave estimates similar to those from traditional time-to-first event statistical Cox analysis, but with higher statistical power. Meaning the HCE method seemed to be better able to detect treatment effects than traditional CKD endpoints by incorporating disease progression at different stages.

The kidney HCE developed by the authors ranked in order of highest to lowest effect was (1) all-cause mortality; (2) **kidney replacement therapy** defined as dialysis for at least 28 days or kidney transplantation; (3) sustained **GFR** <15 ml/min per 1.73 m2 for at least 28 days; (4) sustained GFR decline from baseline for at least 28 days of 57%; (5) sustained GFR decline of 50%; (6) sustained GFR decline of 40%; or (7) GFR slope.

A limitation (disadvantage) the authors highlight in this article is that while HCEs rank kidney failure above GFR decline and include more kidney failure events (since it's not a time-to-event analysis), the GFR slope still becomes the main HCE component. This is because many patients, who progress slowly with CKD during trials, don't experience significant clinical events, meaning GFR slope data primarily drives the results in the HCE.

While some might argue that death should always be considered the most important outcome, including it in HCE's can complicate matters for treatments that don't aim to reduce mortality. The authors suggest redefining the kidney HCE to include **all kidney outcomes** while excluding deaths that occur during follow-up.

Summary of Readings Continued

Little D J, et al. Validity and Utility of a Hierarchical Composite End Point for Clinical Trials of Kidney Disease Progression: A Review. JASN. 2023 Dec; 34(12):1928-1935.

This review article describes the use of **HCEs** in clinical trials of CKD progression, emphasizing the potential to combine **dichotomous** (where the result can only be one of two possible outcomes) **clinical events** with the continuous variable of GFR over time. In addition, all components would be ranked according to clinical significance. See the pros and cons table below of HCE and more.

The authors also identify some of the challenges associated with combining and prioritizing outcomes for HCE in specific settings. For example, creating a heart failure HCE that includes clinical events like cardiovascular death, hospitalizations, and symptoms requires balancing the goals of tracking both survival and symptom relief. It's also challenging to compare events like hospitalizations: one patient might have an early, short hospital stay, while another has a later, longer one, making it unclear which outcome reflects a "win" or "loss" in terms of treatment success.

Pros and Cons of Hierarchical Composite Endpoints

	Pros	Cons
Traditional Kidney Endpoints	• Can handle variable follow-up • Substantial contributions of clinical events in the overall composite.	• Analyses the first events of the patients and disregards the potential more severe events occurring later in the study. • Gives equal importance to GFR declines and kidney failure. • Does not capture treatment effect on patients without clinical events or large decline in GFR • Requires large sample size.
GFR Slope	• All participants contribute to the analysis. • Likely to decrease sample size compared with traditional endpoints.	• Severe events, such as death or kidney failure, do not contribute to the endpoint. • GFR values are noninformative after dialysis initiation.
HCEs	• Prioritises the most severe event of the patient. • The components are not considered of equal importance. • All patients contribute to the analysis, not only patients experiencing clinical events or large decline in kidney function. • Likely to decrease sample size compared to traditional endpoints.	• Requires a fixed time point (follow up duration) for evaluation. • Most patients contribute to the analysis with rate of change of GFR.

Little D J, et al. Validity and Utility of a Hierarchical Composite End Point for Clinical Trials of Kidney Disease Progression: A Review. JASN. 2023 Dec; 34(12):1928-1935.

Glossary

Acceptability is how suitable and user-friendly a tool is for participants.

Adaptive trial: A trial in which researchers can change the treatment plan based on the results they see during the study. For example, if a drug is showing good results in the early stages, more people might be added to that treatment group to get more data on its effectiveness.

Anemia is a condition where your body doesn't have enough healthy red blood cells to carry oxygen to your tissues. This can make you feel tired, weak, dizzy, or short of breath.

Basket trial: A trial in which researchers test one drug on different diseases that share a common feature, like a genetic mutation. For CKD, this would mean testing one drug across different types of kidney diseases that share a common feature. So, one drug can be tested on multiple kidney disease conditions at once to see if it works for all of them.

Biomarker is a biological sign or measurement in your body (something found in your blood, urine, or other body parts) that can help doctors understand your health, diagnose a disease, or to see how well the body responds to a treatment.

Completion rate is the percentage of participants who fully complete a given test or survey.

Clinical trial is a research study involving human participants that evaluates the safety and/or effects of one or more interventions on health outcomes. Interventions include, but are not limited to, drugs, vaccines, radiopharmaceuticals, cells and other biological products, surgical procedures, radiologic procedures, devices, genetic therapies, natural health products, process-of-care changes, preventive care, manual therapies, and psychotherapies. Data from clinical trials can be used to support the approval of drugs or to compare different medicines or treatments. The data can also help us determine which treatments are best for specific populations.

Clinical outcomes assessments (COA) are measures that describe or reflect how a patient feels, functions, or survives due to a medical condition or its treatment. The four types of clinical outcome assessments are patient-reported outcomes (PROs, clinician-reported outcomes (ClinRO), observer-reported outcomes (ObsRO), and performance outcomes (PerfO).

