Abstracts from the SCT 42nd annual meeting (2021)

Session Proposals

SP-1
ETHICS AND FUTURE OF HUMAN CHALLENGE TRIALS WITH SARS-COV-2 TO SPEED VACCINE DEVELOPMENT

Organizers:
SEEMA SHAH
Northwestern University
CHARLES WEIJER
Western University

Contributor(s):
META ROESTENBERG
Leids Universitair Medisch Centrum
ALASTAIR FRASER-URQUHART
1Day Sooner
GAGANDEEP KANG
Christian Medical College

Introduction: Early in the COVID-19 pandemic, human challenge trials with SARS-CoV-2 were proposed as a way to develop a vaccine faster, and thereby save lives. But would such trials in fact speed vaccine development? Is it ethical to infect healthy volunteers with wildtype SARS-CoV-2? And going forward, what should be the role of human challenge trials in the development of SARS-CoV-2 vaccines? This 60-min panel session, chaired by Professor Seema Shah, will explore scientific and ethical issues in SARS-CoV-2 human challenge trials. The panel brings together key stakeholders, including scientists working with a variety of challenge models, research ethicists, and a member of the advocacy group 1Day Sooner who recently participated in a challenge trial. The panel comprises five talks followed by a 10- to 15-min discussion period.

Background: The origins of human challenge trials can be traced back to efforts in the 1790s to determine whether cow pox inoculation protected against smallpox. In the last 50 years, human challenge trials have enrolled thousands of healthy volunteers to be exposed to dozens of diseases, from the common cold to cholera. Challenge trials are used to better understand how a pathogen causes disease, how it is transmitted (and how transmission may be mitigated), and whether a vaccine candidate confers protection against it. In vaccine challenge trials, healthy participants are randomized to receive either vaccine or placebo, and they are subsequently exposed to the infectious agent. If less people who received the vaccine, compared to those who received placebo, become infected, this provides evidence that the vaccine is efficacious. Challenge trials have played important roles in developing vaccines for cholera, typhoid, and malaria. To protect volunteers, in the modern era, scientists have generally restricted their use of such trials to diseases that are well-understood, or for which there is curative treatment. In the face of the global COVID-19 pandemic, human challenge trials with SARS-CoV-2 were proposed as a means to speed vaccine development and save lives. But are such trials ethical? There are risks associated with COVID-19, including in young people, such as myocarditis, stroke, long COVID, and death. Proponents argue ethical norms are too conservative and may be set aside in a pandemic, provided informed volunteers are willing to participate. What should be the role of challenge trials in the development of a SARS-CoV-2 vaccine? While challenge trials were initially proposed as a substitute for phase 3 efficacy trials for vaccines, challenge trials take time to set up and did not keep pace with dramatically accelerated vaccine development in the COVID-19 pandemic. They also have limitations, including generalizability to older adults or people with comorbidities, since they test vaccine efficacy in 15-150 healthy young adults. Nevertheless, challenge trials could test the second round of vaccine candidates, which is important if the first vaccine approved is not very efficacious or safe or to ensure widespread global access of vaccines. Our panel comprising a healthy volunteer, scientists working with a variety of challenge models, and research ethicists will explore and discuss these ethical and scientific issues.

Talk 1: Human challenge trials with SARS-CoV-2
Meta Roestenberg (Leids Universitair Medisch Centrum, The Netherlands) will introduce human challenge trials and their use in vaccine development. She will discuss the development of a SARS-CoV-2 challenge model, and the role of challenge trials in the development of second-generation SARS-CoV-2 vaccines. Bio: https://www.lumc.nl/org/parasitologie/medewerkers/metaroestenberg
Talk 2: The volunteer experience in a SARS-CoV-2 challenge trial Alastair Fraser-Urquhart (1Day Sooner) participated in the UK SARS-CoV-2 challenge trial in Spring 2021, and he will describe his experience in the trial. 1Day Sooner is a grass roots organization that has gathered over 38,000 people globally who have expressed a willingness to participate in SARS-CoV-2 challenge trials. Bio: https://1daysooner.org/

Talk 3: Developing ethical guidance for SARS-CoV-2 challenge trials Seema Shah (Northwestern University Feinberg School of Medicine, USA; session co-chair) will describe the development of ethical guidance for SARS-CoV-2 challenge trials published in Science and Vaccine in 2020 and the World Health Organization’s key criteria document. She will outline the main ethical considerations in such trials and explain the most compelling potential social value they could provide. Bio: https://www.feinberg.northwestern.edu/faculty-profiles/az/profile.html?id=42926

Talk 4: A critical perspective on the ethics of SARS-CoV-2 challenge trials Charles Weijer (Western University, Canada; session co-chair) will examine critically the case for SARS-CoV-2 challenge trials. He will focus on the social value of challenge trials, whether benefits are justified by risks to volunteers, the use of high-risk groups in such trials, and their impact on public trust in a SARS-CoV-2 vaccine. Bio: https://www.charlesweijer.com/about-me

Talk 5: The future of challenge trials in SARS-CoV-2 and future pandemics Gagandeep Kang (Christian Medical College, India) will consider the future of challenge trials in the future of SARS-CoV-2 vaccine development and, indeed, their role in future pandemics. The pace of vaccine development and clinical trials in the COVID-19 pandemic is unprecedented. Does this undermine the need for challenge trials? If not, what ought to be their role in vaccine development? Bio: https://www.charlesweijer.com/about-me

Format: This 60-min panel session will be chaired by Professor Seema Shah. Each of the talks will be 8- to 10-min in duration and they will cover the volunteer experience, the science of challenge trials, and ethical issues. The five talks will be followed by a highly interactive 10- to 15-min discussion period. The discussion will be divided into two sections. The first half of the discussion will explore the ethics of SARS-CoV-2 challenge trials. The second half will discuss the future role of challenge trials in vaccine development.

Contributor(s):

ROBIN BOINEAU
National Center for Complementary and Integrative Health

ANNE TRONTELL
Patient Centered Outcomes Research Institute

WYNNE NORTON
National Cancer Institute

Managing Pragmatic Clinical Trials in the Time of COVID-19

The clinical research enterprise across many disciplines is expanding rapidly to incorporate methods and best practices for conducting pragmatic clinical trials. Such trials can embed interventions within routine clinical care and enable testing of interventions in real-world settings. The sudden onset of the COVID-19 pandemic, however, has resulted in significant disruptions in US healthcare delivery within systems, and presented significant challenges for planning and conducting pragmatic clinical trial. Unique strategies for design and analysis, intervention adaptation, and stakeholder engagement have been considered across several National Institutes of Health- and Patient-Centered Outcomes Research Institute-funded programs. Consideration of dissemination and implementation issues during the pandemic have also raised challenges and opportunities for pragmatic clinical trial researchers. This session will provide examples of how these issues are being addressed in the context of National Institutes of Health and Patient-Centered Outcomes Research Institute projects.

Moderator—Qilu Yu, PhD, National Center for Complementary and Integrative Health (NCCIH)

Design and Statistical Analysis Considerations for Pragmatic Trials during the COVID-19 Pandemic—Qilu Yu, PhD, NCCIH. During the COVID pandemic, pragmatic clinical trials embedded in the healthcare systems face special challenges for both design and statistical analysis. National Institutes of Health programs like the Health Care Systems Research Collaboratory have noted several examples of pandemic induced changes in trial design, including indefinite interruptions in stepped-wedge design trials; post-randomization modifications in number of clusters and/or cluster size in cluster-randomized trials; as well as new considerations of clustering effects due to requisite switching of planned in-person to virtual delivery of interventions. To deal with COVID-19 pandemic impacts, analytic plans may need to consider heterogeneity of treatment effects and COVID-19 confounding with timing/phases of trials. Also, strategies must consider the likely enhanced missing data fraction caused by pandemic disruptions. Several corresponding analytic strategies for pragmatic clinical trials will be discussed, including subgroup analysis of heterogeneity of treatment effect, analysis of modality of intervention...
delivery and dosing and time-specific mediation analysis of the COVID-19 impact.

Pragmatic Clinical Trial Adaptations During the COVID-19 Pandemic—Can We Do This?—Robin Boineau, MD, MA, NCCIH. With the advent of the COVID-19 pandemic, embedded pragmatic clinical trial investigators have considered multiple issues in conducting their trials. Lessons from both the National Institutes of Health’s Health Care Systems Research Collaboratory and the National Institutes of Health/VA/DoD Pain Management Collaboratory have provided insights on maneuvering the disrupted landscape for research embedded into routine clinical care. At first pass, teams needed to address whether stopping study recruitment or study intervention was necessary for the safety of investigators, staff and subjects or required due to closing of healthcare system clinics and pivoting resources to COVID-19 pandemic. Other critical issues included how continuing recruitment would be feasible and under what conditions; potential changes in delivery of interventions from in-person to fully remote; determining when pilot testing of remote interventions would be needed; and how to alter randomization schemes if sites or treatment arms could not provide remote interventions. The likely sustained heterogeneity of intervention delivery across sites and systems also needed to be addressed by teams. Changes to remote interventions frequently required shifting to remote data collection methods as well, including clinical outcome data, patient-reported outcomes, safety data, and bio-samples. Strategies implemented by several trials included pivoting to mail-in testing kits, testing kits with remote viewing of results, video or phone calls, and providing Internet access to subjects with limited access.

To Adapt or Not to Adapt—What is the Role of Methodologists and Stakeholders in Managing Trial Disruptions. Due to the COVID-19 Pandemic?—Anne Trontell, MD, MPH, Patient Centered Outcomes Research Institute. Like other funding institutions, the Patient Centered Outcomes Research Institute is confronting significant disruptions to its portfolio of pragmatic comparative clinical effectiveness trials. In immediate response to suspensions and slowdowns of research activities early in the pandemic, Patient-Centered Outcomes Research Institute instituted rapid response reviews to assist funded researchers in making immediate adjustments to recruitment, informed consent, intervention delivery, trial operations, and data collection. As research institutions and local pandemic control measures have fluctuated over time, Patient-Centered Outcomes Research Institute systematically examines proposed adaptations to determine whether they represent an acceptable modification to the original study design both scientifically and to the involved patient and stakeholder communities. Proposed changes are reviewed using Patient-Centered Outcomes Research Institute’s Methodology Standards with focused attention to standards for complex interventions and causal inference as well as for their acceptability and feasibility to participants and stakeholders. For obligatory in-person interventions which cannot otherwise be carried out, testing of alternative or remote delivery may be allowed to assess if they can be substituted with acceptable validity. Other review considerations include trial status at the time of disruption, possible selection biases due to access differences to a revised intervention or its delivery method, intervention validity and fidelity with modification(s), variation across sites in their regulatory permissions and ability to accommodate changes, and impacts upon the statistical analysis plan and planned subgroup analyses. Analysis of Patient-Centered Outcomes Research Institute’s experience with adaptations will be directed to examination of themes and approaches to discern best practices which may be generalized.

How Can Implementation Science Help During the Pandemic?—Wynne Norton, PhD, National Cancer Institute. With the rapidly changing healthcare delivery landscape during the COVID-19 epidemic, it may be more difficult to both implement pragmatic trials and study the implementation of interventions embedded within pragmatic trials. Implementation science—the scientific study of methods and strategies needed to implement evidence-based interventions into routine care delivery settings—can offer both scientific and practical guidance for ongoing trials and the future implementation of results of such trials. Using an implementation science lens can help better understand and measure context, including adaptations to intervention components and delivery options, trial components, data collection methods, patient-, provider-, and system-level re-engagement in trials, de-implementation of trial components, and strategies for translating and sustaining trial results in practice. Implementation science offers an important perspective to help optimize the conduct of pragmatic trials and the integration of trial results into real-world settings.

SP-3
DESIGN, IMPLEMENTATION, AND ANALYSIS OF N-OF-1 CLINICAL TRIALS

Organizer:
SONIA JAIN
University of California

Contributor(s):
SAMA SHRESTHA
Pharmapace, Inc.
JIABEI YANG
Brown University
**Description of Proposed Invited Session:** N-of-1 clinical trials are randomized, multi-period crossover trials conducted within a single subject to evaluate an individual’s heterogeneity of treatment effect. In an N-of-1 trial, a single subject will serve as one’s own control. These trials are best applied to chronic conditions, in which treatments have an immediate but stable response with minimal washout requirements. By identifying an individual’s response to the intervention, a personalized treatment plan can be administered. Although N-of-1 trials date back to the early 20th century, the popularity of N-of-1 trials, as a patient-centered, clinical decision-making tool has only recently surged in the era of precision medicine, and likely now, in the pandemic era. In this session, we present novel clinical trial design, implementation, and statistical analysis approaches to the N-of-1 trial paradigm.

**Invited Session Format:** The session will consist of four invited talks, each about 18-20 min.

**Talk #1 Title:** A Bayesian-Bandit Adaptive Design for N-of-1 Clinical Trials. Authors: Sama Shrestha and Sonia Jain. Brief Description: We propose a novel Bayesian adaptive design for both the individual and aggregated N-of-1 trials using a multi-armed bandit framework that is estimated via efficient Markov chain Monte Carlo. A Bayesian hierarchical structure is used to jointly model the individual and population treatment effects. Our proposed adaptive trial design is based on the Thompson Sampling, which randomly allocates individuals to treatments based on the Bayesian posterior probability of each treatment being optimal.

**Talk #2 Title:** nof1: an R package for analyzing and presenting n-of-1 trials. Authors: Jiabei Yang and Christopher H Schmid. Brief Description: Combining a series of n-of-1 trials gives extra information for estimating the population average treatment effect by repeatedly measuring participants compared with evaluating participants only once in randomized controlled trials. When participants are assigned to different sets of treatments in a series of n-of-1 trials, the combined trials form a network of treatments where the average treatment effects are estimated from both the direct comparisons within individuals and from the indirect comparisons through the common treatment across individuals. There is currently no standard tool for analyzing and presenting the results for n-of-1 trials. We developed an R package, nof1, for the analysis and the result presentation of individual and a series of n-of-1 trials. We use nof1 to analyze the n-of-1 trials that evaluates the comparative effectiveness of the specific carbohydrate diet and the modified specific carbohydrate diet to usual diet on patients with inflammatory bowel disease.

**Talk #3 Title:** Identifying Individual Goal Achievement Behavior Using Bayesian Networks in The Just Walk Study. Authors: Christian B Pascual, Eric Hekler, and Sonia Jain. Brief Description: Prescribing daily activity goals can motivate individuals to be healthy and active, but achievement of these goals is influenced by both intrinsic and extrinsic factors. We used Bayesian network analysis to identify factors that contribute to or against goal achievement in the completed Just Walk N-of-1 health behavior study.

**Talk #4 Title:** Fixed or Random Intercepts: Does It Make a Difference? Authors: Kexin Qu and Christopher H Schmid. Brief Description: Many meta-analytic models require a multi-level framework in which intercepts may be considered fixed or random. These include models for binary arm-level outcomes, control-risk meta-regression, network meta-analysis, and meta-analysis of a series of N-of-1 trials. Choice of the type of intercept is controversial. Random intercepts facilitate predictions to new studies or new individuals in the case of N-of-1 trials or individual participant data with repeated measurements. They also allow for borrowing of strength across individual units to improve predictions on units observed. But the resulting shrinkage of the intercepts can cause changes in the estimates of the treatment effects themselves. This unsettles many analysts who believe that the randomized estimates should be independent of study prevalences. Because the number of parameters in a fixed intercept meta-analysis model is increasing at the same rate as the sample size, a key principle that underlies the validity of maximum likelihood estimation is violated. This study investigates the sensitivity of Bayesian model parameters to different parameterizations of the multi-level model with fixed or random intercepts. It applies the findings to pairwise meta-analysis, network meta-analysis, control risk meta-regression, and aggregated N-of-1 trial data.
The SCT was founded in 1978 as an interdisciplinary organization dedicated to the development and dissemination of knowledge about the design, conduct, analyses, and reporting of government and industry-sponsored clinical trials and related healthcare research methodologies.

In this session, following an introduction to the work of the Society by the President of the Society, committee chairs will give brief updates on the work of the Program, Education, Communications, Nominations, Development, Membership, Fellows, and Trial of the Year committees. Find out more about the life of the Society!

SP-5
DESIGN AND IMPLEMENTATION OF A VACCINE TRIAL IN A PANDEMIC—MRNA-1273 PHASE 1 TRIAL
Organizer:
JAMES ALBERT
The Emmes Corporation
Contributor(s):
MAT MAKOWSKI
The Emmes Corporation
KAITLYN CROSS
The Emmes Corporation
SEMHAL SELAMAWI
The Emmes Corporation
AVA MANOKIAN
The Emmes Corporation
BRITTA FLACH
National Institutes of Health

Proposed Session Type
This session will use examples from an mRNA-1273 phase 1 COVID-19 vaccine trial to explore reporting and operational considerations for vaccine trials in the pandemic setting. The session is composed of three talks from members of the Statistical and Data Coordinating Center for the mRNA-1273 phase 1 trial followed by two discussants and a question-and-answer session.

The mRNA-1273 phase 1 trial evaluated the safety and immunogenicity of several doses of the mRNA-1273 vaccine. Three age groups were enrolled (18–55, 56–70, and 71 years or older) and followed for safety and immunogenicity. At the time of this writing, subjects are completing Day 209. The results of this trial combined with the phase 2 trial were used to justify the 100 µg dose used in the phase 3 trial.

The expectation is that all Day 209 follow-up will be completed and the results published by the time of the session in May 2021. The speakers and discussants will share knowledge gained from their roles and will address questions from the audience about the design and conduct of the trial as well as the process of reporting trial results. This session is highly relevant to the meeting theme and is expected to be of interest for a wide audience.

Talk one will provide a brief background of the mRNA-1273 vaccine study platform and will highlight issues that arise at the start of the clinical trial process during an outbreak. Ava Manokian, Data Manager, will address considerations for adopting accelerated timelines and quickly getting a trial launched from concept to first subject enrollment.

Talk two will focus on strategies for handling enrollment, data entry, and data cleaning during trials in outbreak settings. Semhal Selamawi, Data manager, will discuss how the changing COVID outbreaks, protocol changes to add dose levels and age cohorts, and the immediate need for clean data presented challenges to data management.

Talk three will introduce the issues related to trial reporting. The pandemic setting raised difficult questions about disseminating the results of clinical trials. Kaitlyn Cross, Senior Biostatistician, will describe all the major reporting occurring during the trial and how the study group navigated the large numbers of reports required in a short time. Topics will also include different assays that are used and how they are handled.

These three speakers will be followed by talks from two discussants who will provide additional perspectives on this vaccine study.

Discussant one, Britta Flach, Director Advanced Clinical Testing, will provide a perspective on the challenges associated with running a lab, specimen receipt and analysis, and reporting results.

Discussant two, Dr Mat Makowski, will discuss the challenges of working on this high-profile trial and how all the different pieces fit together from a logistics and analytical standpoint.

Our expectation is that these talks will generate many questions; hence, the remainder of the session is devoted to a panel style question-and-answer session with the speakers and discussants.
The predominant reporting of clinical trial results today is what we might call the “regulator’s analysis.” This analysis meets requirements that regulators have set for their work. Shouldn’t there also be an analysis for other stakeholders, for example, physician’s analysis, a third-party payor’s analysis, a patient’s analysis, and other analyses? Different analyses for different stakeholders appearing in different journals would allow Bayesian analyses, simulation results, effectiveness data, real-world evidence, more sophisticated safety analyses, and so on to also be published. The need for analyses for different stakeholders has been heightened by the various stakeholders involved in the analysis of COVID-19 trials and the current attention to estimands and concern about the proportional hazards assumption in survival analysis. This session will present the case that “persuasive evidence” is defined differently by different clinical trial stakeholders and provide some examples of innovative clinical trials analyses.

**Chair and Org:** Greg Ball, Merck [greg.ball@merck.com]

**Title:** Is There Only One Analysis? Jay Herson, Johns Hopkins Bloomberg School of Public Health [jay.herson@earthlink.net]. Summary: The main output of clinical trials is to yield persuasive evidence. However, our notion of persuasive evidence has changed over time. The intent-to-treat frequentist analysis preferred by regulatory agencies may no longer be considered persuasive to other stakeholders. This talk will motivate the need for and provide examples of additional analyses that are useful for other stakeholders such as physicians, patients, and insurers.

**Title:** Safety Analysis: A Physician’s Perspective Barbara Hendrickson, AbbVie [barbara.hendrickson@abbvie.com] Summary: Clinical trial safety data presentations are typically confined to rates of different categories of adverse events (e.g. common, fatal, serious, and those leading to discontinuation). However, other types of questions are important to healthcare providers and patients. Specifically, information about time to onset, reversibility with and without intervention, and additional measures of severity by functional impairment or treatments required is frequently asked questions. Also, more meaningful risk factor analyses as well as information about the likelihood for key adverse reactions in different patient subpopulations (e.g. by age and renal impairment) are needed.

**Discussant:** Elizabeth Garrett-Mayer, American Society of Clinical Oncology [Liz.Garrett-Mayer@asco.org].

**SP-7**

LEVERAGING EXTERNAL EVIDENCE IN CLINICAL RESEARCH USING NOVEL STATISTICAL METHODS

**Organizer:** Saurabh Mukhopadhyay AbbVie

**Contributor(s):**

NA Cai Astellas

Jessica Lim GSK

Emmanouil (Manos) Spanakis AbbVie Deutschland GmbH & Co. KG

Mareike Bereswill AbbVie Deutschland GmbH & Co. KG

Conducting rigorous and ethical clinical trials is particularly difficult in pediatric patients and rare disease populations. In these limited populations, it is often difficult to conduct fully powered and adequately controlled randomized clinical trials. Advancing rigorous studies in those as well as in larger populations has
become even more challenging in the pandemic era. Using external data is very appealing to address such challenges.

Many diseases affect both adults and children, but treatments are often developed and tested only in adults. Conducting trials in children is often very challenging for various reasons. For example, diseases may have a low incidence in children, and it is unethical to expose more children than absolutely required to either an experimental treatment or no treatment. Parents are also reluctant to enroll their children in an experiment where too much is unknown.

Many of these problems are also present in rare diseases. There are more than 7000 different rare diseases affecting 25–30 million Americans. For many of these conditions, a well-controlled clinical trial to study new medicines is not feasible. Due to ethical considerations and to reduce subjects’ fears of being randomized to placebo, many clinical trials reduce the size of control arms or proceed without control arms altogether.

Novel statistical approaches can ease some of these concerns by developing methodological frameworks to allow borrowing and quantitatively incorporating external evidence when evaluating new medicines. There are many existing and new methodological and operational questions that remain to be investigated. For example, where to borrow from, how much to borrow, reproducibility of evidence, and so on.

This session will investigate novel and flexible approaches of leveraging external data and explore some of the important scientific and implementation questions. The goal is to reduce subjects’ burden in evaluating new medicines in clinical research. Specifically, the session proposes to have three presentations on outstanding challenges and methodologies with applications to address some of those challenges by leveraging external evidence in clinical trials. The presentations will be followed by discussions and Q&A.

Presentations:

1. Title of the presentation: “A snapshot of facts on rare disease and recent drug development in rare disease” Speaker: Na Cai, PhD. Affiliation: Associate Director, Biostatistics, Astellas. Email: na.cai@astellas.com.

2. Title of the presentation: “Reducing Patient Burden in Clinical Trials Through the Use of Historical Controls: Appropriate Selection of Historical Data to Minimize Risk of Bias” Speaker: Jessica Lim, MA. Affiliation: Director, Clinical Statistics, Biostatistics, GSK. Email: jessica.w.lim@gsk.com

3. Title of the presentation: “A Bayesian framework on borrowing data from historical controls—Experiences at AbbVie” Speakers: Emmanouil (Manos) Spanakis, MSc and Mareike Bereswill, MSc. Affiliation: AbbVie Deutschland GmbH & Co. KG, Data and Statistical Sciences, Ludwigshafen, Germany. Emails: emmanouil.spanakis@abbvie.com, mareike.bereswill@abbvie.com. Collaborators: Martina Kron, PhD. Affiliation: AbbVie Deutschland GmbH & Co. KG, Data and Statistical Sciences, Ludwigshafen, Germany; Saurabh Mukhopadhyay, PhD. Affiliation: Data and Statistical Sciences, AbbVie

Chair: James Myles, PhD, Executive Director Biostatistics TA Head Medical Specialties Astellas. Email: james.myles@astellas.com

SP-8

BENEFIT–RISK ASSESSMENT IN CLINICAL TRIALS WITH INDIVIDUAL PATIENT NEEDS IN MIND

Organizer:
SHAHRUL MT-ISA
MSD

Contributor(s):
DEBORAH ASHBY
Imperial College London

RUTH OWEN
London School of Tropical Hygiene and Medicine

RICHARD BAUMGARTNER
Merck & Co.

Introduction: This session focuses on the recent developments, challenges, and opportunities in medicine in clinical trials. In the era of pandemic when time is precious, and resources are scarce, the use of statistical methods when characterizing individualized benefit–risk balance may help to better elucidate delicate benefit–risk balance between health technologies, and to clarify priorities.

Background: It is known that benefits and risks are multi-dimensional, and these can vary between patients receiving the same treatment. Often the population-level benefit–risk balance is sufficient and acceptable in regulatory decision-making, but in some cases, it is not ideal. When thinking about how individual clinical trial subject respond to treatment, there may be occasions when, albeit rare, that some subjects may benefit without any adverse effect, and vice versa. A population-level benefit–risk analysis would miss this nuance without a rigorous assessment that considers the benefit–risk balance at the individual level.

Statistical uncertainty, among others, also plays a key role in the confidence to determine the true patient outcome and their benefit–risk balance. Therefore, while deterministic analysis may dominate the field of benefit–risk assessment, accounting for uncertainty is also a crucial aspect when it comes to making informed treatment decisions.

In this session, our panel of prominent experts will discuss the merits and challenges of development and
implementation of statistical methodologies in this complex but important area. With a growing number of initiatives and resources, such as, from the US Food and Drug Administration, European Medicines Agency, the Council for International Organizations of Medical Sciences, Innovative Medicines Initiative, American Statistical Association, and the European Federation of Statisticians in the Pharmaceutical Companies, our session would contribute to the active discussions and debates on this topic.

Talk 1: Ruth Owen will present a novel method to assess individual patient’s benefit–risk trade-off with some examples of application using clinical trials data. Multi-variable predictive models for efficacy and safety outcomes are used to quantify the absolute treatment effects on benefit and harm for any individual patient’s profile, therefore facilitating clinical judgment as to which patients have (and not have) a favorable benefit–risk trade-off.

Talk 2: Deborah Ashby will continue the session with key foci on the multi-dimensional aspects of benefit–risk assessment and the importance of addressing and managing uncertainty. The Bayesian statistics paradigm provides the natural framework here, and subsequently makes elegant probabilistic statement.

Discussions and Dialogues: Dr Baumgartner will then provide commentary and extend the dialogues on these emerging methodologies in clinical trials and precision medicine, and the preparedness of sponsors, regulators, and the scientific community to welcome them, as well the transparency they bring to the benefit–risk assessments, especially when patient perspectives and/or preferences may alter the optimal treatment decision.

Target Audience: The topic would be of particular interest to clinicians, pharmaceutical companies, sponsors, regulatory agencies, health technology assessment bodies, payers, and academic researchers in this area.

Disclaimers: The views and opinions presented in this session are of the speakers only, and do not represent their institutions/companies or the views of the pharmaceutical industry.

SP-9

ADAPTIVE COVID-19 TREATMENT TRIAL DATA MANAGEMENT: ENSURING SUCCESS IN A CONSTANTLY-EVOLVING, HIGH-DEMAND PANDEMIC ENVIRONMENT

Organizer:
ASHLEY WEGEL
The Emmes Company

Contributor(s):
ASHLEY BOWERSOX
The Emmes Company, LLC

KELLY CLARK
Duke Human Vaccine Institute

Remdesivir (GS-5734) was identified as a promising therapeutic candidate for COVID-19, at a time when no therapeutic agents had shown to be efficacious for the treatment of coronavirus disease 2019 (COVID-19). A series of rapid randomized, double-blind, placebo-controlled Adaptive COVID-19 Treatment Trials were developed, all written into one Protocol document, to evaluate the safety and efficacy of novel therapeutic agents in hospitalized adults diagnosed with COVID-19. The first study (ACTT-1) was a multi-center trial conducted in 73 sites globally (N = 1062), that demonstrated remdesivir shortened the time to recovery in adults who were hospitalized with COVID-19 and had evidence of lower respiratory tract infection. This presentation focuses on the unique challenges presented by the high-profile Adaptive COVID-19 Treatment Trial and how the data management team ensures its success in this pandemic environment.

Typically, a data management team may take several weeks or months to design and build a database upon receipt of a new protocol but due to the high priority of the Adaptive COVID-19 Treatment Trial study, the database was initially built in 2 days. Having such a short turnaround time meant the team had to focus on critical items such as randomization and forms that collected data related to safety signals. Flexibility was also important since study sites were being identified and activated on a rolling basis, and many were international which meant various formatting differences, specifically with laboratory results. For both internal and external stakeholders to ensure proper communication and organization, it was crucial to maintain a proper balance of collaboration yet still identifying clear leaders of functional groups to avoid having “too many cooks in the kitchen.” Scheduling frequent meetings and utilizing collaboration software, rather than relying on email communication, helped all parties to brainstorm and resolve issues efficiently.

The Adaptive COVID-19 Treatment Trial was divided into multiple stages (called ACTTs) where each new Adaptive COVID-19 Treatment Trial introduced a different therapeutic combination. Each Adaptive COVID-19 Treatment Trial was designed to enroll over 1000 trial participants across 100+ sites, and it was not uncommon for the stages to overlap, creating many challenges with data management. Data queries were often sent to sites weekly or even multiple times a week, and data managers would go above and beyond by working nights and weekends to accommodate the high
workload and shortened timeframe. Safety and endpoint data were the top priority during data cleaning and the volume of queries was not just a challenge for the data management team but also for the sites that had to juggle the trial in addition to clinically managing their COVID patients who had a brand new disease. Meeting deadlines was essential but doing so at the expense of the cleanliness of the data was not an acceptable option, so new processes had to be implemented in order to avoid this from occurring.

It was also critical to be “audit-ready” at all times due to the high-profile nature of the study, so an internal team was dedicated to ensuring that all documentation was present and complete throughout the course of the study. Mock audits were held so that issues could be resolved a Food and Drug Administration audit.

In this presentation, four speakers will introduce the study in more detail and demonstrate how being more adaptable and communicative enabled the data management team, in collaboration with the other study stakeholders, to successfully manage the trial. Each presentation will last approximately 20 min and they are listed below in more detail and in chronological order. The remaining 10 min will be devoted to a question and answer session with audience members.

National Institutes of Health Clinical Project Manager Ms Wegel is a Clinical Project Manager at the National Institute of Allergy and Infectious Disease who, along with several other Clinical Project Managers, was responsible for general oversight of study conduct and contractors on the Adaptive COVID-19 Treatment Trial. Ms Wegel will present site management and data management strategies for meeting incredibly demanding timelines required by sponsor and activation of study sites as well as attention to the unique needs of the international sites participating in the study.

Ashley Bowersox, MPH (Emmes) Ms Bowersox is a Senior Data Manager in the Vaccine and Infectious Diseases group, supporting the Clinical and Epidemiological Studies in Infectious Diseases. She served as the Forms Design Lead and DM Project Manager on the Adaptive COVID-19 Treatment Trial. Ms Bowersox will discuss the ACCT trial and provide some background information, as well as briefly explain the study objectives, design, and procedures.

Ashley Wegel, CCRA, MEd (Emmes) Ms Wegel is an Associate Project Leader in the Vaccine and Infectious Diseases group, supporting the Clinical and Epidemiological Studies in Infectious Diseases. She served as the Forms Design Lead and DM Project Manager on the Adaptive COVID-19 Treatment Trial. Ms Wegel will present site management and data management strategies for meeting incredibly demanding database freeze and lock timelines required by sponsor to support safety reviews and publications and continuously improve upon forms design and interface for each new stage. Kelly Clark (Duke Human Vaccine Institute) Ms Clark is a Clinical Research Coordinator in the Duke Vaccine Trials Unit at the Duke Human Vaccine Institute, one of the enrolling sites on Adaptive COVID-19 Treatment Trial. Ms Clark will discuss the challenges of working on a rapid study with tight timelines, requiring quick responses to a large volume of research data queries, while clinically managing their COVID patients who had an unfamiliar and daunting new disease.

SP-10
SAMPLE SIZE RE-ESTIMATION: REFLECTIONS ON IMPLEMENTATION AND EXTENSIONS TO COMPLEX DESIGNS
Organizer:
ERINN M HADE
The Ohio State University
Contributor(s):
LESLEY A MCCLURE
Drexel University
VALERIE DURKALSKI-MAULDIN
Medical University of South Carolina
CHRISTOPHER S COFFEY
The University of Iowa
In the past 20 years, sample size re-estimation procedures, both blinded and un-blinded, have been developed, debated, and their merits promoted as a potential design enhancement to improve trial efficiency. However, the practical limits to implementation, gains in trial efficiencies, or improvements to inference have been slow to be reported to guide design, particularly in cases when the study is not a standard two-group parallel design. This session will explore real-world examples of sample size re-estimation procedures. Session speakers will discuss results from extended follow-up of trials utilizing these methods, extension of these methods to cluster-randomized trials and to non-inferiority trials, as well as provide insights on current controversies in utilizing these methods in real applications.

This paper session will include a chair and four speakers:

Leslie A McClure, Professor & Chair in the Department of Epidemiology and Biostatistics at Drexel University. Challenges in implementing sample size re-estimation in cluster-randomized trials

Cluster-randomized trials are increasingly popular in clinical and health services research. Cluster-randomized trials can be under-powered because sample size estimation assumes equal cluster sizes, but in reality, cluster sizes can be vastly different. Internal pilot designs, in which nuisance parameters can be re-
estimated during on-going data collection and then used to update the sample size mid-study, may help to ensure cluster-randomized trials have sufficient power to detect interesting treatment differences. In this talk, we describe the results from a simulation study in which we examined an internal pilot approach to re-estimating the sample size using the coefficient of variation in cluster size, and the re-estimated intraclass correlation observed in the accumulating data. We provide results from our simulations, and make recommendations as to when re-estimation of the ICC may provide benefit to the overall study operating characteristics.

Erinn M Hade, Associate Professor in the Department of Biomedical Informatics and Obstetrics & Gynecology at The Ohio State University. Follow-up after sample size re-estimation for disease free survival

While the clinical trials and statistical methodology literature on sample size re-estimation is robust, evaluation of sample size re-estimation procedures following the completion of a clinical trial has been sparsely reported. We review our experience of sample size re-estimation in a large international, National Institutes of Health-funded clinical trial for adjuvant breast cancer treatment, and evaluate our blinded sample size re-estimation procedure for this time to event trial. We evaluate the sample size re-estimation procedure by examining assumptions made during the re-estimation process, estimates resulting from re-estimation, and the impact on final trial results with and without the addition of participants following sample size re-estimation. We found that our re-estimation procedure did reasonably well in estimating the control group failure probabilities at the time of re-estimation. Particularly for time to event outcomes, pre-planned blinded sample size re-estimation procedures are a viable option to aid in maintaining study power.

Valerie Durkalski-Mauldin, Professor in the College of Public Health Sciences, Medical University of South Carolina. Sample size re-estimation: are there additional considerations in the setting of non-inferiority trials?

Non-inferiority (NI) trials present unique challenges in trial design and conduct. The Food and Drug Administration released a guidance on NI trials in 2016 that highlights key concepts and the importance of an appropriate active control, appropriate NI margin, and assurance of assay sensitivity. The guidance also highlights sample size estimation and the potential difficulty in obtaining reliable assumptions, with a recommendation of considering prospective sample size re-estimation to maintain adequate statistical power if the assumptions are incorrect. The literature on sample size re-estimation has primarily evolved around superiority designs with little attention given to potential differences in the non-inferiority setting. We will share a blinded sample size re-estimation approach taken in a current National Institutes of Health-funded clinical trial designed to test non-inferiority of indomethacin versus indomethacin plus stenting in the prevention of post-endoscopic retrograde cholangiopancreatography pancreatitis.

Christopher S Coffey, Professor in the Department of Biostatistics, University of Iowa. Is Sample Size Re-Estimation Controversial—Well, Yes & No

The issue of whether or not sample size re-estimation approaches are controversial is a rather complicated one that causes a great deal of confusion. This is somewhat complicated by the fact that sample size re-estimation encompasses a range of approaches that are both one of the most acceptable one of the more controversial forms of adaptive designs proposed to date. The main distinction involves whether or not one is re-estimating only nuisance parameters unrelated to treatment effect, treatment effect, or both. This talk will attempt to clarify the important distinctions between the various approaches, and use examples to demonstrate where the differences and controversies exist.

