

DCRI THINK TANKS

FROM INSIGHT TO ACTION

Clinical Endpoint Classification and Adjudication: Consolidating Standards and Leveraging Technology

January 27-28, 2021

EXECUTIVE SUMMARY

BACKGROUND

This DCRI Think Tank workshop on clinical endpoint classification (CEC) and adjudication brought together thought leaders from academia, the US Food and Drug Administration (FDA), the European Medicines Agency (EMA), and industry to share expert perspectives and best practices about key topics, including the impact of COVID-19 on event adjudication, emerging therapeutic areas and trends in safety and efficacy assessments; approaches for academic-industry collaboration; innovative uses of new digital approaches for event ascertainment and adjudication; and standards for data quality and compliance.

WELCOME AND OVERVIEW

Cohosts Renato Lopes (Duke University) and Ken Mahaffey (Stanford University) welcomed the attendees. Dr Lopes reminded attendees of the inspiration for the workshop, the inaugural CEC Summit hosted in Chicago in 2018. Dr Mahaffey encouraged attendees to consider practical steps for transforming the field of event adjudication in clinical trials.

SESSION I: COVID-19

Dr Lopes (Duke University) and co-moderator Karen Hicks (FDA) introduced the session. In response to 2 polls, almost all attendees agreed that the COVID-19 pandemic has changed the conduct of clinical trials, and most attendees agreed that the COVID-19 status of participants should be collected in clinical trials and the details included in event adjudication, particularly cause of death.

Panelist Steve Wiviott (Brigham and Women's Hospital) addressed the question of how much COVID-19-related outcome data should be collected in clinical trials. The question is especially relevant for cardiovascular trials, because COVID-19 influences many aspects of the cardiovascular system, leading to potentially higher likelihood of the events being reported. The

European Society of Cardiology has offered guidance on the diagnosis and management of cardiovascular disease during the pandemic (<https://www.escardio.org/Education/COVID-19-and-Cardiology/ESC-COVID-19-Guidance>). Moreover, COVID-19 has emerged as a leading cause of death during certain phases of the pandemic, highlighting the need to consider its contributions to specific clinical events and as a competing risk in outcome analyses.

The appropriate extent of COVID-19–related data collection depends on many factors, including the disease state, patient population, therapeutic interventions, and the primary and secondary outcomes. Considerations include the impact of data collection on trial operations, recruitment, retention, missing data, and data quality. In addition, some trials may need to collect details during adjudication that help will help in evaluating possible causal relationships between COVID-19 and observed outcomes and treatment effects.

The method of data collection should be determined by the information needed and the planned analyses. If the investigators need to know the COVID-19 status of every participant, this information should be collected prospectively as part of the main data collection effort (such as during regular visits or in the case report form). If COVID-19 is a safety outcome or is related to safety outcomes, the information should be collected systematically from the site investigators or included in event adjudication.

Panelist Venu Menon (Cleveland Clinic) described the impacts of the COVID-19 pandemic on the collection and quality of source documents. In ongoing trials, the pandemic has caused reductions in staffing, leading delays in communicating events to CEC committees. Remote work by research staff has contributed to breakdowns in communication between participants and their recruiting sites. The pandemic has also led to a reduction in the clinic visits that often trigger identification of events. Other impacts include challenges in obtaining records, poor clinical documentation and recordkeeping due to staff being focused on COVID-19 surges, delays in event reporting, and decreased site monitoring.

Investigators should consider a variety of solutions for these new challenges. Consider developing a script for study personnel who contact participants for event identification, because these contacts may be the only opportunity to collect data. Consider providing additional education and support to sites and asking participants to signing forms to authorize releases of information. When requesting records for CEC purposes, consider being more expansive than usual, and consider expanding the calling of events and triggering packages, being sure to let sites know to trigger packages in all cases so they can be properly adjudicated.

Panelist Roland Chen (Bristol-Myers Squibb) summarized several direct and indirect impacts of the COVID-19 pandemic on evaluating and interpreting endpoints and clinical events. Direct impacts include challenges assessing contributions of the underlying condition, the intervention, and COVID-19 to the endpoint and assessing effects on patient-reported outcomes, quality of life, and global assessment measurements. Indirect impacts include variable patient access to (and willingness to seek) health care services, which can affect endpoint ascertainment; disproportionate effects by socioeconomic subgroup; effects of

concomitant medications and vaccination; adequacy of management of COVID-19–related complications; and the impacts of using telemedicine tools on endpoint validity. In response to these challenges, it is important to collect data related to potential or confirmed COVID-19; address the impacts of data collection methods on the reliability of the data and validate new methods; capture reasons for noncompliance, discontinuation, missingness, and protocol deviations and consider operational interventions to minimize them; and make thoughtful adjustments to the statistical analysis plan, including sensitivity analyses and prespecified exclusions and rules for handling data. It is important to strive for consistency in study operations while recognizing the need to tailor operations to changing situations.