Clinician-reported outcomes (ClinRO): clinicians assess and report patient condition.

Glossary

Companion study: A study done alongside a larger trial. It helps answer extra questions about the main trial, such as how the treatment works in real-world settings or how it affects certain groups of patients.

Composite endpoints combine multiple outcomes into a single measure to assess the effectiveness of a treatment in clinical trials.

Creatinine is a waste product in the blood that comes from muscle metabolism (it builds up in the blood from normal muscle use). It's measured in blood tests to assess how well the kidneys are working since one of the functions of the kidneys is to filter creatinine out of the blood. Higher levels can indicate reduced kidney function.

Cystatin C is a small protein produced by nearly all cells in the body and is used as a marker of kidney function. Unlike serum creatinine, cystatin C levels in the blood are less influenced by muscle mass, making it a potentially more accurate indicator of how well the kidneys filter waste.

Endpoint is a targeted outcome of a clinical trial that can help determine how well and how safe a new therapy or treatment is. Endpoints may also be used throughout a study to determine if a participant's risk of continuing to be in a study is too great. Endpoints for a clinical trial may include one or more **clinical outcome assessments and/or surrogate endpoints**.

Epidemiology is the study of how diseases and health conditions spread, what causes them, and how they affect people in a population. It looks at patterns, such as who gets sick, where, and when, to find ways to prevent and control illnesses.

Exclusion criteria are the conditions that would prevent someone from joining the study. These are meant to protect patients by not allowing people who may face higher risks, side effects, or complications to take part in the study.

Generalizability refers to how well a test or measure works for different people and situations outside the original group studied.

GFR estimates how much blood the kidneys filter per minute. It's a key measure of kidney health, with lower GFR values indicating decreased kidney function. GFR is often calculated using serum creatinine, age and sex.

Glossary

Health–related quality of life (HRQOL) is a measure of how well a person can live and enjoy their life despite health issues. It looks at different areas like physical health, mental well-being, social connections, and daily activities.

Hierarchical composite endpoints (HCE) are a type of measurement used in clinical trials that combine multiple outcomes or results into one overall measure, arranged in a specific order (hierarchy). The goal is to capture a range of important effects of a treatment or intervention, from most to least important.

Hyperkalemia is a medical condition where there is too much potassium in the blood. Potassium is a mineral that helps nerves, muscles, and the heart work properly. However, too much potassium can cause serious problems, such as irregular heartbeats, muscle weakness, or even heart failure if left untreated**.**

Inclusion criteria are the conditions a person must meet to be part of the study. Researchers choose patients to be in a study if they have aspects of their health that will show whether a treatment works for other patients with the same issues.

Internal consistency indicates how well different items within a test measures the same concept.

Kidney failure GFR occurs when **GFR** drops to 15 mL/min/1.73 m^2 or lower, indicating that the kidneys are no longer able to function well enough to filter waste and fluids effectively from the blood, or balance body chemicals. At this stage, **KFRT** is usually required.

Kidney function replacement therapy (KFRT) refers to treatments that replace the kidney's function when the kidneys can no longer work adequately. This includes dialysis, which artificially filters the blood, and kidney transplantation, where a healthy kidney from a donor is surgically placed in the patient.

Life participation refers to the ability to participate in meaningful activities of daily living, including work, study, and social recreational activities. This is rarely assessed in clinical trials.

"N Equals 1" trial: This trial focuses on just one person. The treatment is tested on a single patient to see how well it works for them specifically, often by comparing periods when they are on the treatment to when they are off it.

Glossary

Observer-reported outcomes (ObsRO)*:* caregivers or others report on the patient's health.

Patient-reported outcomes (PROs)*:* patients self-report their symptoms or quality of life.

Patient-reported outcome measures (PROMs) are standardised, validated survey tools that assess health outcomes reported by patients. These instruments can be used to assess concepts that may be narrow (e.g., pain intensity) or broad (e.g., **health–related quality of life**).

Performance outcomes (PerfO)*:* objective measurements of a patient's ability to perform specific tasks.

Platform trial: A trial in which many treatments are tested at the same time for a single disease. As new treatments become available, they can be added to the trial without starting a whole new study, making the process faster and more efficient.

Psychometric evaluation is a process used to assess the reliability, validity, and overall performance of psychological tests or measurement tools.

Reliability is the consistency of a measurement tool in producing similar results under consistent conditions.

Time-to-Event Analyses are used in trials to evaluate how long it takes for specific clinical events (i.e kidney failure, dialysis initiation) to occur in participants. This method helps identify the effectiveness of treatments over time and allows researchers to compare the timing of events between different groups. It also accounts for participants who may drop out or not experience the event, providing a clearer picture of treatment impact on patient outcomes.

Win statistics compare the number of wins, losses, or ties between the treatment groups.

Validity is the extent to which a tool measures what it is intended to measure.

Surrogate endpoints are indirect measurements used to predict the actual outcomes of interest. For example, instead of waiting to see the real effects of a treatment (which might take years), researchers measure something that is easier or faster to track, which is believed to be related to the real outcome.