---

**SP-II**

**THE ASPIRE GUIDANCE ON SURGICAL PLACEBO CONTROLS IN RANDOMIZED TRIALS**

Organizer:
JONATHAN A COOK
*University of Oxford*

Contributor(s):
DAVID BEARD
*University of Oxford*
SIAN COUSINS
*University of Bristol*
MARION CAMPBELL
*University of Aberdeen*

**Background:** Placebo-controlled randomized trials are generally considered the optimal trial design to investigate healthcare interventions. Having a placebo control potentially addresses a number of biases which could affect the observed effect. Their use outside of pharmacology is limited, and this is particularly true for trials evaluating surgical procedures. Perhaps surprisingly surgical placebos have a history in randomized trials dating back to the 1950s. However, their use is highly controversial for ethical and design reasons and trials using them are relatively rare. Recently, there has been increased interest in their conduct and a number of high-profile trials (e.g. ORBITA, a surgical placebo trial of stenting for stable angina, and CSAW, a trial of sub-acromial decompression for shoulder pain). In this session, we will present findings from the new ASPIRE (Applying Surgical Placebo In Randomized clinical trials) guidance framework.
Evaluations) guidance on the use of placebo surgical in randomized trials and the related checklist developed by an international group of methodologists, trialists, ethicists and surgeons. In addition, related work which provides an up-to-date review of current practice and the limitations in it will also be presented.

Invited session outline:
Session Chair—Dean Fergusson—Introduction to the session

Talk 1 (David Beard): New insights on designing a placebo surgical trial. The latest thinking and understanding of the placebo phenomenon in surgery will be covered in the initial presentation along with an outline of the session. This presentation will address the key principles of the ASPIRE guidance on how to design a surgical trial with a placebo control. It will address issues such as fidelity to the real procedure, mitigation of risk, and related ethical and conduct issues.

Talk 2 (Sian Cousins): When & how have placebo trials been used? Results of a systematic review examining key methodological issues will be presented in this talk. It will also address how to design an invasive placebo control (the DITTO framework), including deconstruction of the treatment intervention and identification of the critical surgical element.

Talk 3 (Marion Campbell): The practical issues of interpretation and implementation of the new guidance. This talk will address interpretation and dissemination issues, as well as presenting the ASPIRE checklist for developing a surgical placebo trial.

The final part of the session will be used to invited comments and questions from the attendees to the presenters in an open moderated discussion.

This invited session will bring together a faculty with direct experience of the design, conduct, and ethical challenges posed by placebo trials in surgery. The session will involve invited talks providing an up-to-date review of practice and the findings from the ASPIRE guidance.

SP-12
INCREASING DIVERSITY IN CARDIOVASCULAR CLINICAL TRIALS: ENGAGEMENT, RECRUITMENT, AND LEADERSHIP

Organizer:
LEHANA THABANE
McMaster University

Contributor(s):
PATRICK GEE
iAdvocate

CLYDE YANCY
Northwestern University

HARRIETTE VAN SPALL
McMaster University

RENATO LOPES
Duke University

PAMELA S DOUGLAS
Duke University

DIEGO ARAIZA GARAYGORDOBIL
National Institute of Cardiology of México

REBECCA ORTEGA
Women as One

MARY NOREEN WALSH
Indiana University School of Medicine

Cardiovascular clinical trials inform practice and policy for conditions that have a high disease burden on populations across the world. While the results of these trials are applied broadly, trial participants often do not reflect the diverse patients in whom the results are applied and we rely on anecdotal and observational data to establish the safety and efficacy of treatments in these groups. Diversity is also lacking among clinical trialists, and this appears to be one of the several factors associated with the homogeneity of clinical trial participants.

In this 90-min session that includes patients, cardiovascular researchers, and national society leaders, we will present data on the under-representation of women, the elderly, and those from vast geographic regions in the world as clinical trial participants. We will discuss the impact of this under-representation on clinical care. Similarly, we will present data on the relative gender and geographic homogeneity among clinical trial leaders, with a focus on lead/senior authors of trial publications and trial executive committees. We will discuss factors associated with the under-representation of demographic groups as participants and leaders of trials and propose strategies at the clinical trial, institutional, medical society, and academic publishing level that could address these gaps.

This session will engage broad participation via the use of Social Media. Questions from Twitter will be fed into the discussion through the use of a Social Media Ambassador. #ResearchDiversity #SCT20.

Chairs: Drs Lehana Thabane and Mary Noreen Walsh.

Social Media Moderators: Dr Araiza Garaygordobil and Ms Ortega.

Patrick Gee: Engaging patients meaningfully in clinical trials: research participants and partners (10 min).

Clyde Yancy: Geographic and racial representation in clinical trials: how diverse are we and why does it matter? (12 min).

Harriette Van Spall: Closing the gender gap in cardiovascular randomized clinical trials: trial recruitment and leadership (12 min).
Renato Lopes: Can pragmatic trial designs increase diversity among trial participants? (12 min).

Pamela Douglas: Diversity in cardiovascular research: do medical societies and journals have a role? (12 min).

Discussion: 30 min.

SP-13

BAYESIAN MODELING IN COVID-19—DISEASE MODELING, THERAPEUTICS, AND VACCINES

Organizer:
KERT VIELE
Berry Consultants

Contributor(s):
ROGER LEWIS
UCLA Emergency Medicine, Berry Consultants

BENJAMIN SAVILLE
Berry Consultants

MARK FITZGERALD
Berry Consultants

SATRAJIT ROYCHOUDHURY
Pfizer

Bayesian Modeling in COVID-19—Disease Modeling, Therapeutics, and Vaccines

COVID-19 efforts have dominated the headlines in 2020. These efforts have involved efforts across the medical and statistical spectrum, from modeling of the pandemic to the development of therapeutics to the testing of possible vaccines. Novel methodologies have been utilized, such as platform trials, Bayesian modeling of pandemic uncertainty, and Bayesian adaptive trials to facilitate timely vaccine delivery. In this session, we will present four real examples of Bayesian methods across this range of activities. These include the official modeling of the epidemic within Los Angeles County by the leader of the team, both design and execution of platforms trials within the COVID-19 pandemic, and the Bayesian Pfizer vaccine trial.

All speakers confirmed.

Roger Lewis is the leader of the COVID-19 epidemic modeling team for Los Angeles County, California, advising government officials on the progress of the epidemic and projecting future developments. He will discuss the Bayesian SEIR modeling performed for Los Angeles, including capturing uncertainty in the predictions and real-world issues in data collection and adjusting modeling in the presence of evolving medical care and government policies.

Ben Saville will discuss therapeutic adaptive platform trials like PRINCIPLE and REMAP-CAP (focus on PRINCIPLE). Both trials are ongoing adaptive platform trials investigating multiple therapies for COVID-19. PRINCIPLE is a UK national priority trial and is focused on ambulatory participants with suspected COVID-19 and a higher risk of morbidity (e.g. >50 years age with comorbidities). The trial is open-label and has co-primary endpoints of subject-reported time to recovery and hospitalization. REMAP-CAP includes both open-label and blinded interventions focused on hospitalized patients in the intensive care unit across eight countries. The primary endpoint is the number of organ support-free days, and includes multiple interventions within multiple therapeutic domains, for example, antiviral agents, corticosteroids, or immunoglobulin. Both trials use innovative Bayesian modeling that account for potential drift over time, with frequent interim analyses allowing early decisions of futility or superiority. Response-adaptive randomization is used to increase allocation to interventions with better observed outcomes, which can increase statistical power for finding effective therapies and result in better participant outcomes.

Mark Fitzgerald will present some of the challenges of execution for the statistical analysis committee for a trial that is rapidly adapting to an ongoing pandemic, with a focus on REMAP-CAP. REMAP-CAP is an adaptive platform trial that explores the efficacy of interventions across a range of treatment domains, including the combinations across domains, that utilizes a novel endpoint: organ-support free days. The international effort combines data from five continents, evaluates thousands of treatment combinations, and rapidly evolves to accommodate information from external sources. The statistical analysis committee faces unique challenges in adjusting to rapid changes when combining data from disparate sources, updating models and reports to incorporate new design features, and producing results for public disclosure for closed domains or interventions, all while ensuring proper communication and maintaining trial integrity.

Satrajit Roychoudhury will discuss the design of the Pfizer Bayesian adaptive vaccine trial. This trial incorporates multiple interim analyses, each based on achieving a sufficiently high Bayesian posterior probability of vaccine efficacy. The trial also incorporates early stopping for futility based on Bayesian predictive probabilities. In November 2020, the trial is currently ongoing. Additional information may be publicly available at the time of SCT 2021 that may be discussed, but this will depend on future events at time of submission.
SP-14
THE IMPACT OF COVID-19 ON THE CONDUCT OF TRIALS IN THE NeuroNEXT NETWORK
Organizer: JANEL FEDLER
The University of Iowa
Contributor(s):
MICHELE COSTIGAN
The University of Iowa
DIXIE ECKLUND
The University of Iowa
TREVIS HUFF
The University of Iowa
DAVID KLEMENTS
Massachusetts General Hospital
BRENDA THORNELL
Massachusetts General Hospital
ELIZABETH KLINGNER
The University of Iowa
ANNA GUDJONSDOTTIR
The University of Iowa

The Network for Excellence in Neuroscience Clinical Trials (NeuroNEXT) is a large, federally funded neurology clinical trial network that was established in 2011. At the start of the COVID-19 pandemic, five clinical studies were active in the network. Rapid response to the evolving COVID-19 situation was critical as clinical study staff transitioned to remote work and on-site research activities were limited. During this session, we will discuss the impacts of the global pandemic on the active clinical trials in the NeuroNEXT network.

The ongoing clinical trials were in various stages of study start-up, participant follow-up, and study close-out. Each study required investigators to develop solutions that were simple, yet flexible, all while maintaining study integrity and patient safety. We will highlight the specific adjustments made to the trials and summarize strategies to be carried forward to future studies. The structure of our session will be five presentations covering the following topics:

1. Introduction of NeuroNEXT and the Active Clinical Trials (Brenda Thornell, BS; ~8 min):
   - Overview of NeuroNEXT Network Structure.
   - Overview of status of the NeuroNEXT Network (NN) trials impacted by COVID.
   - Identification of COVID impact.

2. Site Communication and Monitoring (David Klements, MS, CCRP & Michele Costigan, RN, BSN, CCRC; ~14 min):
   - Drafting guidance to study sites for implementation of protocol modifications.
   - Tracking local restrictions and guidelines by each site and geographic region.
   - Regular meetings/teleconferences with sites to discuss challenges and potential solutions with conducting study activities during the pandemic.
   - Remote monitoring.
   - Drug accountability-dispensing, compliance, and returns.

3. Protocol Changes and Data Collection (Trevis Huff, BSE; ~13 min):
   - Capturing National Institutes of Health/Food and Drug Administration requested COVID-19 impact data.
   - Allowing for and tracking telehealth/remote visits.
   - Accommodating Drug Dispensing/Shipping/Receiving challenges

4. Reporting on COVID Study Impacts (Elizabeth Klingner, MS and Anna Gudjonsdottir, MS; ~11 min):
   - COVID related protocol deviations and adverse events.
   - Recruitment and retention before and after the pandemic.
   - Data quality.
   - Analysis implications

5. Lessons Learned and Recommendations (Janel Fedler, PhD; ~4 min):
   - Adjustments that maintain trial integrity and patient safety.
   - Documentation and communication.
   - Design and conduct of future trials.
   - Following the presentations, there will be a 10-min Q&A.

SP-15
PRAGMATIC TRIALS: KEY UNRESOLVED ETHICAL ISSUES AND THE NEED FOR GUIDANCE
Organizer: MONICA TALJAARD
Ottawa Hospital Research Institute
Introduction: Pragmatic trials raise ethical issues that we believe have not been addressed adequately. The lack of guidance specific to pragmatic trials leads to divergent practices in the design and conduct of randomized clinical trials and delays in ethics approval. Our ethics of pragmatic trials research group is working to provide ethical guidance for pragmatic trials. In this panel session, we will review major empirical and ethical findings of our work and discuss key ethical issues to be addressed in future guidance.

Background: There is a widely recognized need for more pragmatic trials that evaluate interventions in real-world settings to inform decision-making by patients, providers, and health system leaders. Increasing availability of electronic health records, centralized research ethics review, and novel trial designs, combined with support and resources from governments worldwide for patient-centered research, have created an unprecedented opportunity to advance the conduct of pragmatic trials, which can ultimately improve patient health and health system outcomes. Such trials raise ethical issues that have not yet been fully addressed, with existing literature concentrating on regulations in specific jurisdictions rather than arguments grounded in ethical principles. Proposed solutions (e.g. using different regulations in “learning healthcare systems”) are speculative with no guarantee of improvement over existing oversight procedures. Most importantly, the literature does not reflect a broad vision of protecting the core liberty and welfare interests of research participants. Novel ethical guidance is required. We have assembled a team of ethicists, trialists, methodologists, social scientists, knowledge users, and community members with the goal of developing guidance for the ethical design and conduct of pragmatic trials.

Our project combines empirical and conceptual work and a consensus development process. Empirical work will: (1) identify a comprehensive list of ethical issues through interviews with a small group of key informants (e.g. trialists, ethicists, chairs of research ethics committees); (2) document current practices by reviewing a random sample of pragmatic trials and surveying authors; (3) elicit views of chairs of research ethics committees through surveys in Canada, the United Kingdom, the United States, France, and Australia; and (4) elicit views and experiences of community members and health system leaders through focus groups and surveys. Conceptual work consists of ethical analyses of identified issues and the development of new ethical solutions, outlining principles, policy options, and rationales. A consensus development process will involve an independent expert panel to develop a final guidance document.

Our 60-min panel session is divided into three parts. In the first part, we will review major empirical findings. In the second part, we will discuss conceptual projects and their results. In the third part, we will discuss the future of the ethics of pragmatic trials in a panel discussion.

Introduction Speaker: Dean Fergusson (Ottawa Hospital Research Institute, Canada). Bio: http://www.ohri.ca/profile/Dean_A_Fergusson. In this brief introduction (5 min), the session chair will introduce the speakers and provide a brief overview of the ethics of pragmatic trials research project.

Empirical Projects
Talk 1: Ethical issues in published pragmatic trials. Speaker: Monica Taljaard (Ottawa Hospital Research Institute, Canada). Bio: http://www.ohri.ca/profile/mtaljaard. In this talk (12 min), the speaker will describe the identification of published pragmatic trials in the literature and the results from a review of a subset of 1988 pragmatic trials focusing on prevalence and variation in reporting and obtaining informed consent.

Talk 2: Stakeholder views of key ethical issues. Speaker: Stuart Nicholls (Ottawa Hospital Research Institute, Canada). Bio: https://orcid.org/0000-0003-0485-9069. In this talk (12 min), the speaker will discuss key stakeholder perspectives on ethical issues in pragmatic trials identified in 45 interviews with trialists, regulators, and ethicists.

Conceptual Projects
Talk 3: Ethics of informed consent in pragmatic trials. Speaker: Cory Goldstein (Western University, Canada). Bio: https://www.charlesweijer.com/corygoldstein. In this talk (12 min), the speaker will address the tension between the requirement to obtain informed consent and the imperative to conduct socially valuable pragmatic trials. The speaker argues that the ethical principle underlying the regulatory requirement of informed consent can be achieved through the use of alternative models of consent, such as the integrated consent model.

Panel Discussion
Discussant: The ethics of pragmatic trials: where to from here? Speaker: Sandra Eldridge (Queen Mary University London, United Kingdom). Bio: https://www.qmul.ac.uk/blizard/all-staff/profiles/sandra-eldridge.html. The panel discussion will be kicked off...
by a discussant (5 min) who will set out key questions for the ethics of pragmatic trials.

Open discussion (14 min): Panelists will engage with the audience addressing questions about the empirical and conceptual projects presented, gaps identified, and key ethical issues in pragmatic trials to be addressed in future guidance.

**SP-16**

SCT NEW COMMITTEES AND DMC TRAINING INITIATIVE

**Organizer:**

SUSAN HALABI
Duke University

**Contributor(s):**

LEHANA THABANE
McMaster University

RICHARD CHAPPELL
University of Wisconsin–Madison

DAVID L DEMETS
University of Wisconsin–Madison

In this session, SCT members will be updated on creation of new committees and the latest Initiative. The SCT was founded as an interdisciplinary organization, and we have the responsibility to extend that approach to embrace demographic and geographical diversity, as well to be more inclusive of those who serve as patient advocates. Toward these ends, the board of directors approved the creation of the Equity, Diversity and Inclusion Committee, with the mission of creating outreach programs to engage and recruit members from diverse and under-represented populations as well as undergraduate and graduate students. In addition, SCT Board of Directors approved the formation of the Outreach Committee to create and support liaisons with other statistical and medical organizations and to promote international collaboration among both researchers and institutions.

Recognizing the increase in the number and complexity of Phase 3 randomized clinical trials has led to a growing need for more and more Data Monitoring Committees. This highlighted training the shortage of clinical trialists with expertise and experience in Data Monitoring Committees. Dr Dave DeMets launched the Data Monitoring Committee initiative to encourage the training of current and next generation of clinical trialists involved in medical research so they may serve effectively as members of independent Data Monitoring Committees sponsored by government, industry, and not-for-profit organizations.

**SP-17**

DESIGN AND IMPLEMENTATION OF CLINICAL TRIALS IN A PANDEMIC—REFLECTING ON THE ADAPTIVE COVID-19 TREATMENT TRIAL STUDIES

**Organizer:**

TYLER A BONNETT
Frederick National Laboratory

**Contributor(s):**

LORI E DODD
National Institute of Allergy and Infectious Diseases, National Institutes of Health

MAT MAKOWSKI
The Emes Company, LLC

TYLER BONNETT
Frederick National Laboratory for Cancer Research

PETER SASIENI
King’s College London

BIRGIT GRUND
University of Minnesota

This 1-h session will use examples from the Adaptive COVID-19 Treatment Trial to explore the unique challenges that the pandemic setting raises for clinical trials. The session will begin with three 10-min talks from statisticians on the Adaptive COVID-19 Treatment Trial study team. These talks, described in more detail below, are arranged to outline statistical considerations for trial design, trial implementation, and effective communication of trial results. These will be followed by 7-min talks from two discussants: one member of the Adaptive COVID-19 Treatment Trial Data and Safety Monitoring Board and one member of the National Institute of Allergy and Infectious Diseases COVID-19 Treatment Guidelines Panel. The session is designed to highlight lessons learned by the Adaptive COVID-19 Treatment Trial team and to explore applications of those lessons to other trials in similar settings. The session will conclude with 10 min for questions and comments.

Dr Lori Dodd, blinded statistician for the Adaptive COVID-19 Treatment Trial protocol, will be the speaker for talk one—“Planning for the Unknown: Designing trials in a pandemic setting” (12 min). This talk will provide a brief background of the Adaptive COVID-19 Treatment Trial study platform and will then highlight issues that arise at the start of the clinical trial process during an outbreak. The talk will address considerations for endpoint selection in settings where knowledge of the disease is rapidly evolving.
Dr Mat Makowski, unblinded statistician for the Adaptive COVID-19 Treatment Trial protocol, will deliver talk two—“Handling the Curve: Data cleaning and interim monitoring during rapid enrollment” (12 min). This talk will focus on strategies for handling rapid enrollment during trials in outbreak settings and will discuss how the extraordinarily high number of COVID-19 cases during the conduct of the Adaptive COVID-19 Treatment Trials challenged the usual approach to interim monitoring and what was done to address these challenges.

Tyler Bonnett, unblinded statistician for the Adaptive COVID-19 Treatment Trial protocol, will be the speaker for talk three—“Getting the Word Out: Communicating trial results when the world is watching” (12 min). This talk will use examples from multiple Adaptive COVID-19 Treatment Trial studies to highlight how the study team navigated release of trial results with a focus on the tension between the ethical obligation to quickly disseminate findings and the merits of rigorous data cleaning and peer review.

These three speakers will be followed by talks from two discussants.

Dr Peter Sasieni, Adaptive COVID-19 Treatment Trial Data and Safety Monitoring Board member, will then provide the first of two discussant talks—“Data Monitoring like Never Before: Reflections from participating in Data and Safety Monitoring Board deliberations during a pandemic.” This 7-min talk will highlight challenges faced by the Adaptive COVID-19 Treatment Trial Data and Safety Monitoring Board and give guidance on various questions surrounding the role of the Data and Safety Monitoring Board during a pandemic.

Dr Birgit Grund, National Institute of Allergy and Infectious Diseases Treatment Guidelines Panel member, will then conclude the speaking portion of the session with the final discussant talk—“Treatments: Weighing emerging evidence for benefit or harm.” This 7-min talk will address challenges in synthesizing information from multiple clinical trials and observational studies in a pandemic, and challenges in interpreting potentially different signals in subgroups of participants; for example, in ACTT-1, the estimated treatment effect in critically ill patients was different from the effect in less severely ill patients.

The remainder of the session (10 min) will be devoted to a question-and-answer session.

**Contributor(s):**

JOHN SCOTT  
FDA

MILO PUHAN  
University of Zurich

MARC BUYSE  
International Drug Development Institute

CHEN HU  
Johns Hopkins University School of Medicine

**Proposal:** Benefit–risk assessment plays a central role in the decision-making process to define the most appropriate medical practice and treatment regimen through clinical trials. Conventional approaches analyze primary and secondary endpoints separately, and the decision-making process largely rely on qualitatively integrating the segregated marginal treatment effect evaluations. In recent years, formal statistical methods, which synthesize information from multiple endpoints and perspectives from patients and clinicians to provide quantitative benefit–risk assessment, have been proposed and implemented in multiple indications. These methods provide novel frameworks to provide a comprehensive and coherent assessment of the global impacts and totality information of treatment under investigation. This session aims to provide an updated review on these novel methods, applications, challenges, and future research areas. Speakers:

John Scott, PhD. Center for Biologics Evaluation and Research, Food and Drug Administration. Scott, John John.Scott@fda.hhs.gov. Bio: John Scott is Director of the Division of Biostatistics in Center for Biologics Evaluation and Research, where he has also served as Deputy Director and as a statistical reviewer for blood products and for cellular, tissue and gene therapies. He has been heavily involved in a number of Food and Drug Administration’s statistical policy and outreach projects, including those for 21st Century Cures and PDUFA VI. Benefit–risk assessment is the foundation for Food and Drug Administration’s regulatory review of human drugs and biologics. These assessments capture the Agency’s evidence, uncertainties, and reasoning used to arrive at its final determination for specific regulatory decisions. In this talk, Dr Scott will review the current status and challenges of implementing quantitative benefit–risk assessment methods in Food and Drug Administration.

Milo Puhan, MD, PhD. University of Zurich, Switzerland. Milo Puhan miloalan.puhan@uzh.ch. Milo Puhan is the department chair of Epidemiology and Public Health and director of the Epidemiology, Biostatistics and Prevention Institute at University of Zurich. His main research interest is in prevention and management of chronic diseases, the quantitative assessment of benefits and harms of medical...
interventions and in the development of tools to support preference-sensitive healthcare. In this talk, the use of Quantitative Benefit Harm Assessment to estimate a benefit–harm balance to inform decisions at the level of specific populations (e.g., clinical recommendations or regulatory decisions), subgroups of these populations, and the individual patient level (personalized benefit–harm balance) is discussed. A framework is presented on how such a Quantitative Benefit Harm Assessment should be planned, executed, and disseminated that goes far beyond just statistical modeling and includes stakeholder engagement, systematic identification, and selection of evidence and development of guideline recommendations and decision aids.

Marc Buyse, ScD. International Drug Development Institute, marc.buyse@iddi.com. Bio: Marc Buyse is the founder and Chief Scientific Officer of International Drug Development Institute and of CluePoints Inc. He is Associate Professor of Biostatistics at Hasselt University in Belgium. He was President of the International Society for Clinical Biostatistics, President of the Quetelet Society, and Fellow of the Society for Clinical Trials. He worked at the EORTC (European Organization for Research and Treatment of Cancer) in Brussels and at the Dana Farber Cancer Institute in Boston. Dr Buyse will focus on the recent development of generalized pairwise comparisons for benefit–risk assessment. Generalized pairwise comparison uses all pairwise comparisons between two patients in different treatment arms in terms of one or several prioritized outcomes. Examples of generalized pairwise comparison include “net treatment benefit” (difference between the proportion of pairs in favor of treatment less the proportion in favor of control) and “win ratio” (proportion of pairs in favor of treatment over the proportion of pairs in favor of control). Pairwise comparisons can incorporate thresholds of clinical relevance, enable several outcomes to be analyzed simultaneously, and can be used flexibly for the benefit–risk assessment of therapeutic interventions.

Chen Hu, PhD, Johns Hopkins University School of Medicine. Bio: Dr Hu is an Associate Professor of Biostatistics in Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins. His statistical methodology interests include oncology clinical trial design and analysis, benefit–risk assessment and prognostic/predictive modeling. Dr Hu is the Senior Statistician of Lung Cancer Committee of NRG Oncology, one of the four cancer clinical trial cooperative groups funded by the National Cancer Institute. In this talk, the use of a novel summary metric for benefit–risk assessment will be discussed. This novel metric combines information from a longitudinal disease morbidity process and mortality (subject to censoring) into one single summary. Statistical inference tools, for example, hypothesis testing and regression methods will be discussed and applied in oncology trials.

SP-19
IT CAN BE DONE! EQUITABLE, ACCESSIBLE, AND EFFICIENT CLINICAL TRIALS THROUGH STREAMLINED PROTOCOL DESIGNS AND REQUIREMENTS

Organizer:
LAURA A LEVIT
American Society of Clinical Oncology

Contributor(s):
NATHAN PENNELL
Cleveland Clinic
LORA BLACK
Sanford Health
JANE PERLMUTTER
Gemini Group
PAUL KLUETZ
US Food and Drug Administration
MARY BEATTIE
Genentech
PETER FREDETTE
IQVIA

CONNIE SZCZEPANEK
Cancer Research Consortium of West Michigan NCORP

The COVID-19 pandemic disrupted all aspects of clinical care, including cancer clinical trials. The changes to clinical trial procedures made during the pandemic provide a long-term opportunity to improve and transform the clinical trial system. Improvements included expanded use of more pragmatic or streamlined clinical trial designs, fewer clinical trial–related patient visits, and minimized sponsor and contract research organization visits to trial programs. Changes like these are key to enhancing the efficiency of clinical trials and improving patient access to high-quality cancer care and opportunities to participate in research. Future policies and practices should be informed by emerging data and lessons learned from the experiences of patients, physicians, care teams, health systems, researchers, research sponsors, and contract research organizations in the wake of the pandemic.

Dr Nate Pennell, a physician investigator, will open the session with a presentation titled “Overview of ASCO’s Road to Recovery Research Recommendations,” which will highlight the key recommendations from the American Society of Clinical Oncology to create a more equitable, accessible, and
efficient clinical trials system that preserves patient safety and scientific integrity. The recommendations include strategies to ensure that trials are: more accessible, affordable, and equitable for patients; simplified, streamlined, and standardized to improve efficiencies and reduce research site burdens; integrated into clinical practice by a well-trained workforce; designed to be more pragmatic and informative; and conducted with appropriate oversight of clinical trial conduct and results. Widespread adoption and buy-in are key to ensuring that improvements to clinical trials are maintained as the immediacy of the COVID pandemic recedes.

The session will feature different key stakeholder perspectives regarding the impact of the pandemic on clinical trials and the sustainability of American Society of Clinical Oncology’s recommended changes to improve clinical trials.

Dr Jane Perlmutter, a patient advocate, will provide the patient perspective, sharing patient experiences, and emerging needs related to clinical trials in the wake of the pandemic. Her presentation is tentatively titled “Patient experiences and the need for more patient-centric and accessible clinical trials.”

Connie Szczepanek and Lora Black, a research nurse and research administrator, will provide the research team perspective, sharing experiences, and emerging needs related to clinical trials in the wake of the pandemic. Their presentation is tentatively titled “Clinical trial operations, management, and trial accessibility in a post-COVID-19 world.” (15 min presentation).

Dr Paul Kluetz, Deputy Director of the Food and Drug Administration Oncology Center of Excellence, will provide the regulatory perspective, sharing experiences, observations, and emerging changes, in his presentation tentatively titled: “The role of FDA in making clinical trials more efficient, accessible, and affordable.”

Dr Mary Beattie, from Genetech, will provide the biopharmaceutical trial sponsor perspective, sharing experiences, challenges, and emerging changes, in a presentation tentatively titled: “The role of biopharmaceutical companies in making clinical trials more efficient, accessible, and affordable.”

Peter Fredette, from IQVIA, will provide the contract research organization perspective, sharing experiences, challenges, and emerging changes, in a presentation tentatively titled: “The role of CROs in making clinical trials more efficient, accessible, and affordable.”

The session will conclude with a panel discussion.

This session will be particularly timely in 2021 given the expected state of the pandemic. Although it is likely that people will still be contracting the virus, many others will have been vaccinated. Thus, we will be on the road to recovery and at a critical point in ensuring that the positive changes to clinical trials are retained.

**SP-20**

**CHALLENGES AND BEST PRACTICES FOR RECRUITMENT AND RETENTION DURING THE COVID PANDEMIC**

Organizer:
DEE BLUMBERG
Emmes

Contributor(s):
CARMEN ROSA
National Institute on Drug Abuse, NIH

MICHELLE LOFWALL
University of Kentucky College of Medicine

SHARON LEVY
Harvard Medical School

KATHRYN HEFNER
Emmes

Recruitment and retention are key aspects of clinical research trials and critical factors in determining the success of the study. Despite this, even under regular conditions, successfully enrolling the desired number of individuals to achieve the targeted sample size and retaining those participants throughout the trial can be among the most challenging aspects in implementation of randomized controlled trials. This difficulty has been amplified recently, with the novel challenge of implementing studies during the COVID-19 pandemic presenting an even greater hurdle to recruitment goals.

The National Institute on Drug Abuse National Drug Abuse Treatment Clinical Trials Network had several studies facing the task of recruiting and retaining their participants during this difficult time.

In this 60-min panel presentation, a sponsor representative will briefly introduce the Clinical Trials Network and discuss the difficulty of continuing to execute a large research portfolio throughout the pandemic. Representatives from the Clinical Coordinating Center and Data and Statistics Center for these trials at Emmes will then introduce general best practices and techniques for tracking of recruitment and retention particularly relevant for such extreme situations, followed by two presentations from Clinical Trials Network investigators who will share the valuable experiences, innovative strategies, and lessons learned by launching their trials during the pandemic. Each presentation will last approximately 10 minutes and they are listed below in more detail and in chronological order. The remaining 10 or so minutes will be devoted to a question and answer session with audience members.

Carmen Rosa, MS (NIDA, NIH). Ms. Rosa is a Regulatory Affairs Specialist and Scientific Officer in
the Center for Clinical Trials Network which manages National Institute on Drug Abuse’s National Drug Abuse Treatment Clinical Trials Network. The Clinical Trials Network conducts studies of behavioral and pharmacological substance use treatment interventions in multi-site clinical trials. Ms Rosa will introduce the panel discussion topic and discuss the key importance of recruitment and retention during the COVID-19 pandemic in order to continue to successfully implement these Clinical Trials Network multi-site clinical trials.

Dikla Shmueli-Blumberg, PhD (Emmes). Dr Blumberg is a social psychologist and the Co-PI of the Clinical Coordinating Center for the Clinical Trials Network trials, where she oversees regulatory support, safety monitoring, quality assurance and protocol monitoring, and clinical equipment support and logistics. She will introduce strategies and creative solutions that have allowed Clinical Trials Network studies to continue to enroll and retain participants who may have not been able to attend research sites in-person during the COVID-19 pandemic, such as transition to e-consent and remote collection of biospecimen samples for urine drug testing. Michelle Lofwall, MD (University of Kentucky College of Medicine).

Dr Lofwall is a Professor in the Departments of Behavioral Science and Psychiatry in the Center on Drug and Alcohol Research at the University of Kentucky. Her research is focused on the treatment of substance use disorders and she serves as the Co-Investigator. Dr Lofwall will introduce the Clinical Trials Network-0080 study, a pragmatic multi-site, randomized controlled trial with the primary objective of evaluating the impact of treating opioid use disorder in pregnant women with injectable subcutaneous (SC) extended release buprenorphine (BUP-XR), relative to daily sublingual (BUP-SL), on mother–infant outcomes, and provides an account of the recruitment and retention experiences on the MOMs study.

Sharon Levy, PhD (Harvard Medical School). Dr. Levy is an Associate Professor in Pediatrics at the Harvard Medical School, and Director of the Adolescent Substance Use and Addiction Program at the Boston Children’s Hospital. Her research includes outpatient management of substance use disorders in adolescents, including screening and brief advice in primary care. Dr Levy will introduce the Clinical Trials Network-0060A1 study (ASC-THEM), an interview-based survey study which evaluates the psychometric properties of three substance use screening and brief assessment tools for adolescents, and describe the trial and error and strategies her team used to increase recruitment during the pandemic.

Kathryn Hefner, PhD (Emmes). Dr Hefner is a clinical psychologist and the Scientific Project Leader for the Emmes Data and Statistics Center, where she oversees development of the data system and the scientific content of the assessments used in National Institute on Drug Abuse Clinical Trials Network trials. She will discuss data system considerations and adaptations in the context of COVID-19 and their relation to study conduct during the pandemic, as well as the use of trial progress reports to track recruitment and retention.

**SP-21**

**CURRENT ISSUES IN THE DESIGN AND ANALYSIS OF COVID-19 VACCINE EFFICACY TRIALS**

Organizer: HOLLY JANES
Fred Hutchinson Cancer Research Center

Contributor(s):
DAVID BENKESER
Emory University

HONGHONG ZHOU
Moderna Therapeutics, Inc.

MARTHA NASON
National Institutes of Health

ELIZABETH BROWN
Fred Hutchinson Cancer Research Center

The global COVID-19 pandemic has posed unprecedented challenges to public health, not to mention the global economy and way of life. The pace of vaccine development has been unprecedented and major successes have been realized, with three vaccines having received authorization for use in several countries around the globe following efficacy trials demonstrating high efficacy at preventing COVID-19 disease. However, considerable challenges remain. The immunological predictors of and mechanisms for the efficacy have yet to be defined, and identifying such immune correlates will be critical for bridging the efficacy to populations not included in the original efficacy trials. Additional vaccines also require evaluation to meet global demand. And evaluating the effects of vaccines on onward transmission of infection is imperative to guide policy and optimize vaccine uptake. The talks in this session will pertain to these topics of discovering immunological correlates, designing non-inferiority studies of new candidate COVID-19 vaccines, bridging efficacy of efficacious vaccines to new populations, and designing studies of vaccine efficacy against transmission.

This session will include presentations by both academic and industry partners in the field of COVID-19 vaccine development. As leaders in the field, the speakers will articulate the scientific and statistical challenges faced for a given problem, and articulate and illustrate the practical solutions based on ongoing and planned studies. The session chair will work with speakers to
ensure that the talks are appropriately sequenced and complementary. Each talk will be 17 min long, followed by 3 min for question and answer, with 10 min discussion at the end of the session. Given the rapid pace with which the COVID-19 field is moving, the topics will be re-evaluated in coming months to ensure that they are relevant and timely.

The development and deployment of effective COVID-19 vaccines is currently among the highest public health priorities. The general topics for the talks (immunological correlates of vaccine protection; non-inferiority trial design; bridging studies; and transmission studies), are applicable to multiple pathogens, and the statistical problems encountered have features common to clinical trials in other fields.

**SP-22**

**CHOICES, CHOICES: SELECTING, DEFINING, AND MEASURING OUTCOMES FOR YOUNG PEOPLE WITH DEPRESSIVE DISORDERS**

**Organizer:**
DARREN COURTNEY
Centre for Addiction and Mental Health

**Contributor(s):**

KAROLIN R KRAUSE
Centre for Addiction and Mental Health

NANCY J BUTCHER
The Hospital for Sick Children Research Institute

SUNEETA MONGA
The Hospital for Sick Children

**Talk 1:** Presenter: Dr Darren Courtney Title: “Response,” “Remission,” and “Recovery” in randomized clinical trials for the Treatment of Adolescent Depression.