Attendees further discussed determinations for which data to collect and the timing of adjudication. Dr Hicks referred attendees to the FDA guidance, [“Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency” \(June 2020\)](#). Depending on the population being studied, it may be possible to prespecify an analysis of participants based on COVID-19 status or related characteristics. Bob Temple (FDA) agreed that this approach could be useful in some studies, such as heart failure trials, in which COVID-19 is likely to interfere with the analysis.

SESSION II: MOBILE DIGITAL TECHNOLOGIES AND ARTIFICIAL INTELLIGENCE IN DESIGN AND CONDUCT OF CEC

Dr Mahaffey introduced moderators Leandro Boer (AstraZeneca) and Michael Gibson (Beth Israel Deaconess Medical Center), who welcomed presenter Mintu Turakhia (Stanford University).

Dr Turakhia noted that adoption of digital health tools, including wearable devices, is increasing quickly, including within clinical trials. Growing use of these technologies raises questions about how to validate the tools for data collection. Designing, building, and pretesting the tools and related software stacks is complicated. Using the tools also means building software for CEC and adjudication that pulls data from multiple sources, deidentifying the data, and designing for scalable application. Data flow is another complication; data from wearable devices and other patient-generated data may be less structured (messy and noisy) and higher rates of missingness. Tech-enabled trials can offer the advantages of speed, efficiency, and scale, but they may need to be organized differently. They require trialists to think like software builders to deal with exception processing, “app-plus” technology design, pretesting, engagement, and data flow and compliance. Event adjudication is an additional challenge and requires addressing measurement validity/surrogacy, poor engagement, problems with tech and data connectivity, and data missingness, especially informative missingness.

Attendees discussed the use of artificial intelligence (AI) for algorithmic adjudication of clinical events. There was agreement that these approaches could be used for data extraction but may not be ready for direct identification of events. AI can reduce the human labor required to identify obvious events, but some events with conflicting data or limited information may be

more difficult. It may be worthwhile to adopt a hybrid approach in which AI-based algorithms are used to identify relatively clear events and human adjudication is used for less clear events. Algorithms can be tested and compared with standard approaches of identification and adjudication, and can also be used to identify poor or inconsistent operators and adjudicators. Available evidence suggests that more work is needed to better understand algorithmic assessment.

The discussion raised the possibility of moving away from simple yes-no event identification toward a probabilistic approach in which reviewers identify or adjudicate events by stating a level of confidence that the event occurred.

FIREMOUNT CHAT: WHY DO CEC?

Moderators Roxana Mehran (Icahn School of Medicine at Mount Sinai) and Bob Temple (FDA) facilitated a discussion framed as a lively and informative debate between Scott Solomon (Brigham and Women's Hospital) and Chris O'Connor (Inova Heart and Vascular Institute) on the question of whether CEC is really needed.

Dr Solomon presented the rationale for adjudication: In outcomes trials, benefit and harm are characterized by discrete clinical events. Knowledge of these events is imperfect and depends on the type of study, the experience and expertise of the investigators, and the methods and systems for data collection. Precision matters, because nonmodifiable events obscure signals. Central adjudication offers consistency in endpoint classification, especially in multicenter, multinational trials. Inconsistency adds noise, which may reduce statistical power. Adjudication enforces adherence to endpoint definitions, which is harder to do if investigators alone make those determinations. Independent CEC committees composed of experts might be more essential in studies in which investigators do not have expertise in specific types of events.

Chris O'Connor shared examples of major trials that failed to show a statistically significant result on the primary endpoint on the basis of adjudicated events. In the CHARM Preserved trial, reliance on investigator determination of events would likely have led to approval of candesartan for heart failure with preserved ejection fraction and lower rates of heart failure hospitalization. In PARAGON-HF, inclusion of investigator-reported events along with CEC committee-confirmed events found a significant treatment effect of sacubitril/valsartan vs valsartan alone on the combined primary endpoint of cardiovascular death or heart failure hospitalization. Adjudication adds cost, potentially reduces the number of identified events, and lacks important bedside information that is not captured in case report forms or source documentation.

Attendees noted that CEC committee determinations do not always lead to more conservative adjudication of events compared with investigator determinations. It should not be a yes-or-no question about whether to include adjudication. Rather, it is about appropriate selection and definition of events and endpoints and use of well-designed case report forms. It is about

customizing the adjudication process according to the trial's needs. There are trials for which traditional CEC committee adjudication is appropriate, others for which it may not be needed, and still others for which algorithmic adjudication or a hybrid approach would be useful.

SESSION III: ENDPOINT DEFINITIONS AND ENDPOINTS ON NONCARDIOVASCULAR THERAPEUTIC AREAS

Dr Lopes introduced session moderators Maria Ali (George Clinical in Sydney) and Ernest Spitzer (Cardialysis in Rotterdam). Dr Ali stressed the importance of using both endpoints that are clinically relevant and outcomes that are important to patients. The panelists in this session presented information about the use of CEC and adjudication in therapeutic areas other than cardiology.