“Response,” “Remission,” and “Recovery” are often used as dichotomized outcomes to define clinically important change in randomized clinical trials for mental disorders. These dichotomizations can guide clinical decisions to continue, intensify, switch, or stop treatment. Through a rigorous scoping review of randomized clinical trials for the treatment of adolescent major depressive disorder (N = 98), we found high variability in how these terms were operationally defined; with 53 unique outcome definitions of “response” across 45 trials that assessed response, 47 unique definitions of “remission” in 29 trials that assessed remission, and 19 unique definitions of “recovery” across 11 trials that assessed recovery. Moreover, there was a dearth of rationale based on empirical findings. Finally, reports of input from youth with lived experience in establishing these definitions were minimal. In studying and implementing interventions to address mental disorders arising from the COVID-19 pandemic, a new and standardized approach to define clinically important change is needed.

**Talk 2:** Presenter: Dr Nancy Butcher Title: Measurement Matters: Evaluating Methods of Assessing Depression in Adolescent Clinical Trials.

Rates of major depressive disorder in youth are on the rise, creating urgency for well-designed trials using well-measured outcomes during the COVID-19 pandemic era. Timely conduct of psychiatric trials in youth, however, may require innovative and virtual methods for outcome measurement as an alternative design to in-person assessments. The Children’s Depression Rating Scale-Revised is the most commonly used method of measuring depressive symptoms in clinical trials of adolescents with major depressive disorder. Originally developed for use in children, it is unknown whether the Children’s Depression Rating Scale-Revised is fit-for-purpose for measuring depression in adolescents. This study aimed to identify all existing evidence of key measurement properties of the Children’s Depression Rating Scale-Revised in adolescents with major depressive disorder through a systematic review, and to evaluate these properties using a well-established appraisal method developed by the COSMIN (COnsensus-based Standards for the selection of health Measurement I(nstru)ments) Initiative. We found that it is unclear whether the Children’s Depression Rating Scale-Revised is fit-for-purpose for measuring depression symptom severity of adolescents with major depressive disorder, either in person or via virtual assessment. No study assessed content validity, cross-cultural validity/measurement invariance, or measurement error of the Children’s Depression Rating Scale-Revised in adolescents with major depressive disorder. Low-quality evidence was found for sufficient construct validity (n = 4 studies) and responsiveness (n = 2 studies). Very low-quality evidence was found for sufficient inter-rater reliability (n = 2 studies). The results for structural validity (n = 3 studies) and internal consistency (n = 5 studies) were inconclusive. No study evaluated the validity of the Children’s Depression Rating Scale-Revised using a virtual mode of administration. Implications and future directions for depression outcome research will be discussed including tools to guide outcome measurement instrument selection and reporting using innovative and virtual methods.

**Talk 3:** Presenter: Dr. Karolin Krause Title: Are Symptoms All That Matters? Contrasting Outcome Measurement in Youth Depression Trials with Youth Perceptions and Priorities.

Depression is a common mental health problem in adolescence worldwide. The COVID-19 pandemic is putting additional strain on youth wellbeing and may lead to a rising demand for effective treatment in the coming years. Within a person-centered healthcare
framework, treatment effectiveness should be judged with reference to individual patient needs. A recent systematic review suggests that most clinical trials for adolescent depression assess outcomes in the domain of symptom change, and half assess functional impairment, while other outcomes are rarely considered. We conducted a mixed-methods investigation of youth outcome perceptions and priorities, to understand the extent to which this focus reflects what matters to youth. In semi-structured interviews conducted with 34 youth following participation in a psychotherapy trial for depression, symptomatic change was the most frequently discussed outcome (by 65%) but change in coping skills was discussed just as often, followed by improved family functioning (47%) and social functioning (35%). These outcomes were assessed by less than 5% of recently published quantitative treatment studies. Subsequently, Q-methodology was used in a purposive sample of 28 youth with lived experience of depression, to investigate which outcomes they considered most important. Four distinct profiles were identified: “symptom reduction and enhanced well-being”; “improved coping and self-management”; “better understanding past and present”; and “less interference with daily life.” Findings suggest that a narrow focus on symptom metrics fails to cover secondary outcomes that matter to youth. Trialist should draw on emerging guidance provided by Core Outcome Sets to consider a broader range of outcomes. Trialists should further consider including a personalized outcome measure to reflect varying individual priorities and understand whether treatments enable progress toward youth’s individual treatment goals.

Talk 4: Presenter: Dr Suneeta Monga Title: Engaging Youth in Defining What Outcomes Matter: Development of a Core Outcome Set for Adolescent Depression with Active Youth Participation

Reviews of the literature document the wide heterogeneity in what, how, and when outcomes are measured in adolescent major depressive disorder. Such heterogeneity significantly limits the ability to draw conclusions around the best or most effective treatment as individual trial results cannot be pooled together in systematic reviews when outcomes differ across trials. A Core Outcome Set, as defined by the Core Outcome Measures in Effectiveness Trials Initiative, is an agreed, minimum, standardized set of outcomes that should be measured and reported in all clinical trials and audits in a specific area of health or healthcare. In other areas of medicine, development of a Core Outcome Set has increased consistency across trials, maximized the potential for a trial to contribute to systematic reviews of key outcomes, increased measurement of appropriate outcomes, and reduced selective outcome reporting. The development of a Core Outcome Set for adolescent major depressive disorder with a strong emphasis on youth and caregiver involvement to ensure that outcomes that really matter to youth and caregivers are incorporated into the final Core Outcome Set is currently underway. The use of innovative approaches that support engagement and participation of youth with lived experience, and caregivers, in all steps of the development of this Core Outcome Set will be a focus of this presentation. Important adaptations to manage the current impact of COVID-19 on the Core Outcome Set development will also be discussed as part of this presentation.

SP-23

CHALLENGES WITH COVID-19 VACCINE DEVELOPMENT AND ASSESSING IMPACT OF THE PANDEMIC ON ONCOLOGY TRIAL OUTCOMES

Organizer:
PRALAY MUKHOPADHYAY
Otsuka America Pharmaceuticals Inc.

Contributor(s):
SATRAJIT ROYCHOUDHURY
Pfizer

WENMEI HUANG
Moderna

JIABU YE
Astra Zeneca

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) causes a respiratory disease called coronavirus disease (COVID)-19 in infected individuals. Since its first outbreak at the end of 2019, COVID-19 has rapidly evolved into a global pandemic and has infected more than 85 million (M) people worldwide and has resulted in over 1.85 M deaths. The pandemic has led to the race for finding therapeutics and vaccines, and to date, there have been two vaccines approved within the United States and the potential for more to follow in the coming months. In addition, this has also forced clinical researchers to look into the impact this may have on ongoing trials, given the risk of COVID-related complications confounding with the efficacy and safety outcomes and necessitating the collection of additional data and evaluating alternate ways of analyzing trial results. This is of particular concern in patients with cancer who are often immunosuppressed as a result of both their disease and the treatment they receive, which puts them at increased risk of severe complications from COVID-19.

In this session, we will focus on two important but very distinct topics that have emerged due to the onset of the pandemic: (1) the issues and challenges related to vaccine development including trial design, rapid execution, real-time regulatory feedback, and overall
interpretation of trial results and (2) understanding potential impact of COVID-related deaths in Oncology trials, on time to event endpoints of overall survival and progression-free survival.  

**Topic 1: Challenges with COVID-19 Vaccine development**

The SARS-CoV-2, first identified in December 2019, has caused a worldwide pandemic leading to widespread morbidity and mortality. There was no Food and Drug Administration-approved vaccine for the prevention of the coronavirus disease 2019 (COVID-19), the disease caused by SARS-CoV-2. The urgent need for safe and effective interventions to mitigate the global spread of SARS-CoV-2 has prompted international efforts to develop antivirals and vaccines. Numerous vaccine candidates based on traditional and new platforms are currently being evaluated including nucleic acid (DNA and RNA), viral vector (replicating and non-replicating), virus-like particles, peptide-based, recombinant protein, and live attenuated and inactivated virus modalities.

COVID-19 vaccines are in various stages of clinical development, with several candidates in pivotal phase 3 clinical trials, including mRNA-based vaccines, of which two vaccines, one developed by Pfizer and the other by Moderna recently received the emergency use authorization from the Food and Drug Administration in December 2020. In this session, we will share experiences and challenges in the development of these two vaccines.

We will share the practical and statistical challenges and considerations when designing COVID-19 vaccine studies, such as target population(s), endpoint selection and assessment, statistical analysis method, timing of interim analyses, and questions continuously to be answered after receiving the emergency use authorization and/or during the Biologics License Application review. We show that, when planning a vaccine pivotal study against a novel virus causing ongoing worldwide pandemic, special consideration needs to be given for the designing of interim analyses related to efficacy, so that a vaccine with favorable benefit–risk profile can be made available as early as possible; special attention also needs to be given for an independent data safety monitoring board for vaccine-associated enhanced respiratory disease and other safety signal monitoring to mitigate the risk for trial participants during an ongoing pandemic.

**Topic 2: Impact of COVID-19 on Oncology trial outcomes using overall survival or progression-free survival**

Simulations were conducted to assess (1) the impact of the COVID-19-related death and (2) missed RECIST visits on the statistical analysis of time-to-event outcomes in randomized phase 3 oncology trials and explore mitigation options for this risk when COVID-19 cohort is well-defined. Two simulated case studies of Phase 3 randomized controlled trials ongoing during the pandemic outbreak were used to evaluate five approaches (1, ITT approach; 2, modified ITT excluding COVID-19-related deaths; 3, censoring COVID-19-related deaths with target number of non-COVID-19-related deaths; 4, censoring COVID-19-related deaths with original date-cut-off; 5, Fine & Gray modeling of competing risk approach to treat COVID-19-related death and non-COVID-19-related death separately), at presence of COVID-19-related deaths with different pandemic onsets relative to timing of analysis and varying pandemic durations, the impact of COVID-19-related death due to pandemic, on the primary endpoints of overall survival and/or progression-free survival in terms of type 1 error, power, and hazard ratio estimates.

It was found that COVID-19-related deaths would impact time-to-event analysis in terms of type 1 error and power for log rank test, and provide biased treatment effect estimation from Cox model if ITT approach is used; impact would be more severe if there was an imbalance in COVID-19-related deaths mainly in experimental arm. With same number of COVID-19-related deaths, the earlier the timing of the pandemic window, will lead to greater loss in power. Approaches censoring COVID-19-related deaths would minimize impact on power loss and bias in hazard ratio estimation, particularly if data cut-off was extended to mitigate for events loss due to censoring.

The simulations conducted in this analysis provide a framework to help understand how to mitigate the risk to the randomized oncology trials in which COVID-19-related deaths are observed in the blinded assessment during the COVID-19 pandemic.
are particularly susceptible to the disruption caused by the COVID-19 pandemic, as well as the impact of COVID-19 illness on older adults. In this session, we will present the impact of the pandemic on design and analytical issues related to the choice of study population, trial design, clinical endpoints, and methods of outcome ascertainment. We will also discuss operational complexities and approaches to address or evaluate the impact of this global pandemic.

This session will consist of three invited talks and a discussant. Each talk, representing clinical trials with different study designs, will be given by a member of the trial’s statistical leadership team. 

The studies to be discussed are as follows:

- **Pharmacological Trials**—The Anti-Amyloid Treatment in Asymptomatic Alzheimer’s (A4) study: This is an ongoing double-blind, placebo-controlled 240-week Phase 3 trial of an anti-Amyloid monoclonal antibody in older individuals (ages = 65–85 years) who have normal cognition and memory function but who may be at risk for memory loss due to Alzheimer’s disease. This study completed enrollment in December 2017 and is anticipated to be completed in 2022. The primary outcome in the study is the Preclinical Alzheimer Cognitive Composite, with the Cognitive Function Index as a key secondary outcome. It is conducted as a public–private partnership between the National Institute on Aging and Eli Lilly.

- **Non-Pharmacological Trials**—The US Study to Protect Brain Health Through Lifestyle Intervention to Reduce Risk (US POINTER): This is an ongoing 2-year clinical trial, that began recruiting in May 2019, to evaluate whether lifestyle interventions that simultaneously target many risk factors protect cognitive function in older adults (ages 60–79 years) who are at increased risk for cognitive decline. The primary outcome of this study is a global cognitive composite score. It is funded by the Alzheimer’s Association.

- **Pragmatic Trials**—The Pragmatic Evaluation of Events and Benefits of Lipid-Lowering in Older Adults (PREVENTABLE): This is a pragmatic trial that began recruitment in September 2020, studying the effectiveness of statins in adults 75 years or older without known cardiovascular disease on a primary outcome of survival free of dementia and disability. It is funded by the National Institute on Aging and the National Heart, Lung, and Blood Institute.

While the talks in the session will describe the experiences and decisions of the individual study team, the discussant will compare and contrast these experiences across the types of designs and stage of study.

 Speakers: Michael Donohue, PhD, Mark Espeland, PhD, and Nicholas Pajewski, PhD. Discussant: Rema Raman, PhD, University of Southern California.

**SP-25**

**DESIGN AND CONDUCT OF PILOT AND FEASIBILITY TRIALS: FOCUS ON UNRESOLVED ISSUES**

**Organizer:** LEHANA THABANE

*McMaster University*

**Contributor(s):**

- MOHAMMED KHAN
  *University of Toronto*

- KATIE MELLOR
  *University of Oxford*

- SASKIA EDDY
  *Queen Mary University of London*

- SANDRA ELDRIDGE
  *Queen Mary University of London*

- GILLIAN LANCASTER
  *University of Keele*

External pilot and feasibility trials are an essential part of trial preparation, particularly for the planning of complex interventions. However, they rarely published, there is confusion about the definitions of the terms “pilot trial” and “feasibility trial and there are also misunderstandings about their purpose.” Furthermore, the ones that are published often show the objectives are often not about feasibility, but focus on clinical outcomes, and there is lack of clarity on methodological focus. In this session, graduate trainees supervised by the pilot and feasibility study group will discuss some of the unresolved uncertainties in design, conduct, analysis and reporting of external pilot and feasibility trials: (1) consent process in pilot and feasibility study: here, we aimed to assess the transparency of informed consent in pilot and feasibility study by investigating whether researchers communicate, through consent documents, key features of the studies; (2) reporting of progression criteria for pilot and feasibility study: this part will discuss the characteristics of progression criteria reported in external randomized pilot trial and protocol publications, including whether progression criteria are specified a priori and feature in the prepublication peer review reports; (3) does the size of a pilot study affect what happens next?: this part will investigate whether the sample size of pilot studies is connected to their progression on the research pathway through linking published pilot trials to their future pieces of research. Here, we address issues such as: What are the outcomes
following a pilot trial? What was the sample size of the pilot and the planned and achieved sample size of the future work (if applicable)? Was the sample size of the pilot trial used to inform the sample size of the definitive trial/further pilot work? Is there an association between the size of the pilot and whether the pilot progressed?

Suggested organizers, chair, speakers, and discussant:
10 min from the organizer: Lehana Thabane, on update on pilot and feasibility study and the unresolved issues.

20 min: Mohammed Khan (DDS student), Biostatistics Unit, St. Joseph’s Healthcare, Hamilton, Ontario and University of Toronto, Toronto, Canada: Transparency of informed consent in pilot and feasibility studies is inadequate: a single center quality assurance study.

20 min: Katie Mellor (PhD student), Centre for Statistics in Medicine, University of Oxford, Oxford, England: Progression from external pilot to definitive randomized controlled trials are not adequately reported: a methodological review of progression criteria reporting.

20 min: Saskia Eddy (PhD student), Institute of Population Health Sciences, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, England: Does the size of a pilot study affect what happens next?

20 min for Summary and open discussion, co-chaired by: Sandra Eldridge and Gillian Lancaster.

SP-26
MISSING DATA IN CLINICAL TRIALS: PRINCIPLES, PLANNING, AND PREVENTION

Organizer:
THOMAS COOK
University of Wisconsin-Madison

Contributor(s):
KATHLEEN WANNEMUEHLER
University of Wisconsin–Madison

JANET WITTES
Statistics Collaborative Inc.

Randomized clinical trials are considered the gold standard for evaluating the effectiveness of new medical or clinical interventions. In the absence of missing data, randomization of treatment assignment is sufficient to ensure valid causal conclusions regarding the effect of treatment on the outcome of interest.

In 2010, the National Research Council report on “The Prevention and Treatment of Missing Data in Clinical Trials” emphasized the importance of trial designs and conduct that reduce the occurrence of missing data. Factors of trial design and conduct that can affect missing data rates include duration of trial (longer the trial, more risk of missing data), difficult to assess the outcome, type of intervention (surgical or medical), less adherence to study protocol (e.g. in psychiatric disorders), poor communication with study participants (to explain about the study, procedures to be followed, follow-up schedule, poor response to patient queries), and poor interpersonal relation with the study participants. In addition, investigator induced missing data by equating withdrawal from treatment to withdrawal from trial.

The report also recommended further research on the effect of missing data on power, and the need for sensitivity analyses to assess the effect of the assumptions on the missing data mechanism. A decade after the report’s publication, missing data continue to confound both clinical trialists and inference regarding the causal effect of the treatment on outcomes. Since the release of the report there have been many discussions and publications on the topic of missing data in randomized clinical trials, yet it is still common to see trials with substantial missing data, and complete-case analyses that implicitly or explicitly assume data are missing at random without justification. It is critical for trialists, the scientific community at large, and the broader public that may benefit (or be harmed) by new treatments to have a better understanding of the effects of missing data on power to detect treatment effects, and the threat to valid inference caused by data that are missing not at random.

This session will begin with an overview of the findings of the NRC report. This will be followed by a non-technical introduction to the research of Cook and Zea (Stats in Med, 2019) which shows how results from complete-case analyses can be misleading, and that the threat to validity increases not only as the rate of missingness increases, but also as the trial sample size increases. A principled sensitivity analysis approach will be presented that can assess the robustness of conclusions under pre-defined assumptions. From this foundation, a method for adjusting sample size for expected rates of missingness will be discussed. These results show the necessary inflation to sample size may be infeasible in some instances, which may require a complete rethinking of the trial. This brings us full circle back to the need to try every effort to reduce rates of missingness toward zero.

Presentations: 1. Revisiting the 2010 National Research Council report on Missing Data. Kathleen Wannemuehler, PhD, University of Wisconsin–Madison (~10 min). 2. Why missingness should cause more angst. Tom Cook, PhD, CHS, University of Wisconsin–Madison (~30 min). 3. Why does missingness not cause more angst? Janet Wittes, PhD, Statistics Collaborative, Inc. (~10 min).
SP-27
OPERATIONAL INNOVATIONS IN ACADEMIC DATA ANALYSIS AND COORDINATING CENTERS

Organizer:
LAUREN C BALMERT
Northwestern University

Contributor(s):
LOGAN SIRLINE
Medical University of South Carolina

KALEAB Z ABEBE
University of Pittsburgh

DEBBIE FELTON
Wake Forest School of Medicine

JOHN NICHOLS
Wake Forest School of Medicine

This session will focus on interdisciplinary, operational innovations within data analysis and coordinating centers, coalescing expertise and experience from diverse perspectives. Connecting project managers, data managers, programmers, and biostatisticians from academic institutions across the country, the session will highlight operational innovations in data quality assurance, regulatory document management, and medical coding and adverse event reporting. Speakers with various coordinating center roles will contribute their expertise, emphasizing the importance of these innovations in preserving trial integrity amid emerging challenges.

The outline for the session is as follows:

Talk 1 (15 min): Ensuring high-quality data through collaboration. Speaker: Logan Sirline. Summary: Ensuring high-quality data is a complex and critical component of clinical research and requires interdisciplinary collaborations between Principal Investigators, Data and Project Managers, Statisticians, Monitors, and Programmers. This discussion will provide strategies for enhancing data quality through building and maintaining these relationships at an academic data management center. These stakeholders provide the backbone for quality data and provide input as new ideas and errors arise. Conversations regarding the collection of high-quality, reliable data should start at trial design and continue through study close out with resources focused on critical risks including outcomes and human subjects’ protection. When designing biosample tracking modules and imaging tracking, it is imperative to protect the safety of the trial and the participants. By creating interfaces and databases that are sustainable through challenges and an ever-changing environment, having high-quality data is not a goal to achieve, rather it is inherent to the process.

Talk 2 (15 min): An electronic Master Regulatory File (eMRF). Speaker: Kaleab Z Abebe. Summary: Appropriate management of regulatory documents in clinical trials is crucial for any entities whose responsibilities involve the organization and/or oversight of study-, site-, and personnel-specific regulatory documentation. However, this is a very time-consuming task which takes time, effort, and focus away from the most important aspect of the lead coordinator and site coordinator’s role—monitoring the safety of the study subjects enrolled in the trial. In this talk, we describe the motivation for and the development and implementation of a regulatory management system, the electronic Master Regulatory File, which is in keeping with ICH E6 GCP guidelines. The electronic Master Regulatory File provides: (1) a secure, web-based system to facilitate document submission and retrieval, (2) automated email alerts indicating pending expiration of essential documents, (3) a platform to facilitate electronic signatures on key documents, and (4) the ability to track and disseminate institutional review board–related documents within sites. Each of the participating sites will be able to upload to the electronic Master Regulatory File and view site- and participant-level documents specific to their site. They will also be able to view shared documents such as the study protocol, MOP, and case report forms. In addition, we developed the electronic Master Regulatory File to serve as the starting point for all randomized clinical trials that the Center for Clinical Trials and Data Coordination coordinates. This involves the “registration” of a trial by the Center for Clinical Trials and Data Coordination, the designation of a “registrar” at the clinical coordinating center (or lead clinical site), and subsequent designation of registrars at the remaining clinical sites. This allows registrars to upload personnel-level information, documentation, and training prior to study access and ultimate study activation. Finally, we highlight how the electronic Master Regulatory File facilitates remote close out of clinical trials, which can save time and money.

Talk 3 (15 min): Medical coding and adverse event reporting. Speakers: Debbie Felton and John Nichols. Summary: Safety monitoring in clinical trials is required to protect the safety of research participants and to identify, evaluate, minimize, and appropriately manage risk. The US Study to Protect Brain Health through Lifestyle Intervention to Reduce Risk (US POINTER) Data Coordinating Center has developed a comprehensive Safety Tracking System for the reporting and assessment of adverse events reported during the trial and during US POINTER ancillary studies. Development of the system required balance of the functional implementation by the programmer while incorporating feedback from interdisciplinary researchers with expertise in participant...
safety. This talk will address the team’s challenges and considerations in development of the US POINTER Safety Tracking System. It will demonstrate the capabilities of a multi-faceted Safety Tracking System that must incorporate a variety of roles, access, and tasks into a seamless interface, balancing ease of use with restricting access to specific functions of the system based on the user’s role. It integrates the reporting of events through three distinct methods, allows assessment of these events at multiple levels including local site study clinicians and centrally by the US POINTER Safety Officer, as well as central MedDRA coding of all events. The result is a Safety Tracking System that allows for reporting and assessment of adverse events in a timely manner while ensuring that all participants are able to participate safely in the trial and its ancillary studies.

Discussion and Q&A with all panel members (12 min) led by Denise Scholtens, Jody Ciolino, and Lauren Balmert.

CONTRIBUTED PRESENTATIONS

CP-1

SAMPLE SIZE CALCULATION FOR STEPPED-WEDGE CLUSTER-RANDOMIZED TRIALS WITH MULTIPLE LEVELS OF CLUSTERING

Primary Author: KENDRA L DAVIS-PLOURDE
Yale University

Co-Author(s): MONICA TALJAARD
University of Ottawa

FAN LI
Yale University

The stepped-wedge cluster-randomized trial is an attractive design for evaluating health services delivery or policy interventions. In this design, clusters start in the control condition and gradually cross over to the treatment based on a schedule dictated by random assignment. Outcomes may be assessed on the same individuals over time (i.e. a cohort design) or different individuals (i.e. a cross-sectional design). A key consideration in this design is that sample size calculation and analysis must account for within-period as well as between-period intracluster correlations; cohort designs have additional correlations due to repeated measures on the same individuals. While numerous methods have been developed to account for within- and between-period intracluster correlations with a single level of clustering during each time period, a few methods are available to accommodate multiple levels of clustering. Our objectives were to develop computationally efficient sample size procedures that recognize within-period and between-period intracluster correlations in stepped-wedge trials with more than two levels of clustering. Focusing on three levels of clustering and assuming equal cluster-period sizes, we consider three variants, depending on whether each level is treated as a cross-sectional or closed-cohort design. We introduce an extended block exchangeable matrix to characterize the correlation structures both within- and between-clusters in each cluster-period and develop convenient sample size expressions that depend on this correlation structure. With a continuous outcome, we show the sample size expression depends on the correlation structure only through two eigenvalues of the extended block exchangeable matrix. For binary outcomes under a mixed effects framework, we develop a sample size expression based on a first-order Taylor approximation. We conduct simulation studies to examine the finite-sample properties of the proposed sample size algorithms and demonstrate the application of the proposed methods using the Washington State Expedited Partner Therapy trial: a multi-level stepped-wedge trial that randomized local health jurisdictions (level 4) consisting of clinics (level 3) and observed patients (level 2) with respect to their Chlamydia infection status (level 1).

CP-2

STATISTICAL CONSIDERATIONS FOR SEQUENTIAL, MULTIPLE ASSIGNMENT, RANDOMIZED TRIALS

Primary Author: SASHA KRAVETS
University of Illinois at Chicago

Co-Author(s): AMY S RUPPERT
The Ohio State University

JENNIFER G LE-RADEMACHER
Mayo Clinic

SUMITHRA J MANDREKAR
Mayo Clinic

Background/aim: Sequential, multiple assignment, randomized trial (SMART) designs are appropriate for comparing adaptive treatment interventions, in which intermediate outcomes guide subsequent treatment decisions for individual patients. These types of adaptive trials are suitable for designs which apply a sequence of interventions, specifically when patients vary in their responses to treatment. Within a SMART design, patients are re-randomized following an intermediate assessment. By continually randomizing patients, an SMART design reduces bias and balances patient characteristics across embedded treatment regimens. The goal of an SMART design could be to compare the effectiveness of treatments at each stage of the
trial, or across the stages, that is, between the embedded treatment regimens. We focus our discussion on two designs, both with a binary intermediate outcome, and one with a binary final outcome, and another with a time-to-event final outcome. We discuss how design parameters, including choice of randomization ratios for each stage of randomization (1:1, 1:2 and 1:3 for the first stage, and 1:1 and 1:2 for the second stage), and prevalence of the intermediate endpoint affect the statistical power. We also assess the choice of weights on data analyses from an SMART design. Weights are implemented to account for the restricted re-randomization, which occurs when select patients are rere-randomized based on their intermediate assessment. In addition, we propose the use of piece-wise hazard rates (instead of a constant hazard rate) for the design with time-to-event final outcome. In previous SMART designs with a time-to-event final endpoint, an overall hazard rate was assumed for each treatment arm. We propose a constant hazard rate for each treatment arm up to the time of the intermediate assessment, implying that at the start of treatment, the hazard rate is the same for each arm within each treatment assignment, and then different hazard rates can be assumed for each treatment arm after the intermediate assessment.

**Methods:** Using simulations, we consider the impact on power from different design parameters, assess the choice of weights on data analyses, and provide guidance on data generation for an SMART trial.

**Results:** Weights are implemented by computing the inverse of the randomization probabilities. We found that when the first stage randomization is 1:1, it is not necessary to consider the first stage randomization when assigning weights, specifically when a robust standard error is used for computing a test statistic. In addition, we found that the choice of randomization ratio for the second stage has much less impact on the power compared to the choice of first stage randomization, as expected, and that a range of possible prevalences for the intermediate endpoint generally provided sufficient power. We provide an R Shiny applet that estimates power based on simulations with options to modify any design parameters.

**Conclusion:** We provide guidance and a software tool for clinical trials with an SMART design with a binary intermediate outcome and a binary final outcome or a time-to-event final outcome.

**CP-3**

A BAYESIAN EXACT PLATFORM DESIGN FOR MULTI-ARM, PHASE 2 TRIALS WITH HISTORICAL CONTROL

**Primary Author:**
ONDREJ BLAHA
Yale University

**Co-Author(s):**
WEI WEI
Yale University

DENISE ESSERMAN
Yale University

The COVID-19 pandemic and associated urgency for rapid testing of new or re-purposed therapeutic agents have exposed many shortcomings of the traditional one-drug at a time approach for clinical trials. Adaptive platform trials refer to trials designed to evaluate multiple interventions (targeting the same endpoints and the same population), with the flexibility to allow for accelerated selection of promising therapeutic agents, early removal of inefficient treatments, and adding new trial arms throughout the duration of the trial. Some form of adaptive randomization can be possibly implemented in adaptive platform trials to allocate more patients to interventions that are performing better than the other interventions (including control). Hence, adaptive platform trials provide an excellent, resource saving alternative compared to more traditional clinical trial designs. Our Bayesian exact design allows for the continuous learning of efficacy and safety profiles from incoming data and with built in safety and efficacy stopping rules, we can accomplish the objectives of safety monitoring, preliminary efficacy assessment, and confirmatory testing within the same trial. Such a design is ideal in situations like the COVID-19 pandemic when many potential antiviral, anti-inflammatory, and anticoagulant therapeutics are available for treatment of similar conditions but have not yet been tested for the population affected by the novel disease.

Compared to the existing methods, our proposed design is innovative in the following aspects: (1) our method allows for the joint monitoring of multiple endpoints such as hospitalization, ventilation use, and other co-primary endpoints, thereby increasing the statistical power of the study. (2) In addition to implementing early stopping rules for efficacy and futility, the proposed design also allows the trial to be stopped early for safety based on already accrued data. (3) Traditionally, statisticians have to rely on complex and time-consuming simulations to develop the statistical design and estimate the sample size of platform trials, our method is based on a recursive relationship that calculates the exact probability of stopping the trial for any cause at any stage of the trial, without relying on simulations, ultimately making our design more rigorous. (4) This newly discovered recursive relationship allows for adjusting the adaptive randomization, further increasing statistical power of the trial. (5) Our adjusted adaptive randomization is based on Bayesian Play, the winner strategy allowing for the patients to be allocated into the most promising arm.
**CP-4**

HOW BIG SHOULD MY PILOT STUDY BE? A METHODOLOGICAL REVIEW

Primary Author:
SASKIA EDDY
Queen Mary University of London

Co-Author(s):
SALLY KERRY
Queen Mary University of London
GILLIAN LANCASTER
Keele University
SANDRA ELDRIDGE
Queen Mary University of London

Introduction: Pilot and feasibility studies are important pieces of preparatory work used to identify whether certain elements in a trial are feasible before conducting a costly, definitive randomized clinical trial. However, as these studies investigate feasibility, the rationale used to justify the sample size differs from the more formal approach used in randomized clinical trials investigating efficacy and effectiveness. While a number of published papers now exist giving guidance for the sample size of pilot and feasibility study, to our knowledge, there has been no assessment of whether this guidance is being used in practice. We aimed to review the planned and achieved sample sizes reported in randomized controlled external pilot and feasibility trial publications and their corresponding protocols, as well as the methods used to justify these sample sizes.

Methods: A systematic search of PubMed was conducted focusing on randomized controlled external pilot and feasibility trials published in 2019. Publications written in English from any medical speciality were included. Publications were screened against pre-defined eligibility criteria. To reduce selection bias, each publication was assigned a randomly generated number and they were screened in ascending order. A pre-piloted data extraction form was used to abstract methodological information from eligible publications. A target sample size of 100 eligible publications was calculated based on the precision of estimating the proportion of trials that reported their planned sample size. The protocol for this work was prospectively registered on PROSPERO (CRD42020189050).

Results: One thousand six hundred four (n = 1604) publications were identified. Publications were screened until 100 potentially eligible publications were identified. Of the 445 publications screened, 360 were excluded, leaving 85 eligible. All eligible publications reported the achieved sample size at randomization. The median achieved sample was 47 (interquartile range (33–61) and range (2–481)). However, only 65 (76%) publications reported a planned sample size, either at the participant level (n = 64) or cluster level (n = 1). The median planned sample size was 60 (interquartile range (40–80) and range (12–360)). A justification for the planned sample size was reported in 76% (n = 65) of the publications. The most common justification was a sample size calculation (n = 21), namely, conducting a power-based calculation (n = 17) or calculating the precision of an estimate (n = 4), followed by referring to published work (n = 15), such as previously conducted research (n = 2) or methodological guidance (n = 13). Protocols were publicly available for only 36% (n = 31) of the publications and 58% (n = 18) of these reported their planned sample size.

Conclusion: It is important authors clearly justify their sample size to enable any conclusions based on the pilot to be appropriately scrutinized by funders, the research community, and the public, as this research has the potential to influence future definitive trials. Also, the sample size justification may assist future researchers who are conducting work in a similar area. These findings act as a reminder to authors and journal editors that research should be reported according to the CONSORT extension for pilot and feasibility study to ensure scientific rigor is achieved from this important preparatory work.

**CP-5**

UTILIZING BENEFIT–RISK ASSESSMENTS WITHIN CLINICAL TRIALS

Primary Author:
NIKKI TOTTON
University of Sheffield

Co-Author(s):
STEVEN JULIOUS
University of Sheffield
DYFRIG HUGHES
Bangor University
JONATHAN COOK
University of Oxford

The Medical Research Council and the National Institute for Health Research fund randomized controlled trials, mostly within the United Kingdom, to provide evidence to inform UK national policy decisions. These trials can be designed to assess the objectives of either superiority, equivalence or non-inferiority which is determined by the research question and the primary outcome selected. However, in some trials, there may be more than one outcome of importance to consider, particularly when using equivalence or non-inferiority designs. This multi-dimensionality is often not considered when the focus is on a single primary outcome.
Benefit–risk methodology is commonly used within the regulatory setting with much of the available information and guidance relating to regulatory drug trials conducted by innovator pharmaceutical companies. In the context of Medical Research Council/National Institute for Health Research trials, the studies are of health technologies (not just drugs) and often of therapies that are already licensed; therefore, additional considerations such as economic outcomes may be required.

This project aims to provide guidance in the use of benefit–risk methods to assess outcome trade-offs (qualitatively or quantitatively) which includes selecting the appropriate trial design, situations to consider using a benefit–risk method within trial design and the key information to include when reporting on trials which have included benefit–risk methods.

**Methods:** The guidance was produced using the following methods: (1) web-based survey of current practice to known researchers in the field and appropriate targeted email lists, (2) rapid methodological review of currently available methods along with their situations of recommended use, and (3) 2-day expert consensus workshop to discuss with a range of stakeholders their experiences and recommendations.

**Results:** A list of 19 factors was created to aid researchers in selecting the most appropriate trial design which center on the following areas: population, intervention, comparator, outcomes, feasibility, and perspectives.

When designing trials, six key rationales were identified that indicate a benefit–risk methodology should be at least considered: (1) the success of the trial is dependent on more than one outcome, (2) important outcomes within the trial are in competing directions, that is, a health technology is better on one outcome but worse on another, (3) to allow patient preferences to be included and directly influence trial results, (4) provide transparency on subjective recommendations from a trial, (5) provide consistency in an approach to presenting results from a trial, and (6) to synthesize multiple outcomes into a single metric.

To ensure information reported which relates to the benefit–risk is consistent and clear, a checklist of reporting items has been developed. There are five pieces of information recommended to be include when reporting on the trial design (heading, use of benefit–risk term, appropriate plan, anticipated benefits and risks, and discussion with patients) and a further two when reporting the results of a trial (summary table of benefits and risks, Quality-Adjusted Life Years reporting in terms of the benefit–risk).

**Conclusion:** This research provides guidance for researchers designing trials to ensure the research results are selected, completed, and reported appropriately to support policy decisions.

---

**CP-6**

**GUIDELINES FOR THE CONTENT OF STATISTICAL ANALYSIS PLANS FOR EARLY PHASE CLINICAL TRIALS**

**Primary Author:**

VICTORIA HOMER

Cancer Research UK Clinical Trials Unit, University of Birmingham

**Co-Author(s):**

CHRISTINA YAP

The Institute of Cancer Research

SIMON BOND

Cambridge Clinical Trials Unit

CARROL GAMBLE

Liverpool Clinical Trials Centre

DEBORAH STOCKEN

Leeds Institute of Clinical Trials Research

MATTHEW SCHIPPER

Departments of Radiation Oncology and Biostatistics

SAMANTHA HINSLEY

Cancer Research UK Glasgow Clinical Trials Unit

GRAHAM WHEELER

Cancer Research UK & UCL Cancer Trials Centre

SARAH BROWN

Clinical Trials Research Unit

KATRINA WALKER

Clinical Trials Research Unit

Early phase clinical trials aim to determine the safety and initial indicators of efficacy (phases 1 and 2, respectively) of interventions before subsequent phase 3 trials are conducted. The undertaking of later phase trials is often a lengthy and costly process predicated on accurate and robust conclusions from early phase trials. The design, conduct, and analysis of these early phase trials should be performed to the highest standards of rigor and quality, ensuring correct decisions is taken forward. ICH E9 guidelines state “although the early phases of drug development consist mainly of clinical trials that are exploratory in nature, statistical principles are also relevant.” [1]

Detailed statistical analysis plans improve transparency, trial quality, and robustness. Guidelines for the content of statistical analysis plans for randomized trials were published in 2017 [2]. These guidelines acknowledged that despite some recommendations being transferable, specific consideration and guidance would be required for early phase trials.

To date, work on the early phase statistical analysis plan guidelines included a literature search (identifying peer-reviewed publications of applicable guidelines, and
example trial protocols and statistical analysis plans); a review of the EQUATOR network reporting guidelines; and an assessment of regulatory and funder requirements. In addition, a survey of UK Clinical Research Collaboration registered Clinical Trials Units has been undertaken (40/53 Clinical Trials Units responded) to identify those conducting early phase trials (n = 21) and establish their current practice regarding authoring and content of statistical analysis plans.

Following these reviews, incorporating statistical analysis plans provided by Clinical Trials Units, and in light of the requirements of pharmaceutical and academic statisticians, regulators, and funders, the early phase trial extension to existing statistical analysis plan guidelines has been developed. These guidelines were discussed at a review meeting attended by international academic, regulatory, and pharmaceutical representatives and subsequently updated.

This extension uses the general statistical analysis plan structure from the Gamble et al. paper [2], with the current draft keeping 31 of the original items unchanged, amending 24, with a further 12 new items proposed. These alterations include items specific to early phase trials (e.g. design information, dose escalation decisions, and deviations from trial sample sizes), and an update to the outcomes section regarding estimates following the adoption of ICH E9 (R1).

Presently, the guidelines are being trialed in several UK Clinical Trials Units to evaluate suitability for purpose, and identify areas for improvement. Subsequently, feedback from these Clinical Trials Units will be incorporated into the final guidelines, which are anticipated to be finalized and published around the date of SCT 2021.

This work supports clinical trial statisticians, trialists, and peer reviewers to facilitate an improvement in the quality of analysis, the reproducibility of methods and results, and the robustness of conclusions. An outline of the project and overview of the guidelines will be presented.

This project is funded by the National Institute for Health Research Clinical Trials Unit Support Funding scheme. The views expressed are those of the authors and not necessarily those of the National Institute for Health Research or the Department of Health and Social Care.

[1] European Medicines Agency. ICH E9—Statistical Principles for Clinical Trials.

[2] Gamble C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017.

CP-7

DISMANTLING THE FRAGILITY INDEX: A DEMONSTRATION OF STATISTICAL REASONING

Primary Author:
GAIL E POTTER
The Emmes Company, LLC

The Fragility Index has been introduced as a complement to the p-value to summarize the statistical strength of evidence for a trial’s result [1]. The Fragility Index (FI) is defined in trials with two equal treatment group sizes, with a dichotomous or time-to-event outcome, and is calculated as the minimum number of conversions from non-event to event in the treatment group needed to shift the p-value from Fisher’s exact test over the 0.05 threshold. It has been applied to evaluate the robustness of hundreds of trials and was applied in a paper [2] arguing that results from an open-label non-randomized trial of hydroxychloroquine are “statistically robust” according to the FI—despite numerous statistical weaknesses in the trial. As the index lacks a well-defined probability motivation, its interpretation is challenging for consumers. I analyze the FI from a perspective of probability by calculating the posterior probability of a treatment effect—a more scientifically interesting and relevant quantity than the p-value [3]. I demonstrate that when model assumptions hold, the FI promotes an incorrect intuition of probability which can lead to incorrect decision-making. As such, it is a fundamentally flawed statistical tool. I also demonstrate that it does not perform well at quantifying robustness of results to violation of model assumptions (e.g. bias from missing data). The exploration highlights weaknesses of the p-value as a tool for quantifying evidence, and I contextualize this and the FI within current debate around the p-value and the null hypothesis significance testing paradigm. Altogether, the FI creates more confusion than it resolves and does not promote statistical thinking. While the explicit goal of this paper is to discourage its use, the broader goal is to demonstrate and promote statistical reasoning. Holistic interpretation of a trial’s results—taking into account effect estimates, confidence intervals, p-values, and carefully designed sensitivity analyses—is advocated to evaluate and communicate statistical evidence. My analysis demonstrates how to evaluate a novel statistical tool in a thoughtful and systematic way. This process of thoughtful statistical reasoning is greatly needed in order to advance rigorous and ethical trials in the pandemic era and beyond.

[1] Walsh M, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. Journal of clinical epidemiology 2014; 67(6): 622–628.

[2] Chow R, Elsayed S and Lock M. How robust are the results of one of the first positive trials exploring hydroxychloroquine for treatment of COVID-19? medRxiv 2020. DOI: 10.1101/2020.05.06.20093237.
FROM CLINICAL TRIAL TO OBSERVATIONAL STUDY: LONG-TERM RETENTION IN THE TODAY AND TODAY2 STUDIES

Primary Author:
BRIAN K BURKE
The George Washington University

Co-Author(s):
MIHILI GUNARATNE
The George Washington University
PAUL MCGUIGAN
Case Western Reserve University
JENNA MEYER
Saint Louis University
MELINDA TUNG
The George Washington University
MEGHAN TURNEY
The George Washington University

The TODAY (Treatment Options for Type 2 Diabetes in Adolescents and Youth) study was a multi-site randomized clinical trial comparing three interventions to treat adolescents and youth with type 2 diabetes. Individuals aged 10–17 years with diabetes duration of less than 2 years and BMI ≥85th percentile were eligible to participate at one of 15 clinical centers. The TODAY study (2004–2011) was followed by TODAY2 (2011–2020), an observational follow-up study tracking the progression of type 2 diabetes and related comorbidities and complications in the TODAY cohort. TODAY randomized 699 participants with 520 (74.4%) remaining enrolled at the end of TODAY2. At TODAY baseline, study participants were largely of low socio-economic status (41.5% with household income < US $25,000) and from racial/ethnic minority groups (79.5%). The long-term retention of this unique cohort was critical to study success. Retention was operationalized at the study level via dedicated committees, procedures, and practices while also incorporating participant-targeted strategies such as travel reimbursement, monetary incentives, and access to medical care and supplies. Committees included groups dedicated to development of participant-facing materials, standard diabetes education, oversight of study protocol, adherence, and transition between study phases. Clear communication, flexible visit schedules, and comprehensible reporting on study findings and individual results directly to participants were important components of the overall retention strategy. The combined efforts of study staff, including study coordinators, certified diabetes educators, physician investigators, physical activity and nutritional leaders, and psychologists were instrumental in maintaining participant engagement. Since the duration of participation spanned nearly two decades, it was necessary to update retention approaches to reflect the needs of a cohort aging from adolescence to young adulthood. Approaches to participant engagement needed to be age-appropriate. Through adaptability, staff dedication, and an emphasis on retention at both the operation and individual level, the study was able to achieve high retention rates. The study group’s experiences with participant engagement and retention from TODAY through TODAY2 provide valuable insights and lessons learned that can benefit others conducting long-term clinical research studies.

HOW CLINICAL TRIAL MANAGEMENT SYSTEMS CAN HELP REDUCE COORDINATOR BURDEN

Primary Author:
GLORY L KOERBEL
University of Pittsburgh

Co-Author(s):
JASON KOJTEK
University of Pittsburgh
DAVID YANKOVICH
University of Pittsburgh
SUSAN SPILLANE
University of Pittsburgh
ANDREW ALTHOUSE
University of Pittsburgh
KALEAB ABEBE
University of Pittsburgh

The challenges facing clinical research coordinators are well-documented. In spite of technological advances and creative marketing strategies touting a plethora of data management systems and tools, coordinators often struggle to perform all of the necessary tasks of multiple complex clinical trials with sparse budgets and reductions in staffing.

This presentation will focus on the impact of integrated study-specific clinical trial management systems, developed, and implemented by the University of Pittsburgh’s Center for Clinical Trials & Data Coordination, on research coordinator burden. The
Center for Clinical Trials and Data Coordination team is comprised of doctoral level statisticians with expertise in epidemiology, clinical trial design, conduct and analysis; clinical research nurse coordinators with a combined total of 40 years of research experience who also serve as study monitors; and multiple developers and statistical analysts. Thus, Center for Clinical Trials and Data Coordination collaborators benefit from a multi-disciplinary expertise from “conception to closeout.”

In this talk, we highlight aspects of our web-based clinical trial management system that reduce the burden on research coordinators, specifically with enhanced communication, decreased duplication of effort, and improved data quality and patient safety.

**Enhanced Communication:** Access to the portal can be granted not only to site coordinators, but also to participants (where indicated to complete surveys and diary entries), clinicians, medical monitors, pharmacists, and lab personnel. Direct data entry is centralized and can be reviewed by the coordinator in real-time. By creating specific access to each type of user, exposure to data can be controlled. Participants can only view data they enter while the coordinator and PI can review data entered by their local participants as well as members of their research team. Coordinators can access PDF versions of study documents and case report forms directly on the clinical trial management system. The system is easily adaptable to accommodate remote visits.

**Duplication of Effort:** The clinical trial management system is user friendly and streamlines the work of the coordinator. Participant specific data are prepopulated on study drug order forms and shipping inventories that are converted to PDF documents and can be printed to include in pharmacy orders and lab specimen shipments. Likewise, labels for lab kits and study drug containers can be created. Automatically generated emails are programmed in the system to alert key personnel of lab shipments, patient randomization and submission of a Serious Adverse Event form, to name a few. Supplies can also be ordered through the system.

**Data Quality and Patient Safety:** We utilize a risk-based monitoring approach with a document upload feature that permits the study monitor to confirm source document and informed consent verification without traveling to the site. Feedback in real-time decreases data change requests and can eliminate recurrent errors. An adjudication portal is also available in the system to facilitate prompt reporting and confirmation of Serious Adverse Events.

In addition to the above, we will highlight feedback from research coordinators that use our clinical trial management system on several clinical and behavioral trials.
trial feasibility. When a clinical trial moves into the intervention and follow-up phases, it is imperative to cross-check and monitor data completion and accuracy for both study efficacy elements (i.e. study endpoints or outcomes) and participant safety elements (i.e. adverse or serious adverse events) to ensure the data elements are not misreported, under-reported, or late reported, which have a huge impact on trial success. The Corporate Data Warehouse in the Department of Veterans Affairs provides such a unique comprehensive administrative and clinical data source. Administrative information in the Corporate Data Warehouse typically includes participant demographics, socio-economic data, and geographical location while clinical information usually includes inpatient admission and discharge dates, outpatient clinic visits visits dates, patient diagnoses, medications, medical procedures and surgeries, and laboratory dates and results. The Veterans Affairs Cooperative Studies Program has a long history of conducting large multi-center clinical trials within the Veterans Affairs healthcare system and has been a strong proponent of using Veterans Affairs’s electronic health records (administrative and clinical) to improve clinical trial design and execution. In this report, we are presenting the Perry Point, Maryland, Cooperative Studies Program Coordinating Center’s approach and experience of using Corporate Data Warehouse as an auxiliary data source to improve multi-center clinical trial practice by focusing on trial design, participant recruitment planning, and trial monitoring.

CP-II
USE OF COMPREHENSIVE SOLICITED EVENT REPORTING FOR SAFETY MONITORING IN CLINICAL TRIALS WITH STANDARD OF CARE DRUGS

Primary Author:
CHRISTINA FU
The Emmes Company, LLC

Co-Author(s):
VICTORIA EASTON
The Emmes Company, LLC

ELIZABETH REARDON
The Emmes Company, LLC

KELSEY HUMPHREY
The Emmes Company, LLC

DAVID LIPSCOMB
The Emmes Company, LLC

JIA-YUH CHEN
The Emmes Company, LLC

SITORA TURDALIEVA
The Emmes Company, LLC

KAREN MARTZ
The Emmes Company, LLC

ELIZABETH PAYNE
The Emmes Company, LLC

RADHIKA KONDAPAKA
The Emmes Company, LLC

Safety reporting in traditional clinical trials is often accomplished through investigator’s independent identification of events meeting Adverse Event reporting criteria. In some study settings, such as observational or standard of care-based research, the subjectivity in this reporting approach can result in gaps or inconsistencies in safety data, limiting the ability to confidently confirm both presence and absence of key events. The Emmes contract research organization has adopted a new approach for clinical research studies to improve upon the traditional, passive, non-specific collection of adverse events with a more targeted solicitation of study-specific Events of Special Interest. Events of Special Interests may be derived from the drug labels, package inserts, or known disease indications, and ensure consistency and uniformity in safety data collection across sites and similar studies. Events of Special Interest reporting aims to target specific events and avoids over or under reporting by sites. Events of Special Interests also reduce the burden of site reporting of irrelevant events, the need to determine if an event is or is not reportable, such as in a sick patient population, and finally supports efficient data and safety monitoring in alignment with the Food and Drug Administration Investigational New Drug Application Safety reporting guidance. The Emmes contract research organization has adopted this approach to ensure comprehensive data collection in support of improved labeling for marketed products in pediatric and adolescent populations. It has proven most effective in observational or standard of care studies, particularly those with sick and/or inpatient populations. Safety event data are bolstered with site required confirmation of occurrence in an unequivocal No or Yes manner, along with the addition of severity, causality, and association assessments by the clinician/subject matter expert. Events of Special Interests are study-defined events, aiding study staff or programmatic searches of health records in the identification and abstraction of event data. Events of Special Interests also allow researchers to more accurately tabulate event frequency and rates, as well as assist sponsors in identifying site reporting imbalances. Clinical research is constantly evolving, driven by the need to optimize data collection and reduce time, effort, and cost, which is particularly important for trials being implemented during the COVID-19 pandemic. The traditional methods used to collect safety data require a high level of effort to collect, review by numerous subject matter
experts, and can often result in the collection of large amounts of data that are irrelevant to research objectives. This utilization of safety reporting and analysis to include Events of Special Interests has been shown to significantly reduce the level of effort to collect, subject matter expert review, MedDRA code, and tabulate safety data. By improving data collection consistency and uniformity, utilization of Events of Special Interest reporting has improved the overall process of safety event reporting in our clinical research, increasing the efficacy of study data utilized to inform regulatory authorities.

CP-12
CO-ENROLLMENT OF PATIENTS IN TRIALS OF COMPLEX SURGICAL INTERVENTIONS: RECOMMENDATIONS FROM THE INTACT TRIAL

Primary Author:
GEMMA AINSWORTH
Leeds Institute of Clinical Trials Research

Co-Author(s):
NEIL CORRIGAN
Leeds Institute of Clinical Trials Research

JULIE CROFT
Leeds Institute of Clinical Trials Research

JULIA BROWN
Leeds Institute of Clinical Trials Research

DAVID JAYNE
Leeds Institute of Biological and Clinical Sciences

DEBORAH STOCKEN
Leeds Institute of Clinical Trials Research

Issue: Co-enrollment of patients in trials of complex interventions is an emerging phenomenon. Co-enrollment can provide efficiencies, pertinent in the era of a pandemic, but is not without issue. Management of patient co-enrollment, especially in trials investigating multiple interacting components, such as surgical randomized controlled trials, is considered.

Setting: IntAct is an international surgical trial which has recruited throughout the COVID pandemic. It is a prospective, parallel group, randomized controlled trial comparing surgery with intraoperative fluorescence angiography against standard care (surgery alone) to determine the effect on anastomotic leak in patients undergoing resection for rectal cancer. Co-enrollment of patients has been requested by nine other trials to date, ranging from interventional drug trials to observational studies. There is minimal published literature regarding co-enrollment to inform discussion and decision-making by the IntAct Trial Management group.

Background: Co-enrollment of patients in clinical trials poses a number of potential issues spanning ethical, safety, statistical, and practical concepts, which require careful consideration. Effort has been made to quantify the potential impacts of co-enrollment on statistical power (Myles, 2014). For example, the interaction of treatment effects and substantial or imbalanced co-enrollment have been shown to possess the potential for a large detrimental effect on the sample size. In practice, the nature of such an interaction, or the level of co-enrollment, will not be known in advance. Furthermore, where the intervention is complex in nature, such as in surgical trials of multiple component parts like IntAct, not only is there the possibility of an interaction between the treatment effects of co-enrolling interventions, but also of how the co-enrolling intervention may alter the individual intervention components. There is no published guidance on how to assess these risks a priori.

Nonetheless, there are also many potential benefits to allowing co-enrollment. These include increased availability of research opportunities to patients and increased efficiencies to sites when patients contribute to multiple research projects. Disallowing co-enrollment could pose a risk to a trial’s recruitment by narrowing the pool of potential participants. Likewise, there is no published guidance on how to assess these benefits a priori.

Key considerations for co-enrolling trials: the IntAct Trial Management group decision-making process was informed by discussions regarding:

- Internal validity: the ability of the co-enrolling intervention to change the typical operative setting, as well as the post-operative care pathway.
- Potential interaction of treatment effects: particularly with regard to the primary endpoint and participant safety.
- Generalizability: the possible implications for the trial results.
- Recruitment impact: the probable restrictions imposed if co-enrollment was disallowed.
- Compliance: the likely effects of co-enrollment on data collection, follow-up visits and withdrawal, and ethical considerations of overburdening participants.

We present our rationale and recommendations for guiding Trial Management group discussions when considering and approving co-enrollment of patients in clinical trials. By presenting the experiences of the IntAct international surgical trial Trial Management group, we provide a practical reference for trials considering co-enrollment.
THE IMPACT OF PATIENT ENGAGEMENT ON TRIALS AND TRIALISTS: AN INTERVIEW STUDY WITH IMPACT Awardees

Primary Author:
STUART G NICHOLLS
Ottawa Hospital Research Institute

Co-Author(s):
ZARAH MONFAREDI
Ottawa Hospital Research Institute
CHANTELLE GARRITTY
Ottawa Hospital Research Institute
ANNE LYDDIATT
Patient Partners in Arthritis
ALIES MAYBEE
Patient Partner
JUSTIN PRESSEAU
Ottawa Hospital Research Institute
BEVERLEY SHEA
Ottawa Hospital Research Institute
DEAN A FERGUSSON
Ottawa Hospital Research Institute

Background: The Canadian Strategy for Patient-Oriented Research was established in 2011 as a national coalition of stakeholders aiming to ensure that research is guided by patients. Despite this, the literature evaluating the impact of patient engagement on research and research teams is still in its infancy. In November 2014, the Ontario Strategy for Patient-Oriented Research Support Unit announced a call for funding to support translational patient-oriented research, with the explicit requirement to involve patients; providing an opportunity to examine specific impacts that patient engagement has had on the research produced as well as members of the research teams.

Objective: To examine the impact(s) that patient engagement has had on the research produced, and the researchers, research teams, and patient partners of Ontario Strategy for Patient-Oriented Research Support Unit Impact Awardees.

Methods: Semi-structured interviews were conducted with principal investigators and members of the research teams (including patient partners), all of whom were involved in randomized controlled trials, and who received funding through an Ontario Strategy for Patient-Oriented Research Support Unit IMPACT Award. The interview guide included (1) discussion of patient engagement and roles; (2) perceived impact of patient engagement on the research; and (3) perceptions regarding training needs. Interviews were audio-recorded and transcribed verbatim. Transcripts underwent thematic analysis with text coded and labeled inductively.

Results: Ten interviews were completed with individuals from seven projects. Interviews were 49 min on average (range = 28–66). While several studies were informed by a theoretical framework, such as the Knowledge to Action cycle, others took a more à-la-carte approach, picking elements of guidance as needed. Consequently, the number of patients, the methods employed, and stages of engagement varied. Interviewees indicated the need to engage the range of stakeholders affected by the study, with examples including caregivers, clinicians, as well as other healthcare professionals (such as paramedics and nurses). Successes generated by patient engagement included ongoing participation of patients engaged as partners, and development of relationships. Interviewees indicated a perception that patient engagement improved the quality of the trial, but specific examples were limited. Impact on process was commonly cited, with influence on the researcher and their awareness of issues such as power dynamics. Challenges experienced included the lack of guidance for compensation, identification of patient partners in non-clinical areas (e.g. public health), and ongoing engagement during the trial. Interviewees indicated a need to be flexible in arrangements, committed personnel, and time to conduct patient engagement, as well as a training need to develop communication skills such as active listening and trust building.

Conclusion: Perceptions of impact or success commonly related to impact on the researchers or research team and changes to the research process, rather than the research products/outputs. Many noted the need to build relationships, and that the impact may be on subsequent studies due to lessons learned. Advice often focused on skill-based elements such as developing communication skills and respectful interactions rather than improving patient knowledge of research. Future work will develop tools to capture this impact.

FACILITATING STUDY COLLABORATION IN YOUR PAJAMAS: REPLICATING A VIRTUAL OFFICE SPACE WHILE TELEWORKING

Primary Author:
DOUGLAS E LAMMIE
United States Department of Veterans Affairs

Co-Author(s):
KIMBERLY CARLSON
US Department of Veterans Affairs
CAITLIN AUTHIER
US Department of Veterans Affairs
CHRISTOPHER SCHROTH  
*US Department of Veterans Affairs*

Successful multi-site clinical trials are not exempt from experiencing challenge, and when they arise, they are strategically mitigated. This is especially true when a global pandemic disrupts a large, complex research enterprise. In March 2020, Study Sponsor staff at Hines Cooperative Studies Program (Hines CSP) transitioned to “work-from-home” full-time and were faced with several risks.

CSP 2008 identified these risks as: inaccurate information to be disseminated to site staff, delays in study communications, and an eroding of the type of teamwork and relationships that comes from working collaboratively in an office space.

The CSP 2008 had an immediate challenge which needed a clear solution: how would the Sponsor Staff foster teamwork and collaboration in the sudden absence of a shared office space?

CSP 2008 adopted and embraced a workplace collaboration platform (MS Teams) to mitigate the unintentional risks that the study team identified may occur from full-time telework. Since its adoption, CSP 2008 has developed several best practices/key learnings/lessons learned for integrating collaboration software into clinical trial management. Some of these include:

1. A central communication space for all sponsor-level staff to troubleshoot issues in real-time.
2. The ability to efficiently yield questions from site-level staff, discuss and issue documented guidance within five business days within the software ecosystem.
3. Our ability to manage overall workload via project management tools to track action items and ensure accountability with sponsor staff.

Using a new communication platform to assist in managing clinical trials remotely, it has allowed CSP 2008 to improve operational efficiencies, allow for clear communication across the study, and improve collaboration, coordination harmonization throughout the study sponsor staff.

**Co-Author(s):**  
SALLY HOPEWELL  
*University of Oxford*  
SUSAN J DUTTON  
*University of Oxford*

**Objectives:** Pre-specified progression criteria can inform the decision to progress from an external randomized pilot trial to a definitive Randomized Controlled Trial. We assessed the characteristics of progression criteria reported in external randomized pilot trial protocols and results publications, including whether progression criteria were specified a priori and mentioned in prepublication peer review reports.

**Methods:** In this methodological review, four journals indexed in PubMed that are known to publish pilot trials were searched: British Medical Journal Open, Pilot and Feasibility Studies, Trials and Public Library of Science One. Articles were single screened against eligibility criteria. Eligible publications reported external randomized pilot trial protocols or results, were published between January 2018 and December 2019, and reported progression criteria. We double-data extracted 25% of the included publications. The primary outcome was the progression criteria characteristics, including the areas of feasibility that progression criteria were based on, their rationale or justification, and who established and assessed the progression criteria. One set of secondary objectives were to assess whether the progression criteria reported in pilot trial results publications were specified a priori in a published protocol or trial registration and whether the results publication reported the intention to progress to a definitive randomized clinical trial. We also assessed the extent and context in which progression criteria were discussed in prepublication peer review reports where available for the protocol and results publications.

**Results:** We included 160 publications (123 protocols and 37 completed trials). Recruitment and retention were the most frequent indicators contributing to progression criteria. Progression criteria were mostly reported as distinct thresholds (e.g. achieving a specific target; 133/160, 83%). Less than a third of the planned and completed pilot trials that included qualitative research reported how these findings would contribute toward progression criteria (34/108, 31%). The publications seldom stated who established the progression criteria (12/160, 7.5%) or provided rationale or justification for progression criteria (44/160, 28%). Nearly a quarter of publications reported who would be involved in assessing progression criteria (35/160, 22%). Published trial protocols or trial registrations were available for all but one completed pilot trial; over half did not report progression criteria (22/36, 61%). Most completed pilot trials reported that a future randomized clinical trial is feasible or the intention to
proceed (30/37, 81%), but less than half strictly met all their progression criteria (17/37, 46%). Prepublication peer review reports were available for 153/160 publications (96%). Peer review reports for 86/153 (56%) publications mentioned progression criteria, with peer reviewers of 35 publications commenting that progression criteria appeared not to be specified.

Discussion: Many external randomized pilot trial publications did not adequately report or propose prespecified progression criteria to inform whether to proceed to a future definitive randomized clinical trial. Only external randomized pilot trial publications were included in this review, and it is unclear whether the findings can be generalized to other external feasibility study designs. Clearly, evidence-based recommendations for the use and reporting of progression criteria in external randomized pilot trials are required.

**CP-16**

**STRATIFIED RANDOMIZATION FOR PLATFORM TRIALS WITH DIFFERING EXPERIMENTAL ARM ELIGIBILITY**

**Primary Author:**
SUBODH R SELUKAR
University of Washington

**Co-Author(s):**
SUSANNE MAY
University of Washington
MEGAN OTHUS
Fred Hutchinson Cancer Research Center

Platform trials facilitate efficient use of resources by comparing multiple experimental agents to a common standard of care arm. They can accommodate a changing scientific paradigm within a single trial protocol by adding or dropping experimental arms—critical features for trials in rapidly developing disease areas such as COVID-19 or cancer therapeutics. However, in these trials, efficacy and safety issues may render certain participant subgroups ineligible to some experimental arms, and methods for stratified randomization do not readily apply to this setting of differing experimental arm eligibility. We motivate this setting with the LEAP trial, a platform trial for acute myeloid leukemia in older adults. When experimental arms differ in eligibility, existing methods for stratified randomization require changes in trial-wide eligibility, which affects trial accrual and generalizability. This work describes how to extend conventional randomization methods to account for varying experimental arm eligibility. We suggest modifying block randomization by including experimental arm eligibility as a stratifying variable, and we suggest modifying the imbalance score calculation in dynamic balancing by performing pairwise comparisons between each eligible experimental arm and standard of care arm participants eligible to that experimental arm. We also briefly discuss the impact of differing eligibility on the efficiency of platform trials as measured by the size of the common standard of care arm.

**CP-17**

**COVARIATE ADJUSTMENT IN SUBGROUP ANALYSES OF RANDOMIZED CLINICAL TRIALS: A PROPENSITY SCORE APPROACH**

**Primary Author:**
SIYUN YANG
Duke University

**Co-Author(s):**
FAN LI
Duke University
LAINE E THOMAS
Duke University School of Medicine
FAN LI
Yale University School of Public Health

**Background:** Subgroup analyses are frequently conducted in randomized clinical trials to assess evidence of heterogeneous treatment effect across patient subpopulations. Although randomization balances covariates within subgroups in expectation, chance imbalance may be amplified in small subgroups and harm the precision of subgroup analyses. Two main approaches for covariate adjustment include analysis of covariance and propensity score weighting in Randomized Controlled Trials. In this article, we develop propensity score weighting methodology to improve the precision and power of subgroup analyses by eliminating chance imbalances.

**Methods:** We extend the propensity score weighting methodology to subgroup analyses by fitting a logistic regression propensity model with covariate–subgroup interactions. We show that overlap weighting exactly balances the covariates with interaction terms in each subgroup. Extensive simulations are performed to compare the operating characteristics of unadjusted estimator, different propensity score weighting estimators and the analysis of variance estimator. We apply these methods to the HF-ACTION trial to evaluate the effect of exercise training on 6-min walk test in several prespecified subgroups.

**Results:** Efficiency of the adjusted estimators is higher than that of the unadjusted estimator. The propensity score weighting estimator is as efficient as analysis of variance, and may be more efficient when subgroup sample size is small (N < 125), or when
outcome model is mis-specified. The weighting estimators with full-interaction propensity model consistently outperform traditional main-effect propensity model.

**Conclusion:** Propensity score weighting serves as a transparent alternative to adjust important covariates in subgroup analyses of randomized clinical trials. It is important to include the full set of covariate–subgroup interactions in the propensity score model.

**CP-18**

**BAYESIAN MULTI-VARIATE NETWORK META-ANALYSIS MODEL FOR THE DIFFERENCE IN RESTRICTED MEAN SURVIVAL TIMES**

**Primary Author:**
XIAOYU TANG  
Boston University

**Co-Author(s):**
XIAOYU TANG  
Boston University
LUDOVIC TRINQUART  
Boston University

Network meta-analysis is essential for clinical decision-making. Network meta-analysis enables inference for all pair-wise comparisons between interventions available for the same indication, using both direct evidence and indirect evidence. In randomized trials with time-to-event outcome data, such as lung cancer data, conventional Network meta-analysis methods rely on the hazard ratio and the proportional hazards assumption, and ignore the varying follow-up durations across trials. We introduce a novel multi-variate Network meta-analysis model for the difference in restricted mean survival times. Our model synthesizes all the available evidence from multiple timepoints simultaneously and borrows information across timepoints through within-study covariance and between-study covariance for the differences in restricted mean survival time. We derived the within-study covariance and estimated the model under the Bayesian framework. We evaluated our model by conducting a simulation study. Our multiple-timepoint model yields lower mean squared error over the conventional single-timepoint model at all timepoints, especially when the availability of evidence decreases. We illustrated the model on a network of randomized trials of second-line treatments of advanced non-small-cell lung cancer. Our multiple-timepoint model yielded increased precision and detected evidence of benefit at earlier timepoints as compared to the single-timepoint model. Our model has the advantage of providing clinically interpretable measures of treatment effects.

**CP-19**

**ON THE DESIGN OF STRATIFIED BIOMAKER TRIALS WITH MEASUREMENT ERROR**

**Primary Author:**
SUSAN HALABI  
Duke University

**Co-Author(s):**
CHEN-YEN LIN  
Eli Lilly
AIYI LIU  
NIH

A major emphasis in precision medicine is to optimally treat subgroups of patients who may benefit from certain therapeutic agents. And as such, enormous resources and innovative clinical trials designs in oncology are devoted to identifying predictive biomarkers. Predictive biomarkers are ones that will identify patients that are more likely to respond to specific therapies, and they are usually discovered through retrospective analysis from large randomized phase 2 or phase 3 trials. One important design to consider is the stratified biomarker design, where patients will have their specimen obtained and the biomarker status known prior to random assignment and regardless of their biomarker status they will be randomized to either an experimental arm or the standard of care arm. The stratified biomarker design can be used to test for a treatment–biomarker interaction in predicting an outcome. Many biomarkers are derived from tissues from patients, and hence, their levels may be heterogeneous. As a result, biomarker levels may be measured with error and this would have an adverse impact on the power of a stratified biomarker clinical trial. We show that the naïve test is biased and will provide bias-corrected estimators for computing the sample size and the 95% confidence interval when testing for a treatment–biomarker interaction. We will apply the sample size formula in the design of a phase 3 clinical trial in renal cancer.

**CP-20**

**MULTI-OUTCOME TRIALS WITH A GENERALIZED NUMBER OF EFFICACIOUS OUTCOMES**

**Primary Author:**
MARTIN LAW  
University of Cambridge

**Co-Author(s):**
MICHAEL J GRAYLING  
Newcastle University
Single-arm clinical trials typically measure a number of outcomes of interest. Investigators may wish to focus equally on multiple outcomes, rather than just one outcome. Existing multi-outcome designs focus almost entirely on one of two approaches: either evaluating whether all outcomes show evidence of efficacy or evaluating whether at least one outcome shows evidence of efficacy. While a small number of authors provide multi-outcome designs that evaluate when a general number of outcomes show promise, such designs are single-stage in nature only. We therefore propose two designs, of group sequential and drop the loser form, that permit novel flexibility in a multi-stage setting.

Previous such multi-outcome multi-stage designs have allowed a maximum of two outcomes; our designs thus also extend previous related proposals by permitting any number of outcomes. We compare the new multi-stage design to a multi-stage design with a composite outcome and compare the new multi-outcome drop the loser design to a multi-outcome single-stage design. We compare designs in terms of expected sample size, expected number of measurements, and probability of rejecting the null hypothesis when true outcome effects vary.

Our proposed designs are superior in various circumstances. The proposed approaches allow investigators to measure, at least initially, a range of outcomes while reducing the high costs that may be associated with such trials. Furthermore, these approaches offer novel flexibility in the area of multiple-outcome clinical trials, allowing investigators to specify any number of outcomes for which promise must be shown. These designs generalize existing designs, which are special cases in comparison as they require promise to be shown on either all outcomes or at least one outcome.

When testing interventions for a novel disease or condition, it may be challenging to select the most appropriate outcomes for which to power the planned trial. Our proposed approach can ease this challenge.

We encourage investigators to use these designs by providing details of an R package that allows users to design trials of their own, using these approaches.

CP-21
POWER CONSIDERATIONS FOR GENERALIZED ESTIMATING EQUATIONS ANALYSES OF FOUR-LEVEL DESIGNS
Primary Author: XUEQI WANG Duke University
Co-Author(s): ELIZABETH L TURNER Duke University JOHN S PREISSER The University of North Carolina at Chapel Hill FAN LI Yale University

In this article, we develop methods for sample size and power calculations in four-level intervention studies when intervention assignment is carried out at any level, with a particular focus on cluster-randomized trials. Cluster-randomized trials involving four levels are becoming popular in healthcare research, where the effects are measured, for example, from evaluations (level 1) within participants (level 2) in divisions (level 3) that are nested in clusters (level 4). In such multi-level cluster-randomized trials, we consider three types of intraclass correlations between different evaluations to account for such clustering: that of the same participant, that of different participants from the same division, and that of different participants from different divisions in the same cluster. Assuming arbitrary link and variance functions, with the proposed correlation structure as the true correlation structure, closed-form sample size formulas for randomization carried out at any level (including individually randomized trials within a four-level clustered structure) are derived based on the generalized estimating equations approach using the model-based variance and using the sandwich variance with an independence working correlation matrix. We demonstrate that empirical power corresponds well with that predicted by the proposed method for as few as eight clusters, when data are analyzed using the matrix-adjusted estimating equations for the correlation parameters with a bias-corrected sandwich variance estimator, under both balanced and unbalanced designs.

CP-22
WOULD THE RECOMMENDED DOSE HAVE BEEN DIFFERENT USING NOVEL DOSE-FINDING DESIGNS? COMPARING DOSE-FINDING DESIGNS IN PUBLISHED TRIALS
Primary Author: REBECCA B SILVA Columbia University Mailman School of Public Health
Co-Author(s): SHING LEE Columbia University Mailman School of Public Health

In this article, we develop methods for sample size and power calculations in four-level intervention studies when intervention assignment is carried out at any level, with a particular focus on cluster-randomized trials. Cluster-randomized trials involving four levels are becoming popular in healthcare research, where the effects are measured, for example, from evaluations (level 1) within participants (level 2) in divisions (level 3) that are nested in clusters (level 4). In such multi-level cluster-randomized trials, we consider three types of intraclass correlations between different evaluations to account for such clustering: that of the same participant, that of different participants from the same division, and that of different participants from different divisions in the same cluster. Assuming arbitrary link and variance functions, with the proposed correlation structure as the true correlation structure, closed-form sample size formulas for randomization carried out at any level (including individually randomized trials within a four-level clustered structure) are derived based on the generalized estimating equations approach using the model-based variance and using the sandwich variance with an independence working correlation matrix. We demonstrate that empirical power corresponds well with that predicted by the proposed method for as few as eight clusters, when data are analyzed using the matrix-adjusted estimating equations for the correlation parameters with a bias-corrected sandwich variance estimator, under both balanced and unbalanced designs.
CHRISTINA YAP  
Clinical Trials Biostatistics, Institute of Cancer Research

RICHARD CARVAJAL  
Columbia University Herbert Irving Comprehensive Cancer Center

Purpose: Multiple simulation studies have shown that novel designs such as the continual reassessment method and the Bayesian optimal interval design outperform the 3 + 3 design in terms of identifying the true MTD more often while using less patients. However, the 3 + 3 continues to be the most used method for oncology trials. One often cited reason by clinical investigators is that simulation studies are not sufficiently convincing to warrant the additional design complexity and that it is not clear that these novel designs would have recommended different doses or used less patients in the context of real dose-finding trials.