Panelist Adeera Levin (University of British Columbia) noted that clinical research in nephrology has lagged behind other therapeutic areas. There has been great variability in definitions of endpoints, which has led to difficulty comparing results between trials, compromises in study design, and issues with ascertainment. For patients and clinicians, this has led to confusion in interpreting results. Therefore, in 2020, nephrologists developed the first ever international consensus definitions of clinical trial outcomes for kidney failure (Levin et al, *Kidney International*, 2020;98:849-859). The result was a composite outcome for kidney failure that includes clinical outcomes (ie, kidney transplant, initiation of maintenance dialysis, and death from kidney failure) and outcomes based on glomerular filtration rate (GFR; ie, sustained low GFR and sustained percent decline in GFR)—including concise, clear definitions for each.

Panelist Ted Leng (Stanford University) described the world of well-defined, highly regulated endpoints in ophthalmology research. Clinical trials typically use the ETDRS Visual Acuity measure with a standardized procedure and certified rooms, charts, and examiners. (Note that this approach would present challenges for remote assessment.) A recently accepted clinical trial endpoint in studies of age-related macular degeneration (AMD) involves imaging of geographic atrophy in “dry” AMD, for which trained image graders work in centralized reading centers with a highly standardized process (usually 2 graders with a third who resolves disagreements). Optical coherence tomography is used for analysis of “wet” AMD. Geographic atrophy growth marks progression to wet AMD and can be predicted using models based on automated feature extraction.

Panelist Warren Capell (CPC Clinical Research) presented definitions used in peripheral artery disease, for which the outcomes of interest are limb ischemia, amputation, and amputation-free survival. Limb ischemic events occur on a spectrum analogous to cardiac ischemic events. Acute limb ischemia (ALI) involves an acute thrombotic occlusion of an artery threatening the limb, and outcomes are determined by time to acute reperfusion. Various outcome definitions for ALI, amputation, and “major adverse limb events” have been used in clinical trials, with little difference in signal. The Peripheral Academic Research Consortium (PARC) has developed the most comprehensive definitions (Patel et al, *Journal of the American College of Cardiology*,

2015;65:931-941. The GLASS guidelines (Conte et al, *Journal of Vascular Surgery*, 2019;69:3S-125S.e40) define “chronic limb-threatening ischemia” and are intentionally more inclusive, but they introduce more heterogeneous biology and may not be as useful for clinical trials. More work is needed to understand and better define these clinical entities and etiologies for CEC purposes.

Moderator Ernest Spitzer summarized the discussion. There is a need for continued standardization. Kidney specialists have made significant progress with the recent consensus definitions for kidney failure. Ophthalmic endpoints rely heavily on image processing and computational stratification. It will be interesting what other endpoints can be developed from these approaches. In peripheral artery disease, further standardization is clearly required, considering the complex biology of the disease.

SESSION IV: ACADEMIC AND INDUSTRY PARTNERSHIPS – DEFINING INDEPENDENT CEC

Dr Mahaffey introduced session moderators Connie Hess (CPC Clinical Research) and John Strony (Janssen).

Panelist Bram Zuckerman (FDA) discussed the independence of CEC, with a focus on cardiac device trials. These trials usually are unblinded to investigators and patients. Objective event adjudication is essential to limit bias in investigator and patient reporting of events, foster consistent event reporting across sites, and promote complete and accurate ascertainment of all endpoints. Independent CEC and adjudication means that committee members are qualified experts without conflicts of interest who are not under inappropriate sponsor or investigator influence. A well-written CEC charter is essential; it should describe the qualifications of members; the events to be adjudicated; the methods for identification of events; the minimum data required for adjudication; event preparation for CEC committee review; adjudication methods; the process for reaching a decision when there is disagreement or the minimum required data are missing; and quality-assurance procedures.

Panelist Claes Held (Uppsala University) extended the discussion by addressing the qualifications of CEC adjudicators and considerations for including CEC chairs and potentially adjudicators as coauthors on primary study manuscripts.

Attendees discussed the benefits and challenges of including research fellows in adjudication, discordance between adjudicators of varying experience levels and between CEC committees and sites, and the different models for how CEC committees are structured. Attendees expressed interest in developing standards for CEC procedures across organizations and in producing case studies of CEC and adjudication conventions.

Panelist Lynn Perkins (Duke University) summarized several considerations for quality control in CEC activities. Audits by sponsors and internal audits keep the clinical research organization aligned with the sponsor, help the organization prepare for inspections, and keep teams up to

date on documentation. However, more regulatory guidance is needed. Regulatory agencies provide excellent guidance to research teams regarding study start-up but offer little support for effective study execution between start-up and database lock. Quality control targets for discordance rates between adjudicators and between CEC committees and sites should be set and monitored. Attendees from the FDA expressed interest in further discussion about what is happening on various CEC teams in order to develop standardized and meaningful criteria and targets.

WRAP-UP

Drs Lopes and Mahaffey ended the workshop by thanking the moderators, presenters, attendees, and the DCRI Think Tank group. They noted the important opportunity to create working groups to more fully address certain issues raised during the sessions, including (1) models, standards, and best practices for CEC procedures across organizations; (2) independence of CEC activities; and (3) regulatory guidance for CEC teams. There may also be interest in collaborative efforts to work on endpoint definitions that contemplate various therapeutic areas in order to make available the right endpoints for the right trials.

Attendees will also be invited to participate in drafting a white paper from the workshop.