Methods: We randomly sample 60 published dose-finding trials to obtain 22 that used the 3 + 3 design, identified an MTD, published their toxicity data, and had more than two dose levels. We compared the published MTD with the estimated MTD using the continual reassessment method, and Bayesian optimal interval using target toxicity rates of 0.25 and 0.30 based on the published data. Moreover, we compared the estimated MTD, patient allocation, and sample size assuming the continual reassessment method and Bayesian optimal interval had been used instead.

Results: Using the observed published data, the continual reassessment method and Bayesian optimal interval, using a target rate of 0.25, matched the published MTD in 12 trials (55%) and 14 trials (64%), and were higher in nine trials (41%) and in eight trials (36%), respectively. Using the continual reassessment method or accelerated Bayesian optimal interval with a target of 0.25 to assign patients, about nine trials (41%) would have recommended one dose level higher than the recommended MTD in the published paper the majority of the time. For these trials, these designs chose dose levels with observed DLT rates between 0.17 and 0.40. In six out of eight of these trials, the 3 + 3 design chose the dose level below with a DLT rate of 0.

Conclusion: With a target rate of 25% and 30%, model-based designs chose dose levels with observed and estimated DLT rates closer to the target rate compared to the 3 + 3 design, choosing a dose level above the recommended MTD in about 40% of the trials. They also had improved patient allocation to the MTD with faster dose escalation and smaller sample sizes without expansion cohorts. This study comparing the 3 + 3 design to the continual reassessment method and Bayesian optimal interval using published dose-finding studies confirm the advantages of novel designs demonstrated by simulation studies further justifying the use of novel methods.

ADVERSE EVENT BURDEN SCORE—A VERSATILE SUMMARY MEASURE FOR CANCER CLINICAL TRIALS

Primary Author: JENNIFER G LE-RADEMACHER  
Mayo Clinic

Co-Author(s): SHAUNA HILLMAN  
Mayo Clinic  
ELIZABETH STORRICK  
Mayo Clinic  
MICHELLE R MAHONEY  
Mayo Clinic  
PETER F THALL  
MD Anderson Cancer Center  
AMINAH JATOI  
Mayo Clinic  
SUMITHRA J MANDREKAR  
Mayo Clinic

Background: In cancer clinical trials, adverse event data are typically collected after each treatment cycle, using the Common Terminology Criteria for Adverse Events, which includes 837 adverse event categories. Most adverse events are classified into five severity grades: 1 = mild, 2 = moderate, 3 = severe, 4 = life-threatening, and 5 = death. The large number of potentially reportable adverse events over multiple treatment cycles makes summarizing and analyzing adverse event data challenging. Consequently, adverse events often are summarized descriptively by tabulating the frequency of the maximum (worst) grade of the most common adverse events reported during the course of treatment, or in randomized trials, making comparisons between treatment arms with a focus on grade 3 or worse adverse events with only descriptive data on specific adverse event categories. Here, we introduce a single quantitative summary measure, called the adverse event burden score, which incorporates both the frequency and the severity of multiple adverse events over time. This can be applied across trials to facilitate adverse event profile comparisons, and/or serve as a benchmark for new trials during the development of new trials.

Method: The adverse event burden by treatment cycle is a weighted sum of all grades and adverse events that the patient experienced per cycle where the weights are clearly defined a priori. The overall adverse event burden score is the total adverse event burden the
patient experienced across all treatment cycles. Adverse event data from two completed multi-center randomized double-blind placebo-controlled trials conducted by the Alliance for Clinical Trials in Oncology, with different adverse event profiles (NCCTG 97-24-51: 176 patients and A091105: 83 patients), are utilized for illustration. The adverse event burden scores were computed with weight equaling the adverse event grade for grade 1–4 events and a weight of 10 for grade 5 events. We assigned 10 as the weight for grade 5 events to reflect the high burden of death due to the adverse event (counts as twice its severity), while assigning a weight that would not overshadow the burden of lower grade events.

Results: Results of the adverse event burden score analyses corroborated the trials’ primary results. In 97-24-51, the overall adverse event burden for patients on the treatment arm was 2.2 points higher than that on the placebo arm (median adverse event burden score: 5 vs 2.8, Wilcoxon’s p-value < 0.0001). A higher adverse event burden was observed in patients who went off treatment early due to an adverse event (Figure 1(b)). Similarly, in A091105, the overall adverse event burden was 1.6 points higher on the treatment arm (median adverse event burden score: 3.6 vs 2.0, Wilcoxon’s p-value = 0.0042). On the placebo arms, the adverse event burden in 97-24-51 remained constant over time, and, in A091105, increased in later cycles, likely reflective of an increase in disease morbidity.

Conclusion: The adverse event burden score may be used for statistical comparisons analogous to other quantitative endpoints in clinical trials. It quantifies the magnitude of the adverse event burden that patients experience during their cancer treatment and should be considered a new and novel safety endpoint in cancer clinical trials.

CP-24
MEASURING INNOVATION-SPECIFIC OUTCOMES IN STUDIES EVALUATING NEW SURGICAL PROCEDURES AND DEVICES: APPRAISAL AND RECOMMENDATION OF OUTCOME MEASUREMENT INSTRUMENTS USING MODIFIED COSMIN METHODOLOGY

Primary Author:
CHRISTIN HOFFMANN
University of Bristol

Co-Author(s):
ANGUS MCNAIR
North Bristol NHS Trust

RHIANNON MACEFIELD
National Institute for Health Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research, University of Bristol

JANE M BLAZEBY
Bristol Royal Infirmary, University Hospitals Bristol and Weston NHS Foundation Trust

KERRY AVERY
National Institute for Health Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research, University of Bristol

SHELLEY POTTER
Bristol Breast Care Centre, North Bristol NHS Trust

Background: Innovations in surgical care are important to improve patient outcomes and advance healthcare. Innovations in invasive procedures and device, however, lack transparent evaluation causing adoption of surgical interventions to not exclusively be based on evidence. The subjective experience and perception of surgeons performing new invasive procedures play a vital role in the refinement and adoption of surgical innovation. A core outcome set for early phase surgical research was recently developed (the COHESIVE study) and identified surgeon experience as a key outcome domain to be measured and reported in all studies of surgical innovation. Robust measurement instruments to assess surgeons’ experience with innovation, however, are not currently defined. The aim of this study is to recommend an instrument for the measurement of surgeon experience of innovative invasive procedures/devices and suggest a measure for this core outcome domain.

Methods: A modified COSMIN methodology was used in three phases to inform recommendation of a suitable outcome measurement instrument. First, existing outcome measurement instruments were identified through targeted and systematic reviews of early phase studies of surgical procedures/devices. Included were self-report measures of surgeon experience. Excluded were measures exclusively assessing surgical competency. Second, appraisal of content validity determined quality and suitability of existing instruments for surgical innovation. Two reviewers independently assessed relevance, comprehensiveness, and comprehensibility of eligible outcome measurement instruments. Third, qualitative methods were used to (in-depth interviews and an interdisciplinary workshop) further explored suitability for surgical innovation for instruments deemed sufficient quality. Recruitment of participants ensured views of a broad range of specialties and geographical locations were represented. Transcripts of semi-structured interviews were analyzed using methods of thematic analysis.
Results: Some 18 outcome measurement instruments were identified from 233 studies across a wide range of specialties. Fifteen outcome measurement instruments met eligibility criteria and underwent COSMIN content validity assessment. Only two instruments met criteria for sufficient quality: The Surgical Task Load Index and The Imperial Stress Assessment Tool. The majority of instruments were graded insufficient (N = 10), indeterminate (N = 2), or inconsistent content validity (N = 2). Analysis of qualitative data from in-depth interviews with 20 surgeons (33% female; 12 specialties, 3 continents) and one interdisciplinary workshop with 10 participants (including surgeons, methodologists, researchers, and editors) revealed that the Surgical Task Load Index is the preferred outcome measurement instrument. Qualitative interviews confirmed that the six dimensions represented in the Surgical Task Load Index (mental demands, physical demands, task complexity, temporal demands, situational stress, and distractions) are relevant, comprehensible and comprehensive to assess surgeons’ experience with surgical innovation.

Conclusion: A single instrument (Surgical Task Load Index) was found to be a valid and suitable to measure surgeon experience in innovative procedures and devices. Uptake and use of this valid measurement instrument are recommended in future studies of surgical innovation. It suggests a valid measure for a key outcome domain identified in the COHESIVE COS. The use of standardized outcome measurement in studies of surgical innovation is important to promote transparent evaluation and safe introduction of new surgical procedures/devices.

CP-25
MULTI-AGENT AGREEMENT COEFFICIENTS: IMPLICATIONS FOR CLINICAL TRIALS OF DIAGNOSTIC TESTING

Primary Author:
MENELAOS KONSTANTINIDIS
University of Toronto

Co-Author(s):
LISA W LE
Princess Margaret Cancer Centre, University Health Network
XIN GAO
York University

Background: While no single policy is able to bring an end to the present COVID-19 pandemic, diagnostic testing plays an important role in the development and efficacy evaluation of public policy. In this regard, throughout the pandemic, the Food and Drug Administration has approved over 200 diagnostic tests for COVID-19. The natural question then is, which among the many tests should be used, and in what contexts or populations?

Answering this requires expedient and complex clinical trials. In evaluating diagnostic test accuracy, key measures such as the sensitivity and specificity of a given test against baseline are often used. However, having several tests with tolerably comparable results is desired. A natural evaluation of the consistency of various substitutable (or competing tests), is how they all perform relative to a baseline, naturally quantified through the inter-rater agreement (IRA). However, a difficulty arises from the multitude of IRA statistics, not well studied in the multiple-raters context.

Objectives: To evaluate the use of several key IRA statistics in the context of multiple raters (testing methods) with binary outcomes.

Methods: We simulated the responses of several raters from a multi-variate normal distribution with a variable compound correlation matrix (spanning 0 to 1, with a step of 0.001). Responses were then dichotomized at a threshold (0 and 0.5, separately). The two distinct cut-offs provided approximately symmetric and asymmetric binary ratings, respectively. Simulations were done for 2, 3, 4, and 5 raters with 20, 50, 300, and 500 observations. For each number of rater/observations combinations, we estimated the expected value of each IRA statistic (Fleiss’ Kappa, Light’s Kappa, Conger’s Kappa, and Gwet’s AC1) for each correlation matrix, and the estimated variance of each statistic. Finally, we fitted linear regressions between each pair of statistics.

Results: In the symmetric case, the estimated mean values (over the set of latent correlations) of all four statistics are asymptotically equal. In contrast, in the asymmetric case, only the mean values of the Kappa statistics are asymptotically equal as Gwet’s AC1 provides a uniformly higher estimate of IRA. The variance of all four statistics decreases monotonically over each scenario (number of raters and observations). In the symmetric case, Fleiss’ Kappa yielded a higher estimated variance than the other three statistics. Of the remaining three statistics, no difference in the estimated variances was appreciated; however, in the asymmetric case, Gwet’s AC1 yielded a lower estimated variance than the three Kappa statistics for each scenario. Finally, from the linear regressions between the correlations, we observe a strong bijection between each pair of statistics (p < 10^{-15}); however, the relationships between the Kappa statistics and AC1 had neither a slope of 1 nor an intercept of 0.

Discussion: Since population-level prevalence of a set of outcomes may not be known a priori, and in the symmetric case, the four statistics provide similar estimated variances, to address the unknown prevalence, Gwet’s AC1 statistic should be used. Moreover, for meaningful head-to-head comparisons between IRA
measures (especially with AC1), linear transformations between the measures must be accounted for.

**CP-26**

**ADVERSE EVENT RECORDING FAILS TO REFLECT POTENTIAL HAZARDS: A REVIEW OF TRIAL PROTOCOLS OF BEHAVIORAL, LIFESTYLE, AND PSYCHOLOGICAL THERAPY INTERVENTIONS**

**Primary Author:**
DIANA PAPAIOANNOU  
*University of Sheffield*

**Co-Authors:**
CINDY COOPER  
*University of Sheffield*
CARA MOONEY  
*University of Sheffield*
RACHEL GLOVER  
*University of Sheffield*
ELIZABETH COATES  
*University of Sheffield*

**Background:** Clinical trials evaluate beneficial effects of interventions alongside harmful effects in order for risk–benefit analysis. Harms or unintended consequences are possible from behavioral change interventions; however, they are not adequately considered in clinical trials. This could leave harms, typically termed adverse events, unidentified. There are problems in how to identify and record harms in behavioral change intervention trials such as (1) considering what harms might be expected from an intervention, (2) determining if an intervention caused a harm, (3) high-frequency and unrelated harms in some populations (e.g. elderly) causing recording burden.

**Aim:** The aim of this study was to explore whether and how potential harms are assessed in trials of behavioral, lifestyle, and psychological therapy interventions.

**Methods:** This study was a review of protocols from the National Institute of Health Research Health Technology Assessment and Public Health Research programs. Protocols were included if the study was a randomized controlled trial and the intervention intended to change lifestyle or behavior or improve psychological outcomes. Data were extracted on the details given on adverse event recording including definitions and assessments planned for serious and non-serious adverse events. (According to the ICH-GCP definition, serious denotes an event resulting in outcomes such as death, hospitalization, life-threatening episode, or persistent and significant disability).
The PRAgmatic EValuation of evENts And Benefits of Lipid lowering in oldEr adults (PREVENTABLE) Trial is the largest randomized trial ever conducted on adults age 75 years and older that will test the effectiveness of statins for the prevention of new dementia or persisting disability. PREVENTABLE is a multi-center, randomized, parallel group trial, which will recruit 20,000 community-dwelling adults without history of cardiovascular disease, dementia, or significant disability. Participants will be randomly assigned to atorvastatin 40 mg daily or a matching placebo and will have a maximum follow-up of 5 years. Wake Forest School of Medicine, Division of Public Health Sciences serves as the Data Coordinating Center for PREVENTABLE and uses a web-based data management system designed for data collection that includes the tracking of blood samples for lipid panel testing and biorepository storage. The first 2000 participants randomly selected for lipid panel analysis also complete an additional blood draw at 3 months post-randomization to test low-density lipoprotein cholesterol lowering from baseline in response to the intervention. Information from each blood sample collected is entered into the Lab Blood Collection Form in the study database. Each printed form has an associated barcode. Scanning or entering the barcode from the printed Blood Collection Form directs the user to the Lab Blood Collection page that is associated with the specific participant and the visit. Scanning a sample’s barcode into the Blood Collection Form in the PREVENTABLE website then links each sample tube to the participant and the user manually completes additional information such as sample type, size, and box ID. Automatic checks ensure that accurate information has been entered into the form, including the appropriate barcode ID structure. The Lab Tracking Tool is a dynamic system that compiles information from multiple participant’s Blood Collection Forms and other participant information to generate a shipping manifest that automatically creates an electronic manifest (transferred to the receiving laboratory) and a paper version (included in the shipping box). The Lab Tracking Tool provides a summary of the number of samples within a box, the dates of sample collection, the box shipment date, and tracking number. A shipping manifest is created for each individual box with information including the barcode ID and participant information of each sample. The creation of a shipping manifest automatically triggers an email alert to the lab processors that samples are being shipped and includes an Excel file of the shipping manifest. Web-based lab blood data collection allows for an efficient and more accurate method of lab tracking and processing. It is important for accurate and efficient data exchange within and among various system frameworks by providing an easier workflow. This presentation will elaborate on the form and system design for data entry describing the features that make this an efficient system for lab tracking as well as advantages and limitations of the tool.

**CP-28**

**DETERMINING THE OPTIMAL PERCENT OF RESEARCH RECORDS TO REVIEW: REDUCING THE VARIABILITY IN AUDITS WITH HUMAN RESEARCH PARTICIPANTS**

**Primary Author:**
MATTHEW J GOODEN
*National Institutes of Health*

**Co-Author(s):**
GINA NORATO
*SANDRA MARTIN*
*National Institutes of Health*

Risk-based monitoring is a quality-driven, monitoring initiative that uses risk factors to determine the extent of review when not performing a complete audit, thereby minimizing financial cost and safeguarding the welfare of human participants. At National Institutes of Health and in the National Institute of Neurological Disorders and Stroke, auditors use the following standard operating procedure to select the percent of participant records to review during an audit: If there are 10 or fewer participants enrolled at the time of audit, 100% of research records are reviewed; if there are 11–20 participants enrolled, 50% of records are reviewed; and if there are greater than 20 participants enrolled, 10%–50% of records are reviewed.

Applying these current guidelines to clinical trials with greater than 20 participants has a risk of either overestimating the incidence of non-compliance events and potentially wasting study resources or underestimating and missing events which may result in unnecessary harm to participants. This study advances the field by evaluating the variability found while auditing according to current practices in National Institute of Neurological Disorders and Stroke. To date, the variability during an audit with human research participants is rarely documented in the literature. As a result, the aim of this study was to gain additional understanding about the optimal percentage of participant records needed to review to confidently identify events of non-compliance during an audit.

An exhaustive audit of all data collected on a single natural history protocol in National Institute of Neurological Disorders and Stroke was conducted.
Events identified during the audit were categorized by the nature of the deviation (e.g., procedural, eligibility, consent, or regulatory) and by the severity (e.g., major or minor). Simulations were performed, as follows, to understand how varying the extent of the audit affects the accuracy of identifying non-compliance: (1) an initial 5% of total participants were sampled and the number of deviations per participant, for each type of deviation, was calculated; (2) resampling of participants was performed 99 more times at this size of 5%, for a total of 100 simulated samples; and (3) the above procedures were repeated for 10% of participants, 15%, 20%, and so on, until 100%, where 100% captures the true underlying deviation event rate at the time of the audit.

This protocol enrolled 55 participants and a total of 138 events of non-compliance occurred, 96 (70%) of which were procedural non-compliance and 33 (24%) of which were major non-compliance. Results indicated that the current guidelines of reviewing 10% of records for auditing large clinical trials may result in excess variability from the true deviation rate. For both the nature and the severity of each deviation, reviewing at least 45% of records may provide additional accuracy, while reducing the burden of reviewing a larger sample.

This study highlights the importance of understanding the sampling variability that may result from selecting a standard percent of records to review during an audit and its effects on potentially impacting participant safety or data integrity. Future studies are needed to determine if the results are consistent with larger studies and with other trial types.

**CP-29**

**STUDY WITHIN A TRIAL COMPARING ELECTRONIC VersUS PAPER PATIENT-REPORTED OUTCOME COLLECTION**

**Primary Author:**
LARA PHILIPPS
Institute of Cancer Research

**Co-Author(s):**
JO HAVILAND
Institute of Cancer Research
ALEXA GILLMAN
Institute of Cancer Research
EMMA HALL
Institute of Cancer Research
REBECCA LEWIS
Institute of Cancer Research

**Introduction:** Within oncology trials patient perspective and survivorship effects are important considerations in evaluating new treatments. Patient-reported outcomes are collected using validated questionnaires that measure the impact of treatment and health conditions on quality of life. The majority of cancer clinical trials currently utilize paper questionnaires, a time-consuming data collection process. Streamlining via electronically collected patient-reported outcome may improve convenience, efficiency and lead to more complete data. Although electronically collected patient-reported outcome has been widely studied in the general clinical setting, there is little published literature demonstrating that they are as effective as paper patient-reported outcome within clinical trials with the associated additional research ethics, governance, and regulatory requirements. Our study-within-a-trial will test implementation of electronically collected patient-reported outcome collection.

**Methodology:** A survey has been collaboratively designed with the charity Independent Cancer Patient’s Voice assessing attitudes regarding health questionnaires among members of the general public. Survey participants may opt to participate in a focus group supporting development of the study-within-a-trial. The study-within-a-trial will compare electronic versus paper patient-reported outcome collection within randomized controlled cancer trials. The design is a partially randomized patient preference trial to ensure that “real-world” data are collected and to prevent participants being excluded due to lack of Internet access. Eligible patients will be participants of a host trial within which the study-within-a-trial is embedded. They will either be randomly allocated to electronic versus paper patient-reported outcome, or will receive their preferred modality if they are unsuitable to be randomized. All participants will complete a paper baseline questionnaire and will then receive questionnaires either electronically via a secure online database or on paper at appropriate timepoints for the host trial up to 12 months. All participants will receive a paper questionnaire assessing satisfaction of the modality of patient-reported outcome completion at 14 months post-enrollment. The primary endpoint is patient compliance (% questionnaires returned out of those expected) at the host trial’s first post-intervention patient-reported outcome timepoint, and the aim is to test for non-inferiority of electronic versus paper patient-reported outcome. It is unknown what proportion of participants will accept randomization versus having a preference. Therefore, the sample size required for the randomized cohort has been estimated, and the numbers entering the preference versus sample size required for the randomized cohort have been estimated, and the numbers entering the preference versus randomized parts of the study-within-a-trial will be monitored. Currently, compliance with paper questionnaires at the first post-intervention timepoint in Institute of Cancer Research Clinical Trials Statistical Unit trials is around 90%. With this assumption, 244
patients (randomized 1:1) are required to exclude compliance rates with electronically collected patient-reported outcome <80% (i.e., 10% non-inferiority margin), with 80% power and one-sided alpha = 0.05. Secondary endpoints include comparison of distributions of domain scores and item responses between electronic and paper questionnaires at key timepoints and in relation to baseline paper questionnaires, compliance at further timepoints within the host trial, data completeness (% of completed items within the questionnaire), and patient satisfaction with both modalities. Discussion: Electronically collected patient-reported outcome could significantly improve the patient experience within clinical trials. Our work aims to show evidence of non-inferiority of electronically collected patient-reported outcome compliance as part of a rollout within clinical trials.

**CP-30**

**MAKING A DISTINCTION BETWEEN DATA CLEANING AND CENTRAL MONITORING**

**Primary Author:**
SHARON LOVE  
MRC Clinical Trials Unit at UCL

**Co-Author(s):**
VICTORIA YORKE-EDWARDS  
MRC Clinical Trials Unit at UCL
CARLOS DIAZ-MONTANA  
MRC Clinical Trials Unit at UCL
MACEY L MURRAY  
MRC Clinical Trials Unit at UCL
LINDSEY MASTERS  
MRC Clinical Trials Unit at UCL
NICOLA JOFFE  
MRC Clinical Trials Unit at UCL
MATTHEW R SYDES  
MRC Clinical Trials Unit at UCL

The terms “data cleaning” and “central monitoring” have become intertwined to the potential detriment of clinical trial conduct. They are practically and conceptually different. What is data cleaning, what is central monitoring, and why does the difference matter? As an example, a data cleaning activity might be sending out a list of queries for site teams to resolve, whereas a related central monitoring activity might be looking at query resolution rates across different sites and escalating if a certain percentage of queries have remained open for 6 months or more. Without a clear understanding of what is meant by data cleaning and by central monitoring, trial teams and site staff may spend time and effort wastefully or inappropriately. If these activities are not separated, they can each occur at the wrong time: data cleaning too rarely and central monitoring too frequently. The quality of the trial will suffer if the differences between data cleaning and central monitoring are not well appreciated. If they are not separated then either or both could be done inadequately. By considering them as one, it can feel like enough is being done. If they are not done separately, then a risk for a trial may not be adequately covered. If the research community cannot be clear on language, it is difficult to discuss best practice or, importantly, define high-quality methodology projects to determine evidence-based improvements to approaches across trials. In conclusion, it is important to correctly define data cleaning and central monitoring in order to communicate the conduct of a trial, to ensure adequate risks mitigation and to ensure that the data are appropriately corrected.

**CP-31**

**INTERVENTION FIDELITY AND TREATMENT EFFECT ESTIMATES IN CLINICAL TRIALS OF COMPLEX INTERVENTIONS: A META-EPIDEMIIOLOGICAL STUDY**

**Primary Author:**
ARSENIO PAEZ  
University of Oxford

**Co-Author(s):**
DANIEL HIND  
University of Sheffield
DAVID NUNAN  
University of Oxford
PETER MCCULLOCH  
University of Oxford
DAVID BEARD  
University of Oxford

The focus of pragmatic trials is typically intervention effectiveness, or whether or not interventions attain their desired outcomes. Less attention is focused on understanding how or why interventions succeed, or fail to attain, their target outcomes. This may be of particular importance in explanatory trials, when there are conflicting findings of interventions’ effectiveness, and in evaluations of complex interventions, which are multi-faceted and often adapted or tailored to individuals.

The correspondence between the intervention that was planned and the intervention that was actually delivered is the intervention’s fidelity. Careful consideration of fidelity helps to explain study findings, highlight possible associations between study factors and outcomes, and develop and refine interventions for
future testing. High levels of intervention fidelity reduce random and unintended variability, limit potential confounding, and decrease the likelihood of type 1, 2, and 3 errors. Even modest degradation of intervention fidelity may lower statistical power in primary research, making it necessary to increase trials’ sample size considerably in order to compensate for losses in fidelity. As a result, many studies of complex intervention are likely to be underpowered when treatment fidelity is low or unmeasured, and treatment effect sizes may be biased. These issues are magnified when the results of studies with poor fidelity are pooled in meta-analyses. The magnitude and direction of fidelity’s influence on treatment effect sizes in clinical trials had not previously been determined, however. This meta-epidemiological study quantifies the size and direction of fidelity’s influence on treatment effect estimates in complex intervention randomized clinical trials.

A systematic search was undertaken for meta-analyses of randomized clinical trials in physiotherapy, speech therapy, and exercise or physical activity promotion with at least two treatment arms and reporting quantitative data for continuous outcomes, published between 2015 and 2020. Nineteen meta-analyses were included in the meta-epidemiological study, representing over 200 randomized clinical trials across three complex interventions in healthcare. Fidelity within trials was quantified using the CONSIDER Fidelity framework previously developed as part of this program of research, identifying measures or processes associated with intervention fidelity in randomized clinical trials. Meta-epidemiological analysis followed a two-level, meta-meta-analytic design, in which treatment effect estimates are compared between trials with and without fidelity measures or monitoring within each meta-analysis, further examining the influence of fidelity, risk of bias, sample size and other characteristics through meta-regression. Smaller treatment effect sizes and more narrow confidence intervals are found across randomized clinical trials when intervention fidelity is present than in trials without fidelity monitoring or intervention fidelity, even when adjusting for risk of bias or sample size.

This is the first meta-epidemiological study to quantify the influence of fidelity on treatment effect estimates in clinical trials. Determining the influence of fidelity on treatment effect estimates in complex intervention clinical trials can significantly influence the interpretation and appropriateness to change practice of clinical studies of complex interventions and clinical practice guidelines based on their systematic review. With the results of this meta-epidemiological study, we make evidence-based recommendations for the conduct, reporting, and quality assessment of complex intervention randomized clinical trials.

**CP-32**

**IMPROVING THE ANALYSIS OF ADVERSE EVENT DATA IN RANDOMIZED CONTROLLED TRIALS**

**Primary Author:**
VICTORIA R CORNELIUS
Imperial College London

**Co-Author(s):**
RACHEL PHILLIPS
Imperial College London

**Introduction:** Obtaining early information on treatment harm is crucial but difficult with the relatively small sample size that an average trial affords compared to pharmacovigilance studies. Good analysis practice for efficacy (primary and secondary) outcomes are now well-established. In contrast, there has been minimal progress for the analysis of adverse events. The current suboptimal analysis of adverse events in randomized clinical trials means we are missing a valuable opportunity to detect harmful effects early.

**Methods:** In this research, we introduce the complexity of adverse event data before providing a framework of the attributes for the analysis that need to be considered and addressed to improve practice. These start with how we value randomized clinical trial data in the assessment of treatment harm, we then discuss the barriers to analysis due to complexity in data and restricted sample size. We examine the research question for adverse events in trials and discuss analysis solutions using a new inference framework. We end by considering the controversies around methods used to select adverse events for reporting in publications.

**Results:** Data collection for adverse event data are designed to be all encompassing as we are unable to pre-specify all events of interest to be collected at the start of a trial. Multiple collection methods used to gather an unspecified number of outcomes often produce vast amounts of multi-faceted data making the analysis challenging. Current analysis practice predominantly involves the use of frequentist hypothesis testing or no testing at all. With restricted sample size resulting in low power and multiple outcomes, both traditional hypothesis testing and not testing the data are unsuitable. We propose reframing the research question to one for detecting signals of adverse reactions in order to provide a different context for interpretation. In the absence of guidance on suitable analysis approaches, we encourage uptake of established good statistical practice, such as adjusted modeling and avoiding dichotomization, with use of distributional methods when dichotomization is necessary. To improve statistical practice, the most appropriate methods need to be identified and evaluated. We highlight how a Bayesian
approach offers a natural framework for adverse event analysis due to its ability to incorporate prior information and enable cumulative assessment of harm as a treatment progresses through the developmental pathway. A final but crucial challenge lays in presenting a pertinent summary of harm in the main trial publication. The availability of the online supplement means that a popular choice now is to report all recorded adverse events in supplementary materials, but without a summary in the main results paper inference is impeded.

Conclusion: The analysis of adverse events in contemporary clinical trials is outdated and needs attention if we are to progress and realize the value of randomized clinical trial data and enable earlier signaling of harm. Community appraisal of methods, practical implementation and guidance are needed to support change. In the interim adopting well-accepted principled analysis approaches with appropriate inference would afford some immediate benefit.

CP-33
STATISTICAL METHODS FOR HANDLING NON-ADHERENCE TO INTERVENTIONS IN RANDOMIZED NON-INFERIORITY TRIALS: A SYSTEMATIC REVIEW

Primary Author:
MATTHEW DODD
London School of Hygiene & Tropical Medicine

Co-Author(s):
KATHERINE FIELDING
London School of Hygiene & Tropical Medicine

JAMES CARPENTER
London School of Hygiene & Tropical Medicine

JENNIFER THOMPSON
London School of Hygiene & Tropical Medicine

DIANA ELBOURNE
London School of Hygiene & Tropical Medicine

Background: Non-adherence to interventions (or non-compliance) is common in trials. In non-inferiority trials, it may increase the probability of falsely claiming non-inferiority in standard intention-to-treat analyses. Per-protocol analysis, which typically excludes individuals not receiving their allocated intervention as intended, is often used as an alternative. However, if non-adherent individuals differ systematically from those who are adherent, this approach can potentially bias the trial results in either direction. This systematic review aimed to identify statistical methods that attempt to account for the impact of non-adherence to interventions (thereby estimating the causal effects of interventions) in randomized non-inferiority trials.

Methods: A systematic review (PROSPERO registration CRD42020177458) was conducted by searching the Ovid Medline® database (31 December 2020) for terms related to adherence, statistical methods for handling non-adherence, and non-inferiority trials. Two authors independently reviewed the titles and abstracts of each paper identified. The following were excluded: non-randomized comparisons, primary analyses not for non-inferiority, papers not published in English, and cost-effectiveness analyses. Full-text reviews of those remaining and containing keywords (including “adherence” or “compliance”) were performed independently by two authors to identify: (1) randomized trials with a primary analysis for non-inferiority that applied (or planned to apply) statistical methods to account for the impact of non-adherence to interventions, and (2) methodology papers that described such statistical methods and included a non-inferiority trial application. Systematic reviews and meta-analyses were searched for eligible non-inferiority trials. Discrepancies between reviewer pairs were discussed with a third author in order to reach a consensus. Statistical analysis plans were requested for eligible publications.

Results: Results from 3235 publications identified found 24 papers meeting the eligibility criteria for inclusion (4 protocols, 13 results papers, and 7 methodology papers). Relevant statistical methods were reported on 26 occasions. The most common were instrumental variable approaches (n = 9), including observed adherence as a covariate within a regression model (n = 3), modeling adherence as a time-varying covariate in a time-to-event analysis (n = 3). Other methods included rank preserving structural failure time models and G-estimation, inverse-probability-of-treatment weighting, and the tipping point approach. The methods identified in protocols and results papers were more commonly specified in sensitivity analyses (n = 13) than primary analyses (n = 3). Twelve results papers included an alternative analysis of the same outcome; conclusions regarding non-inferiority were in agreement on six occasions and could not be compared on six occasions (different measures of effect or results not provided in full).

Conclusion: While statistical methods that can be used to account for the impact of non-adherence to interventions are available, their use in non-inferiority trials remains infrequent. Since intention-to-treat and per-protocol analyses do not guarantee unbiased conclusions regarding non-inferiority, the methods identified should be considered for use in sensitivity analyses of non-inferiority trials. In particular, those with non-trivial non-adherence should assess the sensitivity of the trial results to different assumptions in order to guard against falsely claiming non-inferiority and accepting a worse intervention.
CP-34

MAKING INTERIM DECISIONS USING CONDITIONAL POWER BASED ON A SURROGATE ENDPOINT

Primary Author:
JODY D CIOLINO
Northwestern University

Co-Author(s):
DANIELA LADNER
Northwestern University
LISA VANWAGNER
Northwestern University

For a given efficacy clinical trial at some interim study fraction, we can define conditional power as the probability of a statistically significant finding at the end the study given the current test statistic, interim fraction, and some assumption on the true intervention effect. Traditionally, an interim futility analysis calling for conditional power will involve calculations under the assumptions of (1) the current trend, (2) the anticipated alternative trend, and (3) the null hypothesis effect assumption at some pre-specified interim fraction. This interim fraction (t) roughly translates to the number of participants for whom we have an observed outcome at a given time (n) divided by the total expected number (N) of study participants by the end of the study to achieve study goals (i.e. t = n/N = the proportion of the way through the trial at a given time).

Obtaining a reasonable information fraction such as t = 0.20 or larger for long-term outcomes (e.g. long-term survival) may take years, thus making a futility analysis based on conditional power impractical. In a recent trial planning endeavor, we derived an estimate for conditional power on primary long-term outcome given the above assumptions plus an added assumption on correlation between the surrogate and final outcome test statistics. Plots of the derived estimates under varying scenarios illustrate the potential utility (or lack thereof) of computing conditional power based on an interim surrogate endpoint. Simulations allow for estimation of proportion of “incorrect” interim decisions depending on scenario along with estimation of sample size savings when relationships between surrogate and final endpoints are strong.

We present an application for this framework as we develop a proposed study design for the Liver Cirrhosis Network to ascertain if a medication (statin) can slow down the progression cirrhosis (liver scarring) which is associated with high morbidity and mortality. For this proposal, we plan to use a continuous measure that may be obtained between 3 and 6 months post-randomization—hepatic venous pressure gradient, a clinically established measure to assess the degree of cirrhosis—as a surrogate for primary long-term follow-up measure: time-to-decompensation, with decompensation defined as the occurrence of liver-related complications (e.g. gastrointestinal bleeding). If associated, the use of short-term changes in hepatic venous pressure gradient in an early conditional power analysis has potential to save hundreds of study participants from enrollment into a futile trial.

CP-35

INCORPORATING BASELINE COVARIATES TO VALIDATE SURROGATE ENDPOINTS WITH A CONSTANT BIOMARKER UNDER CONTROL ARM

Primary Author:
EMILY ROBERTS
University of Michigan

Co-Author(s):
MICHAEL ELLIOTT
University of Michigan
JEREMY TAYLOR
University of Michigan

A surrogate endpoint S in a clinical trial is an outcome that may be measured earlier or more easily than the true outcome of interest T. In this work, we extend causal inference approaches to validate such a surrogate using potential outcomes. The causal association paradigm assesses the relationship of the treatment effect on the surrogate with the treatment effect on the true endpoint. Using the principal surrogacy criteria, we utilize the joint conditional distribution of the potential outcomes T, given the potential outcomes S. In particular, our setting of interest allows us to assume the surrogate under the placebo, S(0), is zero-valued, and we incorporate baseline covariates in the setting of normally distributed endpoints. We develop methods to incorporate conditional independence and other modeling assumptions and explore their impact on the assessment of surrogacy. We demonstrate our approach via simulation and data that mimics an ongoing study of a muscular dystrophy gene therapy.

CP-36

FAST AND INFORMATIVE PHASE 2B DOSE-RANGING TRIALS USING BAYESIAN UNCERTAINTY-DIRECTED DESIGNS WITH MODEL AVERAGING

Primary Author:
DANIEL SCHWARTZ
University of Chicago

Co-Author(s):
YUAN JI
University of Chicago
This paper discusses new Bayesian adaptive designs for Phase 2b dose-ranging. This stage of clinical development is crucial to an efficient drug pipeline and should, as quickly as possible, produce high-quality evidence about whether a Phase 3 trial should be run and if so what dose of the drug is most likely to succeed. However, traditional designs used in Phase 2b are often slower and provide less conclusive evidence than is possible at this stage because they do not use their limited patient resources efficiently. In contrast, a promising new approach called a Bayesian uncertainty-directed design explicitly randomizes each new cohort of patients more often to a dose that earlier maximizes the added information their data will generate about the optimal dose, given what is already known from earlier patients in the trial. This typically means assigning new patients to doses that have been previously understudied relative to how strongly the data suggest they could be the optimal dose. This work develops new BUD approaches for use in the Phase 2b setting. Critically, we incorporate pharmacological knowledge about the dose–response curve to allow data about one dose’s effectiveness to partially inform the evidence about nearby doses. This is achieved through powerful yet robust Bayesian model averaging of candidate dose–response models. Combining BUD designs with dose–response curve modeling can be computationally daunting; we provide an efficient algorithm using Sequential Monte Carlo techniques to make these designs easy to use and study.

**CP-37**

**THE DESIGN AND ANALYSIS OF SMALL POPULATION AND RARE DISEASE TRIALS: A REVIEW OF RANDOMIZED TRIALS**

**Primary Author:**

GILES PARTINGTON  
Imperial College

**Co-Author(s):**

SUZIE CRO  
Imperial College  
ALEXINA MASON  
London School of Hygiene & Tropical Medicine  
RACHEL PHILLIPS  
Imperial College  
VICTORIA CORNELIUS  
Imperial College

**Introduction:** Trials in rare diseases are needed to ensure effective treatments are found. However, rare disease trials face the challenge of restricted sample size. Sample size can also be restricted when there is interest in a subgroup of special interest. Late phase randomized controlled trials designed in a frequentist framework for hypothesis testing often require large sample sizes to ensure good power to detect the minimum clinical important difference. In small populations and rare diseases, this sample size may be unfeasibly large. A temptation can be to go above the minimum clinical important difference, but this risks the trial being underpowered to detect a treatment effect that would be of interest. Rather than treating a trial as stand-alone evidence, one option is to use a Bayesian approach and incorporate valuable existing information through informative prior distributions. Alternatively, frequentist trials can implement several less commonly applied strategies in their design, such as relaxing the type 1 error rate.

**Methods:** We conducted a targeted review of rare disease or small population trials in the Embase and Medline databases, to identify what design and analysis approaches are being used. Included studies were phase 2–4 randomized clinical trials that reported “rare” disease or “small population” in the title or abstract from 2009 onwards. We evaluated the number of trials using a frequentist or Bayesian approach. For Bayesian trials, we further evaluated the types of prior distributions used, approaches to developing prior distributions and effective sample size of this additional information. For frequentist trials, we evaluated sample size parameters and other features related to overcoming issues of small sample size.

**Results:** Six thousand one hundred eight articles were screened, and 64 eligible trials were found, of which 4 (6%) were Bayesian trials and 60 (94%) were frequentist. Only 35/60 frequentist trials stated a planned and actual sample size with mean recruitment 9.6 patients lower than planned. Planned power ranged between 72% and 90%, with 5% powered at <80%. Fifty-two (87%) frequentist trials used 5% type 1 error, five (8%) used 10% type 1 error, one (2%) used 1% type 1 error, and two (3%) did not report. Of the Bayesian trials, two used Bayesian methods only in their analysis, while the other two used them for both design and analysis. Three used informed priors; one used only historical information to inform the prior, another used only expert group consensus, and the final incorporated both historical information and expert group consensus. Three Bayesian trials reported how much larger a sample size was needed to be frequentist (317%-2500%), and one reported effective sample size gain of 85% and 240% for respective trial arms.

**Discussion:** A few rare or small population trials use a Bayesian approach for their design or analysis; those that do are able to achieve improvements to their effective sample size. Most small population frequentist trials do not achieve their planned sample sizes and are therefore likely underpowered. A few small population frequentist trials utilize multiple methods suggested for
appropriate evaluation of treatments in guidelines such as in Parmar 2016 that should be more widely adopted.

**CP-38**

**CHALLENGES AND PROPOSED SOLUTIONS IN MAKING CLINICAL RESEARCH ON COVID-19 ETHICAL. A STATUS QUO ANALYSIS ACROSS GERMAN RESEARCH ETHICS COMMITTEES**

**Primary Author:**
ALICE FAUST  
Charité– Universitätsmedizin Berlin

**Co-Author(s):**
ANNA SIERAWSKA  
Charité– Universitätsmedizin Berlin

KATHARINA KRÜGER  
Association of Medical Ethics Committees in Germany

ANNE WISGALLA  
Association of Medical Ethics Committees in Germany

JOERG HASFORD  
Association of Medical Ethics Committees in Germany;  
Institute for Medical Information Processing, Biometry, and Epidemiology, LMU München, Munich

DANIEL STRECH  
Charité– Universitätsmedizin Berlin

**Background:** In the course of the COVID-19 pandemic, the biomedical research community’s attempt to focus the attention on fighting COVID-19, led to several challenges within the field of research ethics. However, we know little about the practical relevance of these challenges for Research Ethics Committees.

**Methods:** We conducted a qualitative survey across all 52 German Research Ethics Committees on the challenges and potential solutions with reviewing proposals for COVID-19 studies. We de-identified the answers and applied thematic text analysis for the extraction and synthesis of challenges and potential solutions that we grouped under established principles for clinical research ethics.

**Results:** We received an overall response rate of 42%. The 22 responding Research Ethics Committees reported that they had assessed a total of 441 study proposals on COVID-19 until 21 April 2020. For the review of these proposals, the Research Ethics Committees indicated a broad spectrum of challenges regarding (1) social value (e.g. lack of coordination), (2) scientific validity (e.g. provisional study planning), (3) favorable risk–benefit ratio (e.g. difficult benefit assessment), (4) informed consent (e.g. strict isolation measures), (5) independent review (e.g. lack of time), (6) fair selection of trial participants (e.g. inclusion of vulnerable groups), and (7) respect for study participants (e.g. data security). Mentioned solutions ranged from improved local/national coordination, over guidance on modified consent procedures, to priority setting across clinical studies.

**Conclusion:** Research Ethics Committees are facing a broad spectrum of pressing challenges in reviewing COVID-19 studies. Some challenges for consent procedures are well-known from research in intensive care settings but are further aggravated by infection measures. Other challenges such as reviewing several clinical studies at the same time that potentially compete for the recruitment of in-house COVID-19 patients are unique to the current situation. For some of the challenges, the proposed solutions in our survey could relatively easy be translated into practice. Others need further conceptual and empirical research. Our findings together with the increasing body of literature on COVID-19 research ethics, and further stakeholder engagement should inform the development of hands-on guidance for researchers, funders, Research Ethics Committees, and further oversight bodies.

**CP-39**

**REPORTING OF INFORMED CONSENT IN PRAGMATIC TRIALS: RESULTS FROM A SYSTEMATIC LITERATURE SURVEY**

**Primary Author:**
MONICA TALJAARD  
Ottawa Hospital Research Institute

**Co-Author(s):**
JENNIFER ZHANG  
University of Ottawa

STUART G NICOLLS  
Ottawa Hospital Research Institute

KELLY CARROLL  
Ottawa Hospital Research Institute

HAYDEN NIX  
Western University

CORY GOLDSTEIN  
Western University

SPENCER P HEY  
Harvard Medical School

JAMIE C BREHAUT  
Ottawa Hospital Research Institute

PAUL MCLEAN  
Harvard Medical School

CHARLES WEIJER  
Western University
Background: Pragmatic trials aim to evaluate interventions in real-world conditions to support clinical decision-making. Pragmatic trials include a variety of designs and approaches, including comparative effectiveness trials of routinely used medical interventions, public health and health promotion trials, and cluster-randomized trials testing different healthcare policies. Informed consent is an essential requirement for the protection of human research participants. Given that pragmatic trials are normally carried out in usual care settings, some view them as posing only minimal risk to the research participants. This gives rise to differing viewpoints about the need for or type of informed consent, and subsequently to variation in practice. Existing trial reporting guidelines vary in their recommendations about consent.

Objective: The objectives of this study are to describe the reporting of research ethics review and informed consent in pragmatic trials published between 2014 and 2019, and to identify characteristics associated with reporting of and obtaining consent.

Methods: Primary reports of 4337 pragmatic trials were identified using an electronic search filter previously developed by the research team. The subset of N = 1988 with a corresponding registration in ClinicalTrials.gov was identified and used in this review. Five reviewers independently extracted data on reporting of research ethics review and informed consent, including justifications for exemption from review and waivers of consent. Trial characteristics (e.g. country of recruitment, clinical vs public health, and individual vs cluster randomization) were also extracted. Additional trial characteristics were downloaded from ClinicalTrials.gov. Percentages and 95% confidence intervals were used to describe reporting of research ethics review and informed consent. Two separate multi-variable logistic regression analyses examined associations of trial characteristics with (1) not reporting on consent, and (2) not obtaining consent.

Results: A statement about research ethics review was included in the vast majority of trials (1947 (97.9%); 95% confidence interval = 97.3% to 98.6%). Of 22 trials that explicitly stated the study had not been reviewed, just over half (54.5%) provided an explanation. A statement about participant informed consent was included in 1856 trials (93.4%; 95% confidence interval = 92.3% to 94.5%), of which the vast majority (1691 (91.1%); 95% confidence interval = 89.8% to 92.4%) indicated that consent had been obtained, 139 (7.5%; 95% confidence interval = 6.3% to 8.7%) reported consent had not been obtained, and 26 (1.4%; 95% confidence interval = 0.9% to 1.9%) reported consent for one aspect but a waiver for another. Of the 165 trials that reported no consent or waiver of consent, just under half (46.1%) provided a justification. Lower journal impact factor and cluster randomization were significantly associated with not reporting on consent, while later trial start year, cluster randomization, higher income country settings, health services research, and self-identification as pragmatic were significantly associated with not obtaining consent.

Conclusion: We found incomplete reporting of ethical protections in pragmatic trials and variation in ethical practice across trial characteristics. We recommend that reporting guidelines are updated to include explicit items about consent, including justifications for exemptions from research ethics review and waivers of consent.

CP-40

REDUNDANT CLINICAL TRIALS ARE HURTING PATIENTS: AN EXAMPLE OF RANDOMIZED CLINICAL TRIALS CONDUCTED IN MAINLAND CHINA EVALUATING STATINS AMONG PATIENTS WITH CORONARY ARTERY DISEASE

Primary Author:
YUANXI JIA
Johns Hopkins University

Co-Author(s):
JIAJUN WEN
Bloomberg School of Public Health
RIAZ QURESHI
Bloomberg School of Public Health
STEPHAN EHRHARDT
Bloomberg School of Public Health
DAVID D CELENTANO
Bloomberg School of Public Health
XIN WEI
Virginia Commonwealth University
LORI ROSMAN
Johns Hopkins University
YUMENG WEN
Johns Hopkins University
KAREN A ROBINSON
Johns Hopkins University

Objective: To identify redundant clinical trials evaluating statin therapy among patients with coronary artery disease in Mainland China, and to estimate the number of extra major adverse cardiac events experienced by the patients in the control group of those redundant trials due to deprivation of statin therapy.

Design: Cross-sectional study.

Methods: Randomized trials from Mainland China were included if they compared statins with placebo or no treatment among patients with coronary artery disease. Redundant trials were defined as those that were
initiated or continued recruiting after 2008, that is, 1 year after statins were strongly recommended by clinical practice guidelines. Chinese and English bibliographic databases were searched for redundant trials up to December 2019. We also conducted cumulative meta-analyses to establish the timepoints when statins were shown to have a statistically significant effect on coronary artery disease.

**Main Outcome and Measure:** The number of extra major adverse cardiac events that were attributable to the deprivation of statins among patients in the control groups of redundant clinical trials, that is, the number of extra major adverse cardiac events that could have been prevented if patients were given statins.

**Results:** Between 2008 and 2019, 2045 redundant trials were identified, in which 101,486 patients were treated in the control group without statins for 24,638 person-years. Three thousand four hundred seventy (95% confidence interval = 3230–3619) extra major adverse cardiac events were reported, including 559 (95% confidence interval = 506–612) deaths, 973 (95% confidence interval = 897–1052) cases of new or recurrent myocardial infarction, 161 (95% confidence interval = 32–190) cases of stroke, 83 (95% confidence interval = 58–105) cases of revascularization, 398 (95% confidence interval = 352–448) cases of heart failure, and 1197 (95% confidence interval = 1110–1282) cases of recurrent or deteriorated angina pectoris.

**Conclusion:** More than 2000 redundant clinical trials on statins among patients with coronary artery disease were identified from Mainland China. Such trials have been harming patients who have experienced more than 3000 unnecessary major adverse cardiac events, including nearly 600 deaths. The scale of redundancy necessitates urgent reform to protect patients.

---

**CP-41**

THE CHALLENGE OF THE COVID-19 PANDEMIC FACED BY ITALIAN ETHICS COMMITTEES: A RESEARCH SURVEY

**Primary Author:**
DARIO GREGORI
University of Padova

**Co-Author(s):**
MATTEO MARTINATO
University of Padova
GIULIA LORENZONI
University of Padova
DANILA AZZOLINA
University of Piemonte Orientale
ILEANA BALDI
University of Padova

The Italian Ethics Committees protect clinical research participants’ rights and well-being guaranteeing also the scientific integrity of the research. In Italy, one of the first countries severely affected by the COVID-19 pandemic, the Ethics Committees activities have been impacted by the virus spreading and subsequent lockdown from 9 March to 16 May 2020. In mid-June 2020, a cross-sectional survey was conducted in Italy, contacting all the 90 local Ethics Committees. The survey’s purpose was to analyze (1) the activities carried out by the Ethics Committees during the lockdown; (2) the features of the Ethics Committees submitted studies; and (3) the implementation of standard research evaluation protocols during the pandemic. Regions with higher incidence (HI) and lower incidence (LI) of COVID-19 were compared via the chi-square test. Out of 46 Ethics Committees that participated in the study, 258 questionnaires were obtained. A rise in the number of studies submitted to local Ethics Committees was identified in 75% of the HI, in comparison to 53% in LI regions (p < 0.001). The 15% of participants from HI areas documented a reduction in compliance with the standard research application assessment protocols. In the HI regions, a greater percentage of Ethics Committees professionals have moved to smart working in comparison with LI regions (75% vs 59%; p < 0.001). The influence of COVID-19 on Ethics Committees activities in HI territories has been considerable; the Ethics Committees workload increased dramatically especially in these areas. The smart working has been effective in ensuring the activities of Ethics Committees and the subsequent activation of clinical trials that are potentially useful for coping with the pandemic. However, it was stated that an accurate review of the submitted study documents could not be carried out, with the consequence of offering a favorable opinion on studies of not excellent nature. From the point of view of some Ethics Committee participants, the large number of studies performed in the most affected areas, together with the emergency faced during the lockdown, may have exposed Ethics Committees to the risk of reducing the implementation of ethical standards and standard assessment procedures for research applications.

---

**CP-42**

GUIDELINES REGARDING THE ENROLLMENT OF PATIENTS AND EVIDENCE FROM EXPANDED ACCESS PROGRAMS: A CASE STUDY OF CONVALESCENT PLASMA THERAPY FOR THE TREATMENT OF COVID-19

**Primary Author:**
HAYLEY M BELL
New York University School of Medicine
Convalescent plasma is an investigational product, not approved by the Food and Drug Administration for any use, that has been administered for the prevention and treatment of epidemic infections, including Ebola and most recently COVID-19. On 3 April 2020, Food and Drug Administration authorized an Expanded Access Program for treatment use of convalescent plasma in hospitalized COVID-19 patients, with Mayo Clinic as the coordinating site for this nationwide effort. Expanded Access, historically referred to as “compassionate use,” is a regulatory pathway that allows patients to use an investigational product outside of a clinical trial when four conditions are met: the patient has a life-threatening or serious disease, no comparable or satisfactory alternative treatment options are available, clinical trial enrollment is not possible, and non-trial use does not pose a threat to timely clinical development. Food and Drug Administration can authorize Expanded Access through requests for single patients, intermediate-size patient populations, and widespread treatment programs. The convalescent plasma Expanded Access Program was the largest in US history, leading to infusion of more than 94,000 patients. In comparison, the largest prior Expanded Access Programs—for lamivudine (HIV) and gefitinib (cancer)—each provided unapproved drugs to approximately 30,000 patients outside of clinical trials.

Expanded Access Programs of this size are unusual and dwarf most clinical trials. However, Expanded Access Programs are intended for treatment, not research. Therefore, even when they involve sizable patient populations, they lack features of rigorous trial design, including control groups and randomization. Nevertheless, Expanded Access Program data have been considered pivotal by both the Food and Drug Administration and European Medicines Agency, often in the context of rare disease and the collection of safety data, and it was primarily the convalescent plasma Expanded Access Program data that led Food and Drug Administration to issue an emergency use authorization for that product. Although Expanded Access Programs are by regulation not supposed to hinder clinical trials, large Expanded Access Programs can create circular challenges, as when a lack of trial opportunities makes an Expanded Access Program an important pathway for initial access to an investigational product, but then the Expanded Access Program becomes a barrier to launching new trials.

In this session, we will discuss the statistical, ethical, and regulatory issues that arise in the context of Expanded Access Programs, using the convalescent plasma Expanded Access Program as a case study in contrast to clinical trials, with a focus on issues arising in the context of a global pandemic. We will first discuss biostatistical considerations for extracting real-world evidence from Expanded Access Program data, as well as the role of these data in supplementing results from clinical trials. We will also consider the perspective of the patient in deciding whether to enroll in an Expanded Access Program versus clinical trial, including the opportunity for personal benefit or harm, potential for participation to impact the greater good, and the patient understanding of an unproven investigational treatment. Then, we will address the gatekeeping role of clinicians, institutions, and regulators, to ensure that Expanded Access Programs do not interfere with clinical trials. Finally, we will address patient pathways for accessing investigational drugs, especially in the context of a global pandemic, and the capacity for collecting rigorous data via these mechanisms, while prioritizing rigorous trials.

### CP-43

**SITE MONITORING AND SITE MANAGEMENT DURING COVID-19—A CONTRACT RESEARCH ORGANIZATION’S EXPERIENCE**

**Primary Author:**
GEORGINE N PRICE  
Westat, Inc.

**Co-Author(s):**
CHRISTINE ANDERSON  
Westat, Inc.

KATHLEEN TIEFENWERTH  
Westat, Inc.

The challenges brought by the COVID-19 pandemic have forced us all to change our routine practices in drastic ways. The routines of a contract research organization are no exception. These challenges were the catalyst that invigorated Westat to rethink procedures, consider new tools, and strengthen risk assessment protocols that would benefit any kind of clinical trial in the future.

When the COVID-19 pandemic struck, clinical sites restricted access and in-person visits were halted. Before the pandemic, Westat had been conducting
intermittent remote monitoring visits in between on-site visits where adequate electronic technologies existed. Exclusive remote site monitoring was not a routine practice and not conducted at sites without remote access to medical records.

Now, remote access to medical records is required, although some sites did not have the technology to accomplish this task. Westat worked with information technology departments and worked through contract issues. In its monitoring and site management roles, Westat found ways to enhance the infrastructure to ensure that trial sites continued to adhere to protocols, ICH-GCP Guidelines, and regulatory requirements, ultimately to protect the safety of research participants and the integrity of data.

As a starting point, Westat developed a guideline for conducting remote monitoring visits. Table 1 shows a list of the immediate high-level challenges identified and a sample of Westat’s responses. This poster will present overviews of the various tools developed, additional considerations while implementing responses, and lessons learned.

CP-44

ELECTRONIC DATA CAPTURE SYSTEM MODIFICATIONS TO CAPTURE COVID-19-RELATED DATA IN THE NATIONAL DRUG ABUSE TREATMENT CLINICAL TRIAL NETWORK

Primary Author:
MITRA K LEWIS
The Emmes Company

Co-Author(s):
KELSEY KRAUSE
The Emmes Company

NICHOLAS DEVOGEL
The Emmes Company

JACQUELINE KING
The Emmes Company

ASHLEY CASE
The Emmes Company

DAGMAR SALAZAR
The Emmes Company

KATHRYN HEFNER
The Emmes Company

ABIGAIL G MATTHEWS
The Emmes Company

The COVID-19 public health emergency created significant challenges for the safe conduct of clinical trials. The Clinical Trials Network conducts multi-site substance use treatment studies, of which several are in an active recruitment phase. Investigators within the Clinical Trials Network collaborate with the Clinical Coordinating Center and Data and Statistics Center, both at the Emmes Company, to effectively manage these trials and ensure data quality. In March 2020, the US Food and Drug Administration issued guidance for the conduct of clinical trials during the pandemic; notably, the Food and Drug Administration recommended capturing reasons for missing data, protocol deviations, or modified study procedures as related to COVID-19.

The Clinical Coordinating Center and Data and Statistics Center worked with investigators to evaluate and modify study protocols to allow for flexibility in the collection of study assessments (e.g. off-site/home visits and telehealth visits) while ensuring the safety of participants and research staff and maintaining the integrity of trial data. In addition, an interdisciplinary team within the Clinical Coordinating Center and Data and Statistics Center reviewed the Food and Drug Administration guidance, identified changes to case report forms, and proposed these to investigators and sponsor for buy-in and approval. Namely, an existing case report form that previously captured missed visits was expanded to collect relevant information for all visits. The modified case report form is expected to be completed for all visits and was adapted to capture whether a visit occurred outside the prescribed visit-window, if a visit was missed, and if not, where the visit occurred: in clinic, via telemedicine, and/or offsite. Furthermore, if any portion of the visit occurred at an offsite location, follow-up questions capture the location of urine sample collection, which is often the basis of primary outcome, and the location of medication dispensing/administration. Another change included the addition of COVID-19 response options to case report forms collecting data that have potential to be missing or otherwise impacted by COVID-19. To capture the nuances among (1) active COVID-19 infection, (2) lockdown-related isolation/quarantine, or (3) other COVID-19 factors including fears of exposure, we proposed a set of three response options for impacted forms: “COVID-19: Illness,” “COVID-19: Public health measures,” and “COVID-19: Other.” Likewise, to assess the impact on study procedures and assessments, a question was added to evaluate if a protocol deviation was related to COVID-19. Finally, the Data and Statistics Center included automated data queries for operational concerns related to visit flexibility. For example, if a participant’s visit occurred entirely offsite, the system issues an automated query if any incongruous visit data point is indicated as being collected in clinic or via telemedicine (e.g. biospecimen). The interdisciplinary collaboration among the Clinical Coordinating Center, Data and Statistics Center, and investigators allowed for timely updates to the electronic data capture system to capture
COVID-19-related data across Clinical Trials Network trials in a harmonized fashion and brought the Clinical Trials Network into compliance with Food and Drug Administration guidance. These data will also allow statisticians to conduct sensitivity analyses to assess the impact of COVID-19 on trial outcomes. This, along with the reasons for missing data, protocol deviations, and assessment changes due to COVID-19 will be included in final study reports of these trials.

**CP-45**

**RECOMMENDATIONS FOR VISUALIZING THE DRUG HARM PROFILE IN RANDOMIZED CONTROLLED TRIALS: A CONSENSUS**

**Primary Author:**
RACHEL PHILLIPS  
Imperial College London

**Co-Author(s):**  
VICTORIA CORNELIUS  
Imperial College London  
SUZIE CRO  
Imperial College London  
GRAHAM WHEELER  
University College London  
CAROL GAMBLE  
University of Liverpool  
GRAEME MACLENNAN  
University of Aberdeen  
TIM MORRIS  
University College London  
SIOBHAN CREANOR  
University of Exeter  
SARAH PIRRIE  
University of Birmingham  
IRYNA SCHLACKOW  
University of Oxford

**Introduction:** A well-designed graphic can effectively communicate a message to a range of audiences and help identify patterns in data that might otherwise be missed. In randomized controlled trials, when analyzing adverse events where there is an abundance of complex data, graphics can be useful to identify patterns, summarize harm profiles, and identify potential adverse drug reactions. Trial reporting guidelines such as the CONSORT extension to harms and recommendations from industry representatives and journal editors encourage the use of visualizations for exploring adverse event data and previous work shows there are an abundance of visualizations available. However, use in journal articles is limited, with prevailing practice to present data in simple tables of frequencies and percentages, despite the advantages visualizations offer. With a range of visualization options available and the increasing ease of implementation, we sought a consensus to support researchers in their choice of visualizations for randomized clinical trial publications.

**Methods:** A methodology review identified visualizations developed specifically to analyze adverse event data. Presentation of this work at academic conferences resulted in feedback about alternative visualizations that could be adapted to the harm setting. The body of available graphics was categorized according to outcome type and each was considered for endorsement. Academic and industry statisticians with a known interest in visualizations were approached and an advert was placed on the PSI visualization special interest group homepage inviting participation in three virtual consensus meetings. A framework for appraisal was developed and discussed among participants and edited based on feedback and group endorsement. Each plot was presented and discussed among the group. Participants were encouraged to raise any queries they had regarding the plot, highlight what they liked and disliked about it, where they thought it might be useful and to flag any potential problems. Each plot was then critically evaluated against the agreed framework. Appraisals were analyzed and the results used to inform discussions about which plots to retain. Participants voted to take plots forward for further discussion about use and potential refinements. Endorsement for each proposed refinement was gathered, before appearance and accompanying recommendations were drafted. Clinicians with clinical trial experience were then invited to participate in interviews to gather feedback on the drafted recommendations and pertinent comments raised were incorporated into the final recommendations.

**Results:** The group endorsed consideration of the dot plot and stacked bar chart (to incorporate information on severity) to visually summarize the overall harm profile in the main research publication. Full recommendations by outcome type and number of events will be presented using example data sets. We will demonstrate how visualizations offer an alternative way to communicate risks of harms, highlighting the advantages they offer over simple frequency tables and include relevant code for implementation.

**Conclusion:** Visualizations provide a powerful tool to communicate harms, offering alternative perspectives to the traditional frequency tables. Implementation of these recommendations will help to improve adverse event reporting in clinical trials, present clearer summaries of harm profiles, and identify both the burden of harm patients experience and potential signals of adverse drug reactions for further monitoring.
CP-46
ESTIMATING TREATMENT EFFECT HETEROGENEITY IN STEPPED-WEDGE CLUSTER-RANDOMIZED TRIALS WITH BINARY OUTCOMES

Primary Author: LARA W MALEYEFF
Harvard University

Co-Author(s): RUI WANG
Harvard T. H. Chan School of Public Health, Harvard Pilgrim Health Care Institute and Harvard Medical School

A stepped-wedge cluster-randomized trial is a unidirectional cross-over study in which timings of treatment initiation for clusters (e.g. hospitals, clinics) are randomized. We focus on cross-sectional stepped-wedge cluster-randomized trials where data from a random sample of individuals at each time step from each cluster are collected to estimate the treatment effect. For example, we consider a stepped-wedge cluster-randomized trial assessing whether tuberculosis patients diagnosed with XpertMTB/RIF, a molecular-based test that provides rapid diagnosis of tuberculosis, have a higher odds of curative treatment completion, as compared with those diagnosed with sputum spear examination.

Because the timing of treatment initiation is different for each cluster, one question naturally arises: Does treatment effect depend on the exposure time, that is, the time since the initiation of treatment. Treatment effect heterogeneity over exposure time can be modeled using interaction terms between the treatment indicator and exposure time. Existing methods include (1) assuming a linear trend; (2) assuming a different treatment effect at the first step, then a common treatment effect for following steps; and (3) allowing differential treatment effect for each exposure time by modeling step as an indicator function. While Approaches (1) and (2) enjoy fitting advantages due to more parsimonious model formulation, they are less robust and may produce biased results when the assumed functional forms are incorrect. While Approach (3) does not require assumptions on how treatment effect changes over exposure time, the number of parameters increases as the number of steps in the stepped-wedge cluster-randomized trial increases, which complicates the model fitting procedure.

In this work, we propose a random effect formulation for the treatment effect heterogeneity over exposure time. The proposed model provides an alternative solution to model treatment effect heterogeneity to existing approaches, especially when the number of steps is large. It is more flexible than the approaches assuming a specific functional form for the treatment effect heterogeneity (e.g. Approaches 1 and 2), and can be more efficient than Approach 3 by pooling information from different exposure timepoints. We conduct extensive simulation studies to compare the performance of different model formulations and provide guidance on the practice use of these models to evaluate treatment effect heterogeneity across exposure time in cross-sectional stepped-wedge cluster-randomized trials.

CP-47
COVARIATE ADJUSTMENT IN CARDIOVASCULAR RANDOMIZED CONTROLLED TRIALS: A REVIEW OF CURRENT PRACTICE AND SIMULATION STUDIES

Primary Author: LEAH G PIRONDINI
London School of Hygiene and Tropical Medicine

Co-Author(s): RUTH OWE
London School of Hygiene and Tropical Medicine
JOHN GREGSON
London School of Hygiene and Tropical Medicine
STUART POCOCK
London School of Hygiene and Tropical Medicine

Introduction: There is inconsistency regarding the use of covariate adjustment in randomized controlled trials. This has been highlighted previously in surveys of the use of covariate adjustment carried out in 2000 and 2010. A recent review of current practice is needed.

Methods: We surveyed all 84 randomized controlled trials relating to cardiovascular disease published in The New England Journal of Medicine, The Lancet, and The Journal of the American Medical Association during 2019. Detailed information on covariate adjustment was collected and tabulated. Case studies of trials where covariate adjustment affected the trial conclusions were identified and described in more detail. A review of methodological papers relating to the use of covariate adjustment in randomized controlled trials was performed. The effect of covariate adjustment on estimated hazard ratios and corresponding standard errors was investigated using simulated time-to-event data for a hypothetical trial.

Results: Of the 84 trials surveyed, 31 performed only a simple treatment comparison, unadjusted for baseline covariates, whereas 53 used covariate adjustment as part of the primary or secondary analysis. Of these 53 trials, 38 gave primary (or equal) emphasis to the adjusted results. The number of covariates included in the adjusted analysis varied, with a median of two and
a maximum of 13. The reasons for the choice of covariates were not always clearly explained but common themes were adjusting for variables used in the randomization procedure, covariates expected to be predictive of the outcome, and, to a lesser extent, covariates imbalanced between groups. In three of 24 trial reports presenting both unadjusted and covariate-adjusted analyses, the adjustment affected the trial conclusions. These and other topical examples will be presented. Simulation results confirm that covariate adjustment for time-to-event outcomes using proportional hazard models leads to an increase in statistical power, only when chosen covariates are predictive of the outcome. Adjusting for baseline variables that show imbalance between treatment groups but do not affect prognosis will make negligible difference. Conversely, a non-significant covariate imbalance can matter if the covariate is strongly related to the outcome.

**Conclusion:** The use and reporting of covariate-adjusted analyses in randomized controlled trials varies enormously. A clearly defined appropriate covariate-adjusted analysis can lead to an increase in statistical power (and hence, a potential reduction in sample size). Our findings suggest the need for greater consistency across trials in the use of covariate adjustment, the choice of relevant baseline covariates, and the reporting of results.

**CP-48**

**PREDICTION OF TREATMENT OUTCOME AND HETEROGENEOUS TREATMENT EFFECT IN CLINICAL TRIALS: A MACHINE LEARNING APPROACH**

**Primary Author:**
PAOLA BERCHIALLA  
*University of Torino*

**Co-Author(s):**
CORRADO LANERA  
*University of Padova*  
VERONICA SCIANNAMEO  
*University of Padova*  
DARIO GREGORI  
*University of Padova*  
ILEANA BALDI  
*University of Padova*

**Introduction:** A common concern to apply randomized controlled trial-based estimates to a target population is that there may be many clinical features that differ from the randomized clinical trial study population. A central problem that arises in most data-driven personalized medicine scenarios is the estimation of heterogeneous treatment effects to stratify individuals into subpopulations that differ in their response to a specific treatment and identify who benefit most from a particular treatment. Machine Learning algorithms are particularly promising for identifying treatment effect modifiers, given their capability to learn by examples.

**Objectives:** The objective of this work is to exploit Machine Learning predictive capabilities to identify systematic variation in randomized clinical trial–treatment outcome, separate it from the variation due to the sampling error, and target responsive subgroups of patients.

**Materials and Methods:** We conducted a sub-analysis on the PROLOGUE randomized clinical trial, which examines whether DPP-4 inhibitors provide better glycemic control to conventional therapy in patients with type 2 diabetes, and we investigated a potential heterogeneous effect on the improvement of glycated hemoglobin (HbA1c). Since Machine Learning algorithms need to learn the statistical dependencies between clinical features and patients’ treatment outcomes, we focused on the SAIS1 randomized clinical trial to train the outcome prediction model. Both SAIS1 and PROLOGUE randomized clinical trials have studied the effect of sitagliptin on patients with type 2 diabetes and have collected a common subset of patient measures. The outcome prediction model was used to assign a probability to be a responder (to get a reduction at 12 months of HbA1c at least 0.5%) to each patient in the PROLOGUE study. An Machine Learning model based on the combination of gradient boosting machine, generalized and polychotomous linear model, random forest, Bayesian additive regression trees, and support vector machine was adopted.

**Results:** The SAIS1 study enrolled 103 patients with type 2 diabetes who were allocated on glimepiride (n = 55) and on sitagliptin (n = 48). The PROLOGUE study enrolled 385 patients with type 2 diabetes, who were allocated on conventional therapy (n = 193) and on sitagliptin (n = 192). The performance of the outcome predictive model developed on SAIS1 was assessed by the cross-validated area under the curve and resulted to be 92.05%. Overall, 376 out of 385 patients in PROLOGUE randomized clinical trial showed a probability of getting a reduction of HbA1c of at least 19.3%. The best treatment effect is achieved in a subgroup of 250 patients selected at the probability value of 41.3%. The median reduction of HbA1c at 12 months is \( -0.2 \) (interquartile range = \(-0.5; 0\)) among the 122 best responsive patients in the conventional group and \(-0.4 \) (interquartile range = \(-0.7; 0.2\)) among the 131 best responsive patients in the sitagliptin group, \( p = 0.013 \); the two groups are balanced for baseline characteristics.

**Conclusion:** We used an ensemble Machine Learning algorithm to find evidence of patient-specific treatment effect. In the literature has been shown how individual response can be estimated in a replicated crossover
study. Here, we used two randomized clinical trials—similar in terms of inclusion, exclusion criteria, and estimands—to train a Machine Learning model to conduct heterogeneous treatment analysis and target best responders.

**CP-49**

A BAYESIAN GROUP SEQUENTIAL DESIGN FOR RANDOMIZED BIOSIMILAR CLINICAL TRIALS WITH ADAPTIVE INFORMATION BORROWING FROM HISTORICAL DATA

**Primary Author:**
WEN ZHANG
University of Texas Health Science Center at Houston

**Co-Author(s):**
ZHIIYING PAN
Amgen, Inc.

YING YUAN
University of Texas MD Anderson Cancer Center

At the time of developing a biosimilar, the reference product has been on market for years, and thus, ample data are available on its efficacy and characteristics. We develop a Bayesian adaptive design for randomized biosimilar clinical trials to leverage the rich historical data on the reference product. The design takes a group sequential approach. Patients initially are 1:1 randomized into the test arm and the reference arm. At each interim, we employ the elastic meta-analysis predictive prior methodology to adaptively borrow information from the historical data of the reference product, and calculate the posterior probability that the efficacy of the test product is located within the biosimilar margins, with respect to the reference product. This posterior probability is used to make go/no-go decision. If the decision is “go,” we determine the amount of information that is borrowed from the historical data, measured by the effective sample size, and adaptively adjust the subsequent randomization ratio with the goal to balance the sample size of the two arms at the end of trials. Simulation study shows that the proposed Bayesian adaptive design can substantially reduce the sample size of the reference arm, while achieving comparable power as the traditional randomized clinical trials that ignore the historical data. We apply our design to a biosimilar trial for treating breast cancer patients.

**CP-50**

HOW INFORMATIVE IS INFORMATIVE IN SMALL SIZE TRIALS: A SIMULATION STUDY

**Primary Author:**
DANILA AZZOLINA
University of Padua

**Co-Author(s):**
ILEANA BALDI
University of Padova

PAOLA BERCHIALLA
University of Torino

DARIO GREGORI
University of Padova

The recruitment of the total sample given at the design stage is difficult in many clinical trial settings. There may be some issues with the registration of the participants, and the amount of information conveyed by a prematurely terminated trial for inadequate accrual may be negligible. A Bayesian analysis of such trials may salvage the information by providing a framework in which to combine prior evidence with current evidence. However, the Bayesian inference may be severely conditioned by the priors, especially for small trials. A pediatric research that was a candidate for early termination due to underrecruitment, the RESCUE (REnal SCarring Urinary InfEction) study, served as a motivating example to examine the prior impact on small trial inference. The possible trial outcomes were simulated by assuming 50 scenarios; each of them combines different sample sizes (14–240 by a pace of 25) and true Absolute Risk Reduction (−0.07 to −0.27 by a pace of 0.05). For each scenario, the data were simulated via 5000 Monte Carlo resamples and analyzed via the Bayesian approach using (1) beta power prior without discounting (informative); (2) beta power prior with 50% discounting (low-informative); and (3) beta power prior with 100% discounting (uninformative). The frequentist approach has been also considered for comparison purposes.

This simulation study demonstrated that, especially for small sample sizes, the trial results could highly be influenced by the prior choices and weakly influenced by the data when using informative priors. The probability to confirm a treatment effect is data-insensitive when using a full informative prior distribution. The situation is different placing a discounting factor on the prior parameters; the probability to confirm the trial results, in this case, is demonstrated to be more data-dependent and, for sample sizes lower than 50, is higher than 80% only for Absolute Risk Reduction higher than 0.17. This study demonstrated also that the Bayesian inference is more able to detect also a small treatment effect, for small sample sizes (lower than 50), even though the prior is fully uninformative in comparison to a maximum likelihood approach. A full informative Bayesian inference, conducted on small samples, could generate data-insensitive results. However, the use of uninformative prior distribution may include, in the final inference, extreme treatment effect hypotheses not clinically proven. The use of a power prior approach on sample sizes smaller than 50 seems to be a
good compromise between these two approaches. However, the choice of parameters and discounting factors should be negotiated with the expert pediatricians and should be guided by an appropriate consultation of the scientific literature. In agreement with the Food and Drug Administration recommendations, a sensitivity analysis on prior choices is highly recommended.

**CP-51**

SYMBOLIC TWO-STEP METHOD COMPARED WITH SINGLE-STEP METHODS TO MODEL THE CENTER MEAN OUTCOME IN CLUSTER-RANDOMIZED TRIALS

**Primary Author:**
DAVID ZAHRIEH
*Mayo Clinic*

**Co-Author(s):**
JENNIFER LE-RADEMACHER
*Mayo Clinic*

**Background:** The recently developed symbolic two-step method provides an alternative to the single-step multi-level linear model in the design and analysis of cluster-randomized trials evaluating interventions within care delivery research. This method applies the symbolic data analysis framework to adjust for patient-level factors when estimating and testing effects of center-level factors on both the average center-level outcome and its variation. Estimation and testing of center-level effects on center outcome variation of the patient-adjusted outcomes are the innovation of the symbolic two-step method. Moreover, our prior works showed that the two-step method performed comparably to the single-step models even under scenarios that favored single-step models. Prior results were evaluated under the assumptions of uniform center sizes, no correlation between patient-level and center-level factors, and homogeneous within-center variance. The current work was undertaken to evaluate the performance of the symbolic two-step method under more general and challenging settings including (1) small or variable center sizes, (2) correlated patient-level and center-level factors, and (3) heterogeneous within-center variances.

**Methods:** The scenarios corresponding to each setting included (1) large sample setting with severe center imbalance (80% of the N patients belong to 20% of the n centers); small sample setting (n = 16; m = 10; N = 160); small sample setting with m patients per center following a uniform distribution on the range 2–22; (2) correlation between a patient-level and center-level factor with coefficient of 0.7; and (3) the log variances varying from center-to-center as a parametric function of a single explanatory variable. For each scenario, we compared the average estimate over the 10,000 data sets generated as a measure of the true estimate, the corresponding empirical standard error, and the power associated with the t-statistic for the regression coefficient, with model-based standard error, testing the null hypothesis that the parameter equals zero.

**Results:** Under all scenarios, the single-step models provided higher statistical power for patient-level factors while the two-step method was better at controlling the type 1 error for center-level factors. In the presence of varying center sizes (1), the single-step models led to increased statistical power for center-level factors than the two-step method. However, in the presence of correlation (2), the two-step method was more powerful than the single-step models while maintaining the type 1 error. Assuming a parametric function of the log within-center variance that was parameterized to reflect heterogeneous mean log within-center variances according to the level of the explanatory variable (3), the results were similar to those observed assuming homogeneous mean log within-center variances. Results from applying the two-step method and the single-step models to a cluster-randomized trial will be presented.

**Conclusion:** Based on the simulation settings studied, we recommend the single-step multi-level linear models for studies where no correlation between patient-level and center-level factors is expected and there is no center-level factor affecting center outcome variation. Otherwise, we recommend the symbolic two-step method which is not vulnerable to correlation between individual and center-level covariates as is often seen in practice and allows for simultaneous estimation and testing of center-level effects on center outcome variation.

**CP-52**

THE USE OF CONDITIONAL POWER IN VACCINE TRIALS WITH SEASONAL VARIATIONS

**Primary Author:**
XINYU TANG
*US Food and Drug Administration*

**Co-Author(s):**
JOHN A SCOTT
*US Food and Drug Administration*

**Background:** Conditional power, the conditional probability of a significant result at the end of a trial given data observed thus far and specific assumptions about future data, has been widely used in clinical trial monitoring to quantify the evidence for early stopping for futility or sample size re-estimation during the trial. Clinical trials studying vaccine candidates against seasonal infectious diseases are often conducted across multiple seasons due to low disease attack rates. Seasonality and the dynamics of infectious diseases are
likely to produce variations in disease attack rate and/or vaccine efficacy from one season to the other. Seasonal variations impose challenges in monitoring trials using conditional power when uncertainty in future data exists.

Objective: Conditional powers calculated using the effect size hypothesized in the planning phase or the estimated effect size from the interim analysis may not truly reflect future data in vaccine trials with seasonal variations. With the belief that the effect size is likely to change in time in vaccine trials, we aim to investigate the use of six different drift parameters to provide flexibility in computing conditional power.

Methods: Vaccine trials conducted across two seasons (Season 1 and Season 2) were simulated to evaluate the differences in conditional powers calculated using different drift parameters at the end of Season 1 under different scenarios of seasonal variations. We calculated conditional powers under the alternative, under the current trend, and under a combination of both, a weighted conditional power between the alternative and current trend, a predictive power (i.e. a Bayesian alternative to conditional power), and a weighted predictive power. The impact of various conditional powers on the expected sample size and type 2 error rate/power of the trial was further investigated when conditional power was utilized as the statistical measure to inform futility/sample size re-estimation in vaccine trials with seasonal variations.

Results: Our simulations showed a pattern: conditional power under the alternative > predictive power > weighted predictive power > weighted conditional power > conditional power under the current trend, while conditional power under a combination of both was largely impacted by the attack rate during Season 1. When conditional power was used to inform futility at the end of Season 1, the expected sample sizes were comparable across different conditional powers and PPs, while the greatest and least type 2 error rates were observed with conditional power under the current trend and predictive power/weighted predictive power, respectively. As the vaccine efficacy in Season 1 increased, the type 2 error rate increased when futility analysis was performed. When conditional power was used to inform sample size re-estimation at the end of Season 1, all the conditional powers and PPs resulted in similar expected sample sizes while the greatest and least power were observed with the weighted predictive power and conditional power under the current trend, respectively.

Conclusion: Predictive power and weighted predictive power seem to outperform other conditional powers studied in terms of type 2 error rate or power. In vaccine trials with seasonal variations, the performance of conditional powers calculated under reasonable alternatives consistent with the data and with prior hypotheses should be considered before recommending stopping for futility or sample size re-estimation.

CP-53
DEVELOPING AND VALIDATING CLINICAL DYNAMIC PREDICTION JOINT MODELS TO PREDICT PROGNOSIS IN PROSTATE CANCER PATIENTS

Primary Author:
HARRY PARR
The Institute of Cancer Research

Co-Author(s):
NURIA PORTA
The Institute of Cancer Research
ALISON TREE
Royal Marsden NHS Foundation Trust
DAVID DEARNALEY
The Institute of Cancer Research
EMMA HALL
The Institute of Cancer Research

Introduction: Prostate cancer is highly prevalent and the second most diagnosed cancer in men worldwide (1.3 m cases in 2018) [1]. When patients are diagnosed with localized prostate cancer, they usually exhibit high concentrations of prostate-specific antigen, a protein of interest secreted by the prostate. Typical first-line treatment for localized prostate cancer is a combination of neoadjuvant and concurrent hormone therapy, with hypofractionated (fewer but larger doses) external-beam radiotherapy [2]. Hormone therapy inhibits testosterone and radiotherapy ablates the tumoral tissue, resulting in very reduced prostate-specific antigen concentrations. Prognosis is typically predicted using clinical baseline prognostic risk factors, including age, tumor grade, and stage. Following treatment, prostate-specific antigen levels are monitored over time and provide vital information to inform clinicians if reappearance of cancer is likely to occur; a sharp rise in post-treatment prostate-specific antigen being indicative of prostate cancer recurrence. There is therefore an opportunity to improve prognostic models by supplementing or updating baseline information with prostate-specific antigen trajectories over time. This could improve predictions and aid clinicians in creating personalized post-treatment follow-up plans.

Objectives and Methods: We develop clinical dynamic prediction joint models to characterize prognosis of localized prostate cancer patients. The joint modeling methodology is used, developed from CHHiP [2] patients’ data (N = 3071), the largest randomized control clinical trial of treatment for localized prostate cancer. We model both longitudinal prostate-specific
antigen concentrations, and time-to-recurrence outcomes simultaneously, to predict patient prognosis. The nonlinear prostate-specific antigen is captured by fitting cubic splines and the parameters are estimated within a Bayesian framework. Dynamic predictions are obtained, calculating future risk of recurrence by incorporating repeated prostate-specific antigen measurements observed up to a given timepoint $t$.

**Results:** Several association structures are evaluated giving broadly consistent results. The association between the two outcomes is given by a linear combination of the value and gradient parameterization of prostate-specific antigen. The strength of association for both outcomes is quantified, value-gradient log-hazard ratios of prostate-specific antigen are 4.4 and 2.1, respectively (both $p < 0.001$). Performing internal 10-fold cross-validation (repeated 10 times) showed including follow-up PSAs improved dynamic predictive discrimination and calibration. An overall median area under the curve of 0.82; interquartile range = (0.75–0.87) and favorable low Brier score, median = 0.07; interquartile range = (0.05–0.09), these metrics are optimal by five-years’ worth of longitudinal data.

**Conclusion:** These encouraging results show the model is suitable for use within current clinical treatment pathways, with the intention of stratifying and confirming those patients at greater risk of cancer recurrence. For example, directing additional scans if PSAs surpass an unacceptable risk threshold and direct salvage care accordingly, or conversely reduce assessment burden for those patients with low risk. Ongoing work includes external validation from other clinical trials, required to show the model is suitable and generalizable in wider populations for prognostic use.

**References**

[1] Rawla P. World J Oncol 2019; 10: 63–89.
[2] Dearnaley D, et al. The Lancet Oncology 2016; 17: 1047–1060.

**CP-54**

**DESIGN AND ANALYSIS OF PARTIALLY RANDOMIZED PREFERENCE TRIALS WITH PROPENSITY SCORE STRATIFICATION**

**Primary Author:**
YUMIN WANG  
*Yale School of Public Health*

**Co-Author(s):**
FAN LI  
*Yale University*
ONDREJ BLAHA  
*Yale University*
CAN MENG  
*Yale University*

DENISE ESSERMAN  
*Yale University*

Abstract body: The two-stage randomized preference design is a powerful method to evaluate patient preference on treatment outcomes through separating the selection effect and preference effect from the treatment effect. However, the first-stage randomization in these designs can lower the motivation for participation in patients who have preference for one of the treatments under consideration. The partially randomized preference design replaces the first-stage randomization with a procedure that allows patients to state and receive their preferred treatment, at the cost of potentially introducing bias in estimating the treatment, selection, and preference effects. We propose to use propensity score stratification in the partially randomized preference design to mimic the first-stage randomization process of the two-stage design and reduce bias in estimating the effects of interest. We derived test statistics and sample size formulas based on propensity score stratification for the treatment, selection, and preference effects. Simulation studies under different parameter settings were carried out to demonstrate bias reduction properties from propensity score stratification. Our results showed that the number of strata needed to achieve the best match between predicted power and empirical power in a given parameter setting may depend on the bias-variance trade-off under that setting. We recommend careful specification of the propensity score model and inclusion of important key baseline covariates in estimating the propensity score, which reduces confounding bias for both design and analysis of partially randomized preference trials.

**CP-55**

**PATIENT AND PUBLIC INVOLVEMENT IN NUMERICAL ASPECTS OF TRIALS: A THEORY-INFORMED SURVEY OF TRIALISTS CURRENT PRACTICES, BARRIERS, AND FACILITATORS**

**Primary Author:**
BEATRIZ GOULAO  
*University of Aberdeen*

**Co-Author(s):**
CAMILLE POISSON  
*University of Aberdeen*
KATIE GILLIES  
*University of Aberdeen*

**Background:** Patient and Public Involvement is increasingly common in trials and acknowledged as a way to increase research relevance as well as a requirement by funders. Numerical aspects are crucial in the design, conduct, analysis, and dissemination of trials; however,
the extent to which patient and public partners have a say in these aspects is unknown.

Objective: We aimed to find out if trialists involve patients and the public in numerical aspects of trials, how and what are the barriers and facilitators to doing it.

Methods: We developed a survey based on the theoretical domains framework, a framework developed to understand adoption of professional behavior. We used a mixed-methods approach to analyze the data calculating descriptive statistics for closed questions and using a deductive and inductive thematic qualitative analysis for open questions. Our survey included the UK-based trialists and was disseminated using relevant mailing lists and social media accounts.

Results: We included 187 responders. The majority were female (70%), trial managers (67%) and involved public and patient partners in numerical aspects of trials (60%). Responders to the full survey were similar to non-responders in most aspects measured. We found lack of knowledge, trialists’ perception of public and patient partners’ skills, capabilities and motivations, scarce resources, lack of reinforcement, and lack of guidance were barriers to involving public and patient partners in numerical aspects of trials. Positive beliefs about consequences were an incentive to doing it.

Discussion: More training, guidance, and funding can help trialists involve patient and public partners in numerical aspects, although they were uncertain about public and patient partners’ motivation to be involved. Future research should focus on finding public and patient partners’ motivations, and develop strategies to improve communication of numerical aspects.

CP-56

RESEARCH READY: STRATEGIES FOR SUCCESSFUL IMPLEMENTATION OF RESEARCH IN CLINICAL SETTINGS

Primary Author:
DANIELE FARRISI
Louisiana Public Health Institute

Co-Author(s):
PEGGY SANDERS
Louisiana Public Health Institute
BETH NAUMAN
Louisiana Public Health Institute

Through our experience implementing pragmatic studies, the Louisiana Public Health Institute recognized clinic support staff as unique stakeholders in clinic-based research whose role in study implementation is often overlooked. Our project, Research Ready, aimed to find innovative ways to engage staff in the design and implementation of research studies. Specifically, the project team designed, piloted, and disseminated materials to improve clinic staff capacity to partner in research. During this pandemic, reliance on clinical staff for study adherence is critical. COVID-related research has been rapidly implemented, which relies upon having well-prepared staff to handle the rapid implementation of new protocols.

The team developed and piloted two tools for improving staff engagement in research activities: a training for clinic support staff and a guide for researchers. The staff training was developed to inform clinic staff, such as medical assistants and nurses, about basic research principles and considerations for supporting the implementation of research in a clinical setting. The training is available in three formats: e-learning, facilitated session, and self-guided workbook. The researcher guide was created to share insights and best practices for engaging and partnering with clinic staff to successfully implement pragmatic research. Both resources are available on Louisiana Public Health Institute’s website.

The Research Ready resources were informed by interviews conducted with clinic support staff and researchers (including principal investigators, study managers, and clinical research coordinators) who had implemented studies in outpatient settings. Clinic staff from a variety of settings were interviewed, including Federally Qualified Health Centers, private healthcare systems, and academic medical centers. Research staff were either affiliated with academic institutions or clinical research firms.

Major themes identified in the key-informant interviews include (1) clinic staff play key roles in implementing research: they are gatekeepers of clinic workflow and brokers of patient trust; (2) clinic staff lack knowledge about research and the research process, which is a barrier to implementing studies in clinic settings; (3) communication and relationship-building are important facilitators for researchers seeking to work with clinical staff; (4) clinic staff prioritize the care and wellbeing of their patients, which can be both a barrier and a facilitator of clinic-based research.

The training was piloted in three sites with a total of 52 participants. Participants were surveyed after completing the training. Survey results showed that participants thought the training was easy to understand and increased their knowledge about research. Results also showed that participants felt the information from their training was applicable to their jobs.

As more research is conducted in clinic settings, researchers and clinic staff will benefit from best practices to assure a mutual understanding of research objectives and processes. Identifying strategies for successful implementation of research in clinical settings will enhance the conduct of pragmatic research and allow it to equitably reach patients in diverse outpatient care settings. Using the Research Ready materials,
researchers can ensure that clinic staff have adequate understanding of research principles and that staff concerns about time and competing priorities are addressed and accounted for in study workflows.

**CP-57**

**COSTS AND STAFFING RESOURCE REQUIREMENTS FOR ADAPTIVE TRIALS: RESULTS FROM THE COSTING ADAPTIVE TRIALS PROJECT**

**Primary Author:** JAMES WASON  
*Newcastle University*

**Co-Author(s):**  
NINA WILSON  
*Newcastle University*  
KATIE BIGGS  
*Sheffield University*  
REBECCA MAIER  
*Newcastle University*  
HELEN HANCOCK  
*Newcastle University*

**Background:** Adaptive designs are increasingly recognized as an important tool in improving the efficiency of clinical trials. Nevertheless, some barriers to their use remain. One important barrier is lack of understanding about whether additional staffing and financial resources are required to conduct a high-quality adaptive clinical trial, compared to a traditional non-adaptive trial.

**Aims:** The Costing Adaptive Trials project, funded by the National Institute for Health Research in the United Kingdom, was set up to:

1. Investigate the additional costs that result from designing, conducting, and analyzing academic-led adaptive trials.
2. Provide guidance on what additional costs should be included in future funding applications.
3. Identify methodology research needs to reduce any additional costs of conducting adaptive trials in the future.

In this presentation, we concentrate on aim (1).

**Methods:** We conducted a mock costing exercise in seven UK academic Clinical Trials Units. Five scenarios were developed, based on actual clinical trials, each describing a non-adaptive version and an adaptive version. Each scenario covered a different type of adaptive feature: (1) sample size re-estimation, (2) group sequential, (3) multi-arm multi-stage, (4) phase 2b dose-finding, and (5) adaptive umbrella design.

Clinical Trials Unit staff were asked to provide the costs and staff time they estimated would be needed to support the trial, categorized into specified tasks (e.g. statistics, data management, and trial management). This was done separately for the non-adaptive and adaptive version of the trial, allowing paired comparisons. Clinical Trials Units did not necessarily cost all scenarios, for example, if the trial laid outside their unit’s scope. After the costing exercise, semi-structured interviews with all Clinical Trials Unit staff who conducted the exercise were conducted to explore reasons for cost differences.

**Results:** The median percentage increase in total costs depended on the scenario. The highest median (across the different Clinical Trials Units) percentage increase, 29.7%, was from using a sample size re-estimation design, although this was due to the potential increase in the length of the trial for additional recruitment. Other adaptive designs, which did not affect the length of the study, had more modest median percentage increases (<5.5% in each case). Qualitative interviews captured drivers for additional costs: these were predominantly statistics, data management/programming, and study management. Analyses are ongoing in Q1-2021 and will be presented in detail at the conference.

**Discussion:** This work sheds light on additional financial and staffing resources required to adequately support a high-quality adaptive trial. Building on the results of this research, we will develop guidance on the factors to consider and how to appropriately resource an adaptive trial. This research will lead to guidance that will reassure funders about the use of adaptive trials and should facilitate and accelerate their implementation.

**CP-58**

**USING IDEAL TO EVALUATE RESEARCH WASTE IN SURGERY MORE REALISTICALLY**

**Primary Author:** PETER MCCULLOCH  
*University of Oxford*

**Co-Author(s):**  
JIJIE YU  
*IDEAL China centre*  
FEI SHAN  
*Peking University Cancer Hospital & Institute*  
MUDATHIR IBRAHIM  
*Maimonides Hospital*  
ALLISON HIRST  
*University of Oxford*

**Background:** Broad analyses of randomized trials of treatments has suggested that up to 85% of research
effort is wasted due to addressing the wrong questions, faulty study design and management, and inadequate reporting. However, the majority of clinical studies in many specialties including surgery are comprised of study types other than randomized clinical trials. These cannot be evaluated with tools designed for randomized clinical trials, and no existing tool is able to deal with all of them. The IDEAL Recommendations are derived from basic principles of good clinical and ethical practice, and may provide a way of analyzing waste in both randomized clinical trials and other study types. We evaluated its potential in a randomized sampling study of current literature.

Methods: We developed and used highly specific search strategies to identify studies focused on the main outcomes of (1) surgical operations and (2) therapeutic devices, in Ovid Medline between the of the periods of January to December 2018. We randomly sampled these studies to select 100 reports from each pool, and used a newly devised tool for estimating compliance with the IDEAL Recommendations to score them as high, medium and low compliance. We calculated the total number and percentage of studies affected by waste and the number with low compliance, and analyzed the percentage in randomized clinical trials and in other study designs.

Results: The study is still ongoing. Preliminary review of data suggests that (a) research waste is prevalent in non-randomized clinical trials studies and (b) since non-randomized clinical trials make up the majority of studies in both groups (a) and (b), their overall contribution to research waste is considerably large. The contributions to waste of specific aspects of studies will be presented.

Conclusion Ideal: Recommendations can be used to estimate research waste in outcome studies for complex treatments. The major sources of waste can easily be identified. This can be valuable in making recommendations to improve methodology.

CP-59
EVALUATING THE EFFECT OF A MOBILE AUDIO COMPANION TO REDUCE FEELINGS OF DISTRESS AND LONELINESS IN CANCER PATIENTS DURING COVID-19 (COMPANION TRIAL)

Primary Author:
GILLIAN GRESHAM
Cedars-Sinai Medical Center

Co-Author(s):
CELINA H SHIRAZIPOUR
Cedars-Sinai Medical Center

Background: Given the current physical isolation and distancing measures in place as a result of the COVID-19 pandemic, and the health risks associated with leaving one’s home, cancer patients face new and unprecedented challenges in connecting with their communities and conducting activities of daily living. They are also experiencing changes in their treatment and medical follow-up plans. Consequently, they are at even greater risk of experiencing loneliness, anxiety, depression, and reduced quality of life compared to the general population. There is emerging evidence for the use of mobile health applications to improve social connections, engage patients in their own health, and provide psychosocial support among cancer survivors. Thus, we are conducting the COMPANION study to examine the use of a mobile health application, Elly, to reduce levels of anxiety, stress, and feelings of social isolation and loneliness as a result of isolation during the COVID-19 pandemic among cancer patients and survivors (NCT04604158).

Methods: This is a prospective, interventional study of patients who self-identify as currently receiving cancer-targeted treatment. All study recruitment and consent procedures, assessments, and follow-up are completed remotely using REDCap. Participants are given access to the Elly phone application developed by Elly Health Inc. and followed for 6 months. Elly is a multi-platform mobile solution that enables cancer patients to receive support and assistance in the management of their illness. Elly delivers audio content through a mobile app to comfort, inspire and motivate its users including tools to manage their diagnosis and symptoms including interactive sessions on motivation, unspoken (taboo topics), exercise, mindfulness, sleep, nutrition, self-care, and symptom as well as COVID-19 specific topics.

Results: The COMPANION trial is currently open to accrual. Approximately 100 patients will be enrolled between November 2020 and December 2021. Trial design and protocol development, regulatory activities
including submission to the local ethics review board, and communication was completed remotely and included an interdisciplinary team of clinical trialists, behavioral scientists, oncologists, and psychiatrists. This study provides an example of how clinical trials may be developed and managed in a decentralized manner and under budget and time constraints. Surveys will be administered through REDCap at baseline and monthly thereafter to assess the patient’s social status, medical history (including history of COVID-19), and National Institutes of Health PROMIS anxiety, depression, perceived stress, pain, loneliness, social support, and global health scales. Participants may opt to wear and link activity trackers in order to obtain objective measures of daily activity. Patients will also have the option to consent to the optional qualitative interview sub-study.

**Conclusion:** Using COMPANION as a case example, we demonstrate the feasibility of designing and conducting a trial to monitor patient outcomes and deliver a mobile health intervention in a remote setting. Findings from this work will provide fundamental knowledge of the additional impact COVID-19 has had on cancer patients’ anxiety, depression, and stress levels. This study will also improve our understanding of the short- and long-term effects of COVID-19 and how it’s trajectory over the course of the next 6 months and beyond will impact these outcomes.

**CP-60**

**MOTIVATION FOR PARTICIPATION IN A CLINICAL TRIAL AMONG YOUTH AND ADULTS WITH PREDIABETES OR RECENTLY DIAGNOSED TYPE 2 DIABETES**

**Primary Author:**
ASHLEY H TJADEN
George Washington University

**Co-Author(s):**
KAREN ATKINSON
VA Puget Sound; University of Washington

MARY SAVOYE
Yale University

KARLA TEMPLE
University of Chicago Medical Center

ELENA BARENGOLTS
University of Illinois

TAMARA S HANNON
Indiana University

SHARON L EDELSTEIN
George Washington University

**Background:** Designing and implementing retention strategies to engage participants starting at the first recruitment visit are key to the successful completion of clinical trials. Difficulty in recruiting and retaining participants can threaten the validity of clinical trial results, and may have financial and ethical implications. While previous research evaluated motivation for participation in clinical trials, there is little research comparing the differences in motivation between youth and adults. Feedback from participants in the Restoring Insulin Secretion Study provided this insight and highlight participant willingness to take part in a trial with extensive and invasive study procedures, including the completion of five oral glucose tolerance tests and three hyperglycemic clamp tests over 21 months.

**Methods:** The Restoring Insulin Secretion Adult Medication Study and the Restoring Insulin Secretion Pediatric Medication Study randomized 267 adults (20–65 years) and 91 youth (10–19 years) with impaired glucose tolerance or recently diagnosed, type 2 diabetes to receive 12 months of pharmacologic treatment followed by observation for up to 9 months off treatment. Of those participants, 218 (81.6%) adults (54.1 ± 8.6 years; 44% female) and 79 (86.8%) youth (14.5 ± 2.1 years; 73.4% female) completed an exit survey in which they were asked to rate the importance of reasons for participating (e.g. free medical tests, learning about diabetes, relationship with staff, monetary, free medication, and taking care of health) and problems encountered while participating (scheduling, transportation, length of visits, medical problems, staff, medications, extensive tests, etc.).

**Results:** For 85.7% of the respondents, Restoring Insulin Secretion was the first research study in which they had participated (86.6% adults and 83.1% youth). Based on their experience, nearly all participants (97.6%) reported that they would participate in another research study, and a similar proportion (96.3%) felt that they gained something personally from their participation. When comparing youth to adults, adult participation was slightly more altruistic than youth participation. For example, youth were more likely to agree or strongly agree that payment (77.9% in youth vs 60.5% in adults, p = 0.006) and free medications (92.4% in youth vs 73.5% in adults, p = 0.005) were reasons to participate. Youth were more likely to strongly agree that part of their motivation to participate was a family history of diabetes (69.2% of youth and 52.8% of adults, p = 0.01), and they were more likely to report not liking to take Restoring Insulin Secretion medications (10.3% in youth vs 7.4% in adults, p < 0.046), and not liking to test their blood sugar (16.2% in youth vs 4.3% in adults, p < 0.012) than adult participants. Within adults and youth, there were no appreciable differences by age group, race/ethnicity, sex, or baseline body mass index subgroups. In free text responses, both youth and adult participants included comments about their relationship with the Restoring Insulin Secretion staff being their motivating factor.
force. Future analyses will include sentiment analyses of the free text responses.

Conclusion: The Restoring Insulin Secretion results show that despite challenging and complicated clinical trial protocols, participants are willing to participate and feel personal gain from their involvement. Future analyses may include assessing potential relationships between participant motivation and study outcomes.

CP-61
CENTRALIZED REMOTE AND RISK-BASED MONITORING IN THE INFLUENZA VACCINE TO EFFECTIVELY STOP CARDIO THORACIC EVENTS AND DECOMPENSATED HEART FAILURE TRIAL: A COST-EFFECTIVE AND EASY-TO-IMPLEMENT APPROACH

Primary Author:
HOLLY SHAW
Frontier Science Foundation

Co-Authors:
ANDREW LOHRUM
Frontier Science Foundation
LU MAO
University of Wisconsin—Madison
DORIANNE KALISH-HUZA
University of Wisconsin—Madison
KENNETH WOOD
Frontier Science Foundation
SUZANNE SIMINSKI
Frontier Science Foundation
KYUNG-MANN KIM
University of Wisconsin—Madison

Monitoring of clinical trials is crucial for the protection of human subjects, the integrity of the study, validity of the data, and for safeguarding the highest standards in clinical research. It is required by good clinical practice, but the expense of traditional on-site monitoring can be prohibitive. Striking a balance between the level of monitoring needed and meeting the good clinical practice requirements while remaining fiscally viable is a delicate act, but can be successfully achieved with a targeted use of resources.

This paper details the use of a cost-effective strategy of using a risk-based remote monitoring model adopted in the INfluenza Vaccine to Effectively Stop Cardio Thoracic Events and Decompensated heart failure trial (ClinicalTrials.gov Identifier: NCT02787044), a pragmatic large simple trial, conducted in the United States and Canada. These strategies provide protection to human subjects and ensure the enrollment of eligible patients to the trial, two areas deemed most critical for INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure as a pragmatic trial.

When combined appropriately strategies common to centralized, remote, and risk-based monitoring ensure the safety of participants, study quality, and allow the study to remain within budgetary constraints. The INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure trial is monitored without a single planned on-site visit. The monitors use the electronic trial master file used in the study along with low-tech and easy to implement, yet secure methods to review the informed consent. The monitors also use an electronic data capture system in the study and a secure cloud-based system to access and share protected health information—redacted medical records to verify eligibility of enrolled patients.

The key to the successful remote monitoring of the trial is based on the low-tech methodologies that all sites could easily access and implement while still providing a high level of security for the subjects’ protected health information. The processes developed for INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure trial allowed for the study teams to react quickly to any issues, while still ensuring high-quality monitoring at a much lower cost than on-site monitoring. The strategies for monitoring the INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure trial could be feasible for other clinical trials similar in design and risk especially during the COVID-19 pandemic and in the future.

The INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure trial was supported by the National Heart, Lung, and Blood Institute under cooperative agreements U01HL130163 and U01HL130204 and by Sanofi Pasteur to the INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure Clinical Coordinating Center and the INfluenza Vaccine to Effectively Stop cardio Thoracic Events and Decompensated heart failure Data Coordinating Center.

CP-62
USE OF VANGUARD ENROLLMENT TO DEMONSTRATE FEASIBILITY OF A HIGH-ENROLLMENT SURGICAL TRIAL: THE SAFE CHOLECYSTECTOMY OUTCOMES CLINICAL TRIAL OF LAPAROSCOPIC CHOLECYSTECTOMY

Primary Author:
JENNIFER J GASSMAN
Cleveland Clinic

Co-Authors:
DIYA ALA青海EN
Cleveland Clinic
Conducting a randomized trial testing an intervention to make a safe procedure safer is challenging. The Cleveland Clinic SCOUT Trial will test laparoscopic cholecystectomy (minimally invasive gall bladder removal) under White plus Near-Infrared Fluorescent light versus Standard of Care White light. This goal of SCOUT is to go beyond a previous Visualization Trial [1] showing Near-Infrared Fluorescent during laparoscopic cholecystectomy improves surgeons’ ability to visualize anatomic structures to show Near-Infrared Fluorescent is safer than Standard of Care White light. The SCOUT team aims to conduct a pragmatic surgical trial investigating whether Near-Infrared Fluorescent lighting makes laparoscopic cholecystectomy even safer than it already is.

SCOUT has a primary binary composite outcome comprised of clinically important safety events, including conversion to open surgery, hemorrhage, bile duct injury, and need for reoperation. Outcome will be ascertained via routine post-surgical visit/virtual visits, medical records, and 30-day SCOUT calls/virtual visits. Literature and local data suggest that no more than 7.5% of laparoscopic cholecystectomy patients experience some component of the SCOUT composite outcome, and we aim to show a 30% reduction (from 7.50% to 5.25%) randomizing 3900 patients at five Cleveland Clinic Surgical Units. Laparoscopic cholecystectomy is a safe procedure, so high enrollment is required to show improved safety. SCOUT shows this incremental reduction, >13,500 safety outcomes would be prevented annually.

Grant funding logistics will require <48 months recruitment. Our Vanguard Enrollment aims to show SCOUT randomization is feasible at five Units in 48 months with randomization of 100 patients at five Cleveland Clinic Surgical Units. Laparoscopic cholecystectomy is a safe procedure, so high enrollment is required to show improved safety. SCOUT shows this incremental reduction, >13,500 safety outcomes would be prevented annually.

Potential SCOUT participants will be identified with surgical scheduling records and consented during routine pre-operative visits. Data show 98% of local routine cases will be eligible. After Vanguard Enrollment, we anticipate 9800 laparoscopic cholecystectomies will be performed at our five Units in 48 months, so less than 40% of Cleveland Clinic laparoscopic cholecystectomy patients will need to be randomized. (We expect most patients approached will consent, with adherence to treatment or visit schedule being unnecessary. Consent rates in the similar Visualization Trial were >90%. Also, additional Cleveland Clinic Surgical Units are available.) SCOUT’s diversity will be enhanced with African American enrollment at the Cleveland Unit and Hispanic enrollment in Weston. Because all surgical units are at Cleveland Clinic hospitals, trial conduct will be facilitated by the Cleveland Clinic’s experienced institutional review board and will benefit from the Cleveland Clinic’s seasoned National Institutes of Health/NIDDK Data Coordinating Center personnel in Quantitative Health Sciences. Vanguard Enrollment will begin this summer.

Reference
1. Dip F, LoMenzo E, White K, et al. Randomized Trial of Near-infrared Incisionless Fluorescent Cholangiography. Annals of Surgery 2019; 270(6): 992–999.

CP-63
THE TRUE POWER OF CLINICAL TRIALS IN PEDIATRIC CANCERS AND OTHER RARE DISEASES
Primary Author:
AUDREY MAUGUEN
Memorial Sloan Kettering Cancer Center

Background: Conducting efficient clinical trials is challenging in rare diseases like pediatric cancers, where the possibility of accrual is limited by the low incidence of the disease. In addition, designs of clinical trials in rare diseases make the same inference assumptions than the trials in common diseases, and assume that the sample comes from a quasi-infinite population. This assumption leads to overestimating the variance of the treatment effect when the population is small. While the finite-population correction factor is often used for prevalence estimation and in survey context, its use in clinical trials has been limited. However, by making few assumptions and assuming that the inference drawn in clinical trials is subject to sampling theory, the use of the finite-population correction factor can improve trials efficiency and show that the power of those trials is higher than it appears.

Methods: The finite-population correction factor is defined by (N – n)/(N – 1), where N is the population
size, and \( n \) is the sample size. A large simulation study was conducted to (1) assess the standard error of the mean treatment effect and coverage of the 95% confidence interval with and without the correction factor and (2) assess the power of a test comparing the treatment effect between two treatment groups. This covers the framework of both phase 2 and phase 3 trials. The impact of using the correction factor is assessed for varying treatment effect (25%–70%) and sample size \((n = 20–150)\). The impact on the calculation of the needed sample size is also investigated.

**Results:** The first simulation results confirmed the overestimation of the standard error of the mean with the naïve estimator, while the corrected standard error of the mean is closer to the true one. Depending on the scenario, the gain in power reached 14.1%, 8.7%, and 3.4% to detect a difference in treatment effect between two groups of 10%, 20%, and 25%, respectively. The gain increased with the sample size; it was negligible for \( n = 20 \), and in scenarios where the power is already very high. Additional scenarios are shown in the figure, for a difference in treatment effect of 20% (25% vs 45%). This gain in power translated into a decrease in sample size that is directly related to the proportion of the population needed: if the naïve calculation leads to a sample size of 10% the population size, then the sample size can be decreased by 9%; if the naïve calculation leads to a sample size of 50% the population size, then the sample size can be decreased by 33%, in order to reach the planned type I error and power.

**Conclusion:** This study shows that when dealing with rare diseases such as pediatric cancers, the power of clinical trials is higher than it appears with naïve estimates, and the planned level of confidence when estimating treatment efficacy can be reached with fewer patients. The gain in efficiency was seen with reasonable sample sizes and treatment differences, showing it can be useful in clinical research conducted in pediatric cancers, when the population size is approximately known.

---

**CP-64**

*(HOW) ARE CLAIMS OF PRAGMATISM JUSTIFIED IN PRIMARY TRIAL REPORTS? RESULTS FROM A REVIEW OF 415 SELF-DECLARED PRAGMATIC TRIALS*

**Primary Author:**
Monica Taljaard

**Co-Author(s):**
STUART G NICHOLLS
Ottawa Hospital Research Institute

ALISON HOWIE
Western University

**Abstracts**

**HAYDEN NIX**
Western University

**KELLY CARROLL**
Ottawa Hospital Research Institute

**PAXTON M MOON**
Western University

**NATALIE M NIGHTINGALE**
Western University

**BRUNO GIRAudeau**
University of Tours

**JAMIE C BREHAUT**
Ottawa Hospital Research Institute

**SPENCER P HEY**
Harvard Medical School

**Background:** Randomized trials designed with a pragmatic attitude are increasingly valued by funders, patients, and health system stakeholders. The pragmatic label may, however, be used rhetorically or without justification. The PRECIS (PRagmatic Explanatory Continuum Indicator Summary)-2 tool identifies nine design domains associated with pragmatism: eligibility, recruitment, setting, organization, flexibility-delivery, flexibility-adherence, follow-up, primary outcome, and primary analysis. Other trial features may also be associated with pragmatism. Within a large sample of self-declared pragmatic trials, we sought to describe: (1) prevalence of pragmatic features; (2) whether authors provided a rationale for the choice of pragmatic (as opposed to explanatory) trial; (3) whether and how authors justified claims of pragmatic design.

**Methods:** We used a Medline search filter to identify primary reports of trials likely to be pragmatic published 2014–2019, and focused on the subset that were registered in ClinicalTrials.gov. We identified all reports that explicitly characterized themselves as “pragmatic” anywhere in the text. We extracted trial characteristics from the report and ClinicalTrials.gov. We extracted any provided rationale for the choice of pragmatic trial. We identified all occurrences of the term “pragmatic” and inspected the surrounding text for relevant statements that seem to support claims of pragmatism. Statements were coded into the following categories: reference to any of the PRECIS-2 domains, the trial “architecture” (e.g. cluster randomization), type of intervention or comparator, lack of placebo and/or blinding, ethical approach (e.g. recruitment without consent), engagement with stakeholders, use of routinely collected data, cost-efficiency, large sample size or multiple centers, and other/unclear. Reference to a PRECIS wheel was noted.

**Results:** We identified 415 eligible trials; primary purpose was commonly treatment (38%), health services
research (23%), or prevention (20%); interventions were commonly patient educational or behavioral (34%), drugs (13%), procedure/surgery (9%), or device (8.4%). Trials were typically multi-center (77%) and randomized a median (Q1–Q3) of 523 participants (250–1876). The control arm was usual care in 59% of trials, and active comparator in 25%; 40% used no blinding; 19% used exclusively routinely collected data; 36% had patient-reported primary outcomes; 9% reported patient engagement and 12% other stakeholder engagement; 4% indicated some aspect of the trial was “minimal risk.” A rationale for choice of a pragmatic trial was provided in 27%. The claim that the design was pragmatic was supported by reference to at least one trial feature in 73%: 58% referenced one or more PRECIS-2 domains, 12% the type of intervention or comparator, 5.3% enrollment without consent, 5.3% the use of routinely collected data, 4.6% lack of placebo and/or blinding, and 4.1% sample size. Almost 10% made a broad statement (e.g. reference to generalizability) without providing details. Of those referencing PRECIS-2, 47% referenced only one, 30% two, and 14% three domains. The most commonly referenced PRECIS-2 domains were eligibility, setting, and flexibility-delivery. Only nine trials (2%) provided a PRECIS wheel. “Pragmatic” was used on average 3.3 times (range = 1–18).

Conclusion: “Pragmatic” is an attractive label which is often not justified or based on limited design features. Inclusion of a PRECIS-2 table or wheel can help make claims of pragmatism transparent.

CP-65
NUMBER OF STRATA IN CLINICAL TRIAL DESIGN AND ITS IMPACT ON TREATMENT IMBALANCE

Primary Author: NEAL L ODEN
The Emmes Company

Co-Author(s): PAUL VAN VELDHUISEN
The Emmes Company

Determining the number of factors to stratify on is often debated when designing clinical trials. In a recent discussion on clinical trial design, it was suggested that extra strata added to a design could be considered “free” in the sense that they cannot hurt a permuted block design because they can be ignored in the primary analysis, and may help in later secondary investigations. This is false: we show by simple calculation that extra strata and longer block sizes increase expected absolute imbalance in marginal (i.e. top level) strata across various sample sizes. We propose a novel measure of the effect of imbalance (worst power loss due to imbalance) and investigate by simulation the effect upon this metric of the number of strata, block length, and sample size in permuted block, Pocock–Simon minimization, and simple randomization designs. Findings from this investigation may provide guidance on a reasonable number of strata to consider when designing a clinical trial.

CP-66
DYNAMIC SCORING AND GUIDED DATA ENTRY IN A TELEHEALTH COGNITIVE BATTERY

Primary Author: EMILY H RIVES
Wake Forest School of Medicine

Co-Author(s): SARAH C BRUSCHI
Wake Forest School of Medicine
SARAH A GAUSSOIN
Wake Forest School of Medicine
LETITIA H PERDUE
Wake Forest School of Medicine

The PRagmatic EVaLuation of evEN Ts And Benefits of Lipid lowering in oldEr adults (PREVENTABLE) Trial is a double-blind, randomized, multi-site clinical trial. PREVENTABLE will randomly assign 20,000 adults ≥75 years of age without dementia, significant disability, or clinically evident cardiovascular disease in a 1:1 allocation to daily atorvastatin 40 mg or placebo. Participants are followed for up to 5 years. Endpoints are the new occurrence of disability, Mild Cognitive Impairment/dementia, and death. Wake Forest School of Medicine, Division of Public Health Sciences serves as the Data Coordinating Center.

Baseline and annual telephone-based cognitive assessments are conducted by trained data collectors as part of a central Call Center at Wake Forest. Complete baseline cognitive assessments are necessary for randomization. All participants are administered the Modified Telephone Interview for Cognitive Status, PROMIS-PF, and Katz ADL.

An innovative call tracking system was developed by the Data Coordinating Center. The system generates participant-specific intro and outro caller scripts comprised of information obtained from the clinic-completed forms. The intro script page includes information such as participant’s full name, preferred name, clinic site, phone number, and age. The ability to dynamically score the Modified Telephone Interview for Cognitive Status was developed as part of the extensive system. Validated cut-points, using the Modified Telephone Interview for Cognitive Status score, determine if the participant goes on to complete seven additional cognitive assessments. The guided data entry
system automatically generates those additional measures if needed, which allows the data collector to directly enter data and easily navigate through the additional tests in real-time with the participant. This minimizes participant burden and allows for the battery to be administered accurately and efficiently. Intro and outro scripts generate in either English or Spanish, depending on which is the participant’s preferred language. This allows the cognitive phone battery experience to be more personalized for the participant.

For participants scoring below the Modified Telephone Interview for Cognitive Status cut-point, the Call Center administers additional telephone-based assessments to a trusted contact. The trusted contact is a friend or family member named by the participant during a prior clinic visit. The outro script for the cognitive assessment generates the names and phone numbers from the clinic-completed Contact Information form for up to two trusted contacts. This allows the data collectors to immediately verify, and update, if needed, contact information for the trusted contact(s).

Dynamic scoring within a guided data-entry call tracking system ensures that scoring is accurate and efficient, which is beneficial for decreasing human scoring errors and necessary for real-time collection of needed additional data. Specifically, tailored scripts not only improve the experience for the participant by making it personalized, but also for the data collector. This makes their workflow more efficient, so they are able to focus on the call recipient and accurate data entry. This presentation will expand upon the guided data entry flow and showcase dynamic scoring and scripts as well as explore the advantages and limitations of the system.

As clinical trials are becoming more complex and multi-site trials more popular, there is a growing need for efficiently monitoring the performance of these trials to ensure data quality and trial success. With the growing trend toward global research and the need to coordinate and collect data across multiple sites, it has become more challenging to guarantee the integrity of trial data and results while keeping costs under control. Traditional on-site monitoring methods may not be the best approach for examining trial conduct as these extensive visits involving source data verification lead to higher costs. There is little improvement in data quality, and they may not evaluate 100% of the patient data entered at a site. Risk-based monitoring evaluates site performance and the risk to a trial represented by each site by tracking data on pre-identified, high-risk trial parameters collected from the respective sites. Risk-based monitoring further determines where monitoring efforts would optimally be placed and identifies sites with the best potential to deliver the greatest benefit to a trial. While on-site monitoring takes a “more is better” frequency-based approach, risk-based monitoring relies largely on a few key pre-defined risk indicators and thresholds. It can potentially reveal deviations in trial conduct and can identify poor-performing sites, but risk-based monitoring does not detect all data and program or protocol-related issues. More recently, there have been multiple initiatives underway to promote risk-based monitoring paradigms such as building Quality by Design into trials and performing a cross-functional risk assessment of program and protocol-level risks and monitoring these risks in order to improve the quality and efficiency of clinical trials.

Centralized statistical monitoring can aid in these efforts by assessing all data using a more holistic approach, and testing for anomalies in the values of variables that can be used to detect site variation, as well as monitor program or protocol-specific risks having the potential to affect trial conduct. Centralized statistical monitoring is commonly used to ensure data quality by detection of data entry/transcription issues and potentially fraudulent data before significant problems occur that would bias study findings and lead to invalid study interpretation but can also be applied to monitor protocol-level risks and program-wide concerns. Furthermore, through off-site centralized statistical monitoring, on-site monitoring can be more efficiently targeted, saving time, costs, and other resources. Consequently, we aimed to develop an approach for centralized statistical monitoring of our multi-site clinical trials to monitor various aspects related to the data quality and performance of a clinical trial. A user-friendly program for the application of centralized statistical monitoring was developed and tested against one of our trials. Tests for differences in the distributions of continuous and categorical variables, number of outliers and inliers, variability, etc.
intraclass correlation, internal consistency, and digit preference were performed. We summarize the methods underlying our application of centralized statistical monitoring as a complementary effort to risk-based monitoring and show how to implement centralized statistical monitoring and interpret the results.

**CP-68**

**ARE ELECTRONIC STUDY MANAGEMENT SYSTEMS ACCEPTABLE TO PARTICIPATING SITES?**

**Primary Author:**
MADELEINE CLOUT
University of Bristol

**Co-Author(s):**
HANA TABUSA
St George’s, University of London

**Introduction:** At the Bristol Trials Centre, we have created an electronic study management system which is currently being used in three studies. The reasons for using this system are to (1) minimize storage requirements, (2) allow remote access by the coordinating center and the study teams, and (3) reduce paper use in the study. The system includes the delegation log and investigator site files. We assessed user acceptability of the electronic systems.

**Methods:** We created an online survey and sent it to sites participating in these three studies (CIPHER HTA Ref:14/166/01, MONARCH HTA Ref:15/97/02, and SUNFLOWER HTA Ref:16/142/04). We asked about the difficulty of using the delegation log and investigator site files, preferences between electronic and paper systems and reasons for these preferences.

**Results:** There were 73 responses from 51 participating sites across the three studies. Responders included research nurses (48%), members of the administrative team (23%), Principal Investigators (19%), clinicians (7%), and specialist nurses (3%). The electronic delegation log was used by 84% of responders. Of these, 87% found the system either easy or quite easy to use. This included registering for an account, making changes and viewing their site delegation log. The survey showed that 42% expressed a preference for electronic delegation logs, 32% for paper, and 26% had no preference. The main reasons for the preference for electronic delegation logs were accessibility (48%) and ease of use (36%). The investigator site file was used by 66% of responders. Of these, 81% found the system either easy or quite easy to use. This included finding, downloading, and uploading documents to the site file. The survey showed that 59% expressed a preference for electronic site files, 24% for paper, and 17% had no preference. The main reasons for the preference for electronic site files were ease of use (40%) and accessibility (36%). Overall, 80% of responders found the online study management system easy or quite easy to navigate. Across all surveyed staff groups (research nurses, members of the administrative team, Principal Investigators, clinicians and specialist nurses), more than half of responders either preferred the electronic format of the delegation log and site file or had no preference.

**Conclusion:** The survey showed that users found the electronic study management systems accessible and easy to use. This suggests that the online systems can be used by participating sites as an acceptable alternative to paper delegation logs and site files. The studies included in this survey were observational and interventional; however, further consideration should be given to the use of these systems in Clinical Trials of Investigational Medicinal Products. The Bristol Trials Centre has successfully bid for funding for development of a generic electronic study management system (MANGO). Work on the system has started and the final product will be made available to UK-based Clinical Trials Units at minimal cost.

**CP-69**

**A CORE OUTCOME SET FOR STANDARDIZING AND REPORTING SURGICAL INNOVATION: RESULTS OF AN INTERNATIONAL DELPHI SURVEY AND CONSENSUS MEETING**

**Primary Author:**
CHRISTIN HOFFMANN
University of Bristol

**Co-Author(s):**
KERRY AVERY
National Institute for Health Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research, University of Bristol

RHIANNON MACEFIELD
National Institute for Health Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research, University of Bristol

NICHOLAS WILSON
National Institute for Health Research Bristol Biomedical Research Centre, Bristol Centre for Surgical Research, University of Bristol

ANGUS MCNAIR
North Bristol NHS Trust

JANE BLAZEBY
Bristol Royal Infirmary, University Hospitals Bristol and Weston NHS Foundation Trust
Background: The introduction of new surgical procedures and devices is important for advancing healthcare and patient outcomes. Rigorous evaluation of surgical innovation from their early development through to randomized controlled trials, however, is hampered by unstandardized measurement and reporting of outcomes. New procedures and devices are commonly introduced without evidence-based practice due to a lack of regulation and oversight. Unstandardized outcome selection, measurement, and reporting can also lead to increased risks of outcome reporting bias and hinders data synthesis. Improved guidance is needed to prevent patient harm, facilitate shared learning, and avoid research waste. The aim of the COHESIVE study was to develop a core outcome set, an agreed minimum set of outcomes to report in studies of surgical innovations, to promote the safe and efficient evaluation of new procedures/devices.

Methods: A core outcome set for early phase surgical innovation and devices was developed according to established guidance (Core Outcome Measures in Effectiveness Trials and core outcome set-STAD). Outcomes identified from published studies, UK healthcare policies, regulatory documentation and interviews with surgeon-innovators and patients informed an international Delphi consensus survey. Patients and professionals, representing a wide range of specialties and professions, rated the importance of each outcome in two survey rounds conducted online. A multi-stakeholder international consensus meeting with anonymized voting agreed the final core outcome set.

Results: Seven thousand six hundred sixty six verbatim outcomes were categorized into 32 outcome domains. A total of 251 participants (148 professionals and 103 patients) across 15 international regions completed both survey rounds. A consensus meeting (28 participants; 8 patient representatives and 20 professionals) was convened in February 2020. Voting resulted in eight outcome domains to be included in the final core outcome set. Six innovation-domains included modifications, unexpected disadvantages, device problems, technical procedure success, whether the overall desired effect was achieved and surgeon experience. Two domains shared with effectiveness studies were expected disadvantages and intended benefits of surgical innovation.

Conclusion: The core outcome set is recommended for use in all studies of new procedures/devices prior to definite assessment in randomized clinical trials. Standardized outcome reporting will support rigorous assessment of surgical innovation and facilitate design of further surgical research comparing new and established invasive procedures.

Primary Author: SHELLEY POTTER
Bristol Breast Care Centre, North Bristol NHS Trust

Co-Author(s): MONICA TALJAARD
University of Ottawa
DEAN FERGUSSON
University of Ottawa
JONATHAN COOK
Oxford University

Background: Patient-reported outcomes are subjective measures of health and well-being that come directly from patients and commonly used to measure patient experience, quality of life, and symptoms. Pragmatic randomized controlled trials aim to provide evidence to inform real-world decisions. Patient-reported outcomes seem particularly well suited to pragmatic trials; however, practices surrounding their use and reporting in pragmatic trials have not been described. We sought to review pragmatic trials to describe (1) the prevalence and types of patient-reported outcomes used, (2) whether the use of patient-reported outcomes varied across trial characteristics, and (3) how sample sizes and target differences were determined for trials with patient-reported outcomes.

Methods: An electronic search filter in Medline was used to identify primary reports of pragmatic randomized controlled trials in health research published 2014–2019, that were registered in ClinicalTrials.gov and self-identified as pragmatic. Trial descriptors were downloaded from ClinicalTrials.gov; information about patient-reported outcomes and sample size calculations were extracted from each report. Data were summarized descriptively. The chi-square, Cochran–Armitage, and Wilcoxon rank-sum tests were used to examine associations with use of patient-reported outcomes.

Results: Four thousand three hundred thirty seven primary randomized clinical trial reports were identified, of which 415 trials were eligible for this review. Most (n = 263; 63%) were individually randomized (n = 152; 37% cluster randomized), half (n = 207; 50%) were conducted in North America, a third (n = 141; 34%) in Europe, and 11% (n = 45) in low- or middle-income countries. Primary purpose of trials was treatment (38%), health services research (23%), or prevention (20%); interventions were patient educational or behavioral (34%), drugs (13%), procedure/
surgery (9%), or device (8.4%); 9% reported patient engagement and 12% other stakeholder engagement. In the review sample, 235 (57%) measured patient-reported outcomes (n = 144; 35% as primary/co-primary and n = 91; 22% as only secondary outcomes). A rationale for the chosen patient-reported outcome was provided in 25 (16%): based on literature or explanation (n = 20; 14%), patient consultation (n = 3; 2%), or other stakeholder consultation (n = 1; 1%). Primary patient-reported outcomes were most commonly symptoms (n = 64; 44%), followed by health behaviors (n = 36; 25%), quality of life (n = 18; 13%), functional status (n = 16; 11%), and patient experience (n = 10; 7%). For the 144 trials with a patient-reported outcome as primary/co-primary outcome, 126 (88%) reported a sample size calculation for that outcome. Of these, 53 (42%) did not provide justification for the chosen target difference. Studies published in higher impact journals or funded by industry were less likely to use primary patient-reported outcomes, whereas individually (vs cluster) randomized studies, those conducted in Europe, and dietary or behavioral interventions were more likely to use patient-reported outcomes as primary/co-primary outcomes. For the 144 trials with a patient-reported outcome as primary/co-primary outcome, 126 (88%) reported a sample size calculation for that outcome. Of these, 53 (42%) did not provide justification for the chosen target difference. Studies published in higher impact journals or funded by industry were less likely to use primary patient-reported outcomes, whereas individually (vs cluster) randomized studies, those conducted in Europe, and dietary or behavioral interventions were more likely to use patient-reported outcomes as primary/co-primary outcomes.

Implications: Patient-reported outcomes are not routinely selected as primary or co-primary outcomes in pragmatic trials, and patient and stakeholder engagement in determining target differences and sample sizes is rare. There is room for improvement in designing pragmatic trials to better incorporate patient-centered outcomes.

**CP-71**

**DISCRIMINANT VALIDITY OF THE MISUSE, ABUSE, AND DIVERSION DRUG EVENT REPORTING SYSTEM IN ADULT AND PEDIATRIC PHASE 2 AND 3 CLINICAL TRIALS**

**Primary Author:**
KATHRYN EVANS  
WCG Analgesic Solutions

**Co-Author(s):**  
ANDREA MARRAFFINO  
WCG Analgesic Solutions  
NATHANIEL KATZ  
WCG Analgesic Solutions

The misuse, abuse, and diversion drug event reporting system (MADDERS®) system was developed to meet Food and Drug Administration recommendations for assessing abuse potential in safety and efficacy clinical trials of CNS-active drugs, and was found in a systematic review to be the only available system that accomplishes this purpose. Traditional methodologies such as retrospective queries of treatment-emergent side effects and adverse events as well as the use of patient-reported and clinician-reported instruments have limited reliability and can lead to the misclassification of potentially abuse-related events, resulting in an overestimation or underestimation of a drug’s true abuse potential. The objective of this study was to evaluate the ability of the system to discriminate between drugs with higher and lower potential for abuse, based on an analysis of completed clinical trials that employed MADDERS, and to determine if results are similar in the pediatric population. Data from 10 clinical trials from drug development programs investigating cannabinoid and opioid formulations were analyzed for the incidence of MADDERS events, types of events, and relative rates of abuse-related final classifications made by adjudicators. The trials included in the analysis were: two placebo-controlled trials of a cannabinoid formulation in adults and adolescents (<12 years) and two in a pediatric population (ages 3–17 years); four were trials of a delta-9-tetrahydrocannabinol and cannabidiol combination product in adults, and two were trials of a novel opioid oral formulation in adults. Data were pooled for all trials of each respective formulation to create separate treatment populations for comparison. The rates of total MADDERS events classified as abuse were 0 (0%), 0 (0%), and 5 (0.4%) in pooled studies of cannabidiol, tetrahydrocannabinol/cannabidiol, and an opioid, respectively. Both pediatric and adult populations in cannabinoid trials had 0 (0%) events of abuse. Drugs with higher expected abuse potential had a greater number of events classified as abuse and misuse than those with no known potential for abuse, demonstrating the discriminant validity of MADDERS for assessing abuse potential. Initial evaluation in the pediatric population suggests similar results in detecting abuse as in the adult populations. Further research is needed to determine if the same level of discriminant validity is seen in the pediatric population with drugs with higher expected abuse potential.

**CP-72**

**CONSIDERATIONS IN IMPLEMENTING THE ESTIMAND FRAMEWORK IN MULTI-SITE RANDOMIZED CLINICAL TRIALS**

**Primary Author:**
MIN ZHAN  
U.S. Department of Veterans Affairs

The misuse, abuse, and diversion drug event reporting system (MADDERS®) system was developed to meet
Randomized controlled trials are the gold standard for evaluating efficacy of a new treatment, therapy, or device. Randomization provides a mechanism to control confounding. One necessary condition to conduct a successful randomized clinical trial is to clearly specify the treatment of interest at the trial design stage. However, the protection provided by randomization does not apply to intercurrent (post-randomization) events that affect the measurements of the outcomes, yet clinical trials often provide insufficient details regarding how to handle the intercurrent events. In view of this, recently the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) published an addendum ICH-E9 (R1) advising an estimand framework to align planning, design, conduct, analysis, and interpretation of randomized clinical trials, where an estimand is a precise description of the treatment effect of interest and the target of estimation linking to a trial objective. A clear description of an estimand includes description of four attributes—the treatment condition of interest, the patient population, the variable (or endpoint), and a population-level summary. This estimand framework also requires strategies for dealing with intercurrent events. In this report, we review the estimand framework, and discuss strategies and challenges for handling different intercurrent events, including discontinuation of assigned treatment, rest period from assigned treatment, use of an additional treatment, and terminal events, using examples motivated by multi-site randomized clinical trials sponsored and conducted by Cooperative Studies Program, Office of Research and Development, U.S. Department of Veterans Affairs.

Co-Author(s):
ZHIBAO MI
U.S. Department of Veterans Affairs
KOUSICK BISWAS
U.S. Department of Veterans Affairs

Randomized controlled trials are the gold standard for evaluating efficacy of a new treatment, therapy, or device. Randomization provides a mechanism to control confounding. One necessary condition to conduct a successful randomized clinical trial is to clearly specify the treatment of interest at the trial design stage. However, the protection provided by randomization does not apply to intercurrent (post-randomization) events that affect the measurements of the outcomes, yet clinical trials often provide insufficient details regarding how to handle the intercurrent events. In view of this, recently the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) published an addendum ICH-E9 (R1) advising an estimand framework to align planning, design, conduct, analysis, and interpretation of randomized clinical trials, where an estimand is a precise description of the treatment effect of interest and the target of estimation linking to a trial objective. A clear description of an estimand includes description of four attributes—the treatment condition of interest, the patient population, the variable (or endpoint), and a population-level summary. This estimand framework also requires strategies for dealing with intercurrent events. In this report, we review the estimand framework, and discuss strategies and challenges for handling different intercurrent events, including discontinuation of assigned treatment, rest period from assigned treatment, use of an additional treatment, and terminal events, using examples motivated by multi-site randomized clinical trials sponsored and conducted by Cooperative Studies Program, Office of Research and Development, U.S. Department of Veterans Affairs.

Co-Author(s):
AHMED MAMANE ARZIKA
Centre de Recherche et Interventions en Sante Publique
ABDOU AMZA
Programme Nationale de Sante Oculaire
BOUBACAR KADRI
Programme Nationale de Sante Oculaire
RAMATOU MALIKI
Centre de Recherche et Interventions en Sante Publique
ALIO KARAMBA MANKARA
Centre de Recherche et Interventions en Sante Publique
ABDOUL NASER HAROUNA
Centre de Recherche et Interventions en Sante Publique
VICTORIA LE
University of California, San Francisco
TRAVIS C PORCO
University of California, San Francisco
THUY DOAN
University of California, San Francisco

The Macrolides Oraux pour Réduire les Décès avec un Oeil Sur la Résistance cluster-randomized trial demonstrated that biannual distribution of azithromycin to children 1–59 months of age reduced mortality by 14% in Malawi, Niger, and Tanzania (Macrolides Oraux pour Réduire les Décès avec un Oeil Sur la Résistance). In response, the World Health Organization suggested targeting this intervention to the subgroup of children 1–11 months of age in high mortality settings in order to reduce potential selection for antimicrobial resistance. Azithromycine pour la Vie des Enfants au Niger: Implementation et Recherche is a double-masked cluster-randomized placebo-controlled adaptive large simple trial in Niger designed to determine the effect of age-based targeting of biannual azithromycin distribution on mortality and antimicrobial resistance (Azithromycine pour la Vie des Enfants au Niger: Implementation et Recherche). The use of a large simple trial design ensures adequate power to detect a modest intervention effect on a rare outcome like mortality while ensuring feasibility through limited data collection, a simple intervention, and simple outcome monitoring. Response-adaptive randomization allows for ethical allocation of interventions by increasing the probability that communities are randomized to the most effective strategy over time. In addition, this approach enables the study to adapt to changing guidelines by randomizing in new intervention arms, for example, rather than being constrained by the original design.

The target sample size is 3350 rural and peri-urban communities from the Dosso, Tahoua, and Maradi

Primary Author:
KIERAN SUNANDA O’BRIEN
University of California, San Francisco
regions. Eligible communities in the Dosso region will be enrolled in the first year and randomized 1:1:1 to (1) azithromycin 1-11: biannual azithromycin to children 1–11 months old with placebo to children 12–59 months old, (2) azithromycin 1-59: biannual azithromycin to children 1–59 months old, or (3) placebo: biannual placebo to children 1–59 months old. After a 1-year run-in period, communities in subsequent regions will be randomized with an updated allocation based on the probability of mortality in children 1–59 months in each arm during the preceding study period. In all regions, communities will retain their allocation for four distributions before being re-randomized with updated allocations. Trained study teams will conduct a biannual door-to-door census to enumerate the population, distribute azithromycin and placebo, and monitor vital status. Primary mortality outcomes will be assessed after 2 years of distributions and are defined as mortality rate (deaths per 1000 person-years) in (1) children 1–59 months old comparing the azithromycin 1-59 and placebo arms, (2) children 1–11 months old comparing the azithromycin 1-11 and placebo arm, and (3) children 12–59 months in the azithromycin 1-11 and azithromycin 1-59 arms. From each arm, 50 communities in the Dosso region will be followed to monitor antimicrobial resistance. Primary resistance outcomes will be assessed after 2 years of distributions and include (1) prevalence of genetic determinants of macrolide resistance in nasopharyngeal samples from children 1–59 months old, and (2) load of genetic determinants of macrolide resistance in rectal samples from children 1–59 months old.

CP-74
IMPACT OF UNEQUAL CLUSTER SIZES FOR GENERALIZED ESTIMATING EQUATION ANALYSES OF STEPPED-WEDGE DESIGNS WITH BINARY OUTCOMES

Primary Author:
ZIBO TIAN
Yale School of Public Health

Co-Author(s):
PAUL J RATHOUZ
The University of Texas at Austin
DENISE ESSERMAN
Yale School of Public Health
ELIZABETH L TURNER
Duke University
JOHN S PREISSER
University of North Carolina
FAN LI
Yale School of Public Health

Stepped-wedge design is a type of unidirectional crossover design where different units switch from control to intervention condition at different pre-specified time points. While a convention in study planning is to assume the cluster-period sizes are identical, stepped-wedge cluster-randomized trials involving repeated cross-sectional designs frequently have unequal cluster-period sizes, which can impact the efficiency of the treatment effect estimator. In this article, we provide a comprehensive investigation of the efficiency impact of unequal cluster sizes for generalized estimating equation analyses of stepped-wedge cluster-randomized trials, with a focus on binary outcomes as in the Washington State EPT trial. Several major distinctions between our work and existing work include (1) we consider multi-level correlation structures appropriate for binary outcomes; (2) we study the implications of both the between-cluster and within-cluster imbalances in sizes; and (3) we provide a comparison between the independence working correlation versus the true working correlation and detail the consequence of ignoring correlation estimation in stepped-wedge cluster-randomized trials with unequal cluster sizes. We conclude that the working independence assumption can lead to substantial efficiency loss and a large sample size regardless of cluster-period size variability in stepped-wedge cluster-randomized trials, and recommend accounting for correlations in the analysis. To improve study planning, we additionally provide a computationally efficient search algorithm to estimate the sample size in stepped-wedge cluster-randomized trials accounting for unequal cluster-period sizes and conclude by illustrating the proposed approach in the context of the Washington State EPT study.

CP-75
IMPROVING EFFICIENCY BY EMBEDDING TRIALS INTO HEALTHCARE SETTINGS

Primary Author:
LINDSAY S KEHOE
Clinical Trials Transformation Initiative

Co-Author(s):
SUANNA BRUGIONNE
American Society of Clinical Oncology ASCO

Randomized controlled trials are a powerful way to evaluate the effectiveness of new or established therapies. However, randomized controlled trials traditionally do not collect data through integration with clinical care. This leads to inefficiency through duplicative data collection, increased trial expense, and results that may not reflect the real-world performance of medical products in the populations that will use them. Current clinical trials like RECOVERY and I-SPY COVID have highlighted that the integration of research components
into clinical care is not only possible, but has the potential to improve trial efficiency and broaden access to new therapies.

The presenters will review Clinical Trials Transformation Initiatives that have paved the way toward embedding trials into clinical care settings. The presenters will also provide examples of relevant American Society of Clinical Oncology trials and current initiatives as a means of discussing elements to consider when integrating trials. Attendees will leave the presentation with a better appreciation of the issues and opportunities that exist with embedding trials into clinical care and a consideration of how embedded trials can utilize real-world data sources while maintaining scientific rigor.

The presenter will be Suanna Bruinooge, MPH. Suanna is the Division Director of Research Strategy and Operations in American Society of Clinical Oncology’s Center for Research and Analytics (CENTRA). CENTRA generates, integrates, analyzes, and shares oncology data to foster innovation in research and patient care. Ms Bruinooge also works with American Society of Clinical Oncology members and staff colleagues to develop and evaluate American Society of Clinical Oncology’s positions related to research design, conduct, and reporting. CENTRA develops and implements American Society of Clinical Oncology’s research projects, including the Targeted Agent Profiling and Utilization clinical trial, the American Society of Clinical Oncology Survey on COVID-19 in Oncology Registry, and projects to advance clinical trial design and methodology. CENTRA also staffs American Society of Clinical Oncology’s Cancer Research Committee and Research Community Forum.

Prior to joining American Society of Clinical Oncology, Suanna worked for seven and a half years in the US House of Representatives, working for Congresswoman Nancy Johnson (R-CT) and Congressman Vernon Ehlers (R-MI). Ms Bruinooge earned a Master’s of Public Health in Health Policy at The George Washington University’s Milken Institute School of Public Health in 2015. Suanna also has a BA in political science from Calvin College in Grand Rapids, MI, USA.

Lindsay Kehoe manages the development and implementation of projects at the Clinical Trials Transformation Initiative or (CTTI). Lindsay has close to 10 years of experience coordinating clinical research. Prior to joining Clinical Trials Transformation Initiative, she was a rare disease specialist in collaboration with the National Institutes of Health and served as a clinical trial lead at Children’s National Medical Center, coordinating both physician-driven and industry-sponsored trials. Lindsay is a certified genetic counselor with extensive experience in both pediatric and adult genetic counseling. Early in her career, Lindsay worked in early phase drug development at Millennium Pharmaceuticals (now Takeda Oncology) and then in post-marketing laboratory surveillance at Sanofi Genzyme. Lindsay has a BA from the University of Virginia and an MS in Genetic Counseling from the Boston University School of Medicine.

CP-76
THE IMPACT OF COVID-19 RESTRICTIONS ON PARTICIPANT ENROLLMENT IN THE PREPARE TRIAL
Primary Author:
SHEILA SPRAGUE
McMaster University
Co-Author(s):
DAVID POGORZELSKI
McMaster University
BECKY SNIDER
Prisma Health
MARKO GJORGJEVSKI
McMaster University
JODI GALLANT
McMaster University
DEVIN MULLIN
Dartmouth-Hitchcock Medical Center
ELIZABETH ALLEN
Harvard University
ASHLEY IWU
University of Pennsylvania Health System
JAMES AHN
Inova Fairfax Hospital
DMITRY POKHVASHCHEV
University of California, San Francisco

Introduction: PREPARE is a pragmatic cluster-randomized crossover trial that compares the effectiveness of two common pre-operative antiseptic skin solutions to reduce the risk of surgical site infection after orthopedic fracture surgery. The trial compares 2% chlorhexidine in 70% isopropyl alcohol (ChloraPrep™) versus 0.7% iodine povacrylex in 74% isopropyl alcohol (DuraPrep™), and recruiting sites alternate study solutions every 2 months. Before the US national response to the COVID-19 pandemic, all PREPARE trial clinical sites obtained informed consent in person at the hospitals or fracture clinics. However, after 13 March 2020, the COVID-19 restrictions limited in-person consenting at some hospitals. Affected clinical sites were encouraged to transition to telephone consenting, which was already included as a consent option in the PREPARE protocol. We aimed to determine how COVID-19 restrictions impacted the number of
enrolling clinical sites and participant enrollment in the PREPARE trial.

Methods: Prior to implementing telephone consent, clinical sites had to determine local logistics and obtain institutional review board approval for telephone consent scripts and procedures from the central or local institutional review boards. We descriptively evaluated the number of clinical sites that switched to telephone consent, the number of clinical sites that had to pause enrollment, and the length of the enrollment pauses. We evaluated monthly enrollment at the following time periods: (1) prior to the COVID-19 restrictions (1 July 2019 to 13 March 2020), (2) immediately after the COVID-19 restrictions in place (March, April, and May 2020), and (3) from 1 June 2020 to 30 November 2020. Results are stratified by open and closed fractures and are summarized using descriptive statistics.

Results: At the time of the pandemic, 13 clinical sites were participating in the PREPARE trial. Eleven (84.6%) clinical sites paused enrollment due to COVID-19 restrictions. The median length of enrollment pause was 44 days (range = 7–92 days; interquartile range = 54 days). By 16 June 2020, all clinical sites resumed enrollment. The average monthly enrollment before the COVID-19 restrictions was 198 closed fracture participants (SD = 22, range = 161–227) and 41 open fracture participants (SD = 16, range = 22–60). The enrollment rate was the lowest in April 2020, when 47 closed fracture participants and nine open fracture participants were enrolled. By June 2020, enrollment began increasing. From 1 June 2020 to 30 November 2020, the average monthly enrollment rate was 183 closed fracture participants (SD = 30, range = 129–206) and 44 open fracture participants (SD = 12, range = 24–61), which was close to pre-COVID enrollment. Monthly enrollment between 2019 and 2020 was similar, except for the months of March to May 2020 for closed fracture enrollment (p = 0.001) and March and April 2020 for open fracture enrollment (p = 0.04) cohorts.

Conclusion: By pre-emptively including telephone consent in the PREPARE protocol, clinical sites were quickly and ethically adapting their procedures for obtaining informed consent via telephone. Although multiple sites paused enrollment, the enrollment pause was brief and had minimal impact on enrollment. A highly pragmatic design allowed for minimal interruptions to enrollment during the pandemic.

CP-77

ACTIONS OF A NATIONAL ONCOLOGY GROUP TO MANAGE THE IMPACT OF THE COVID-19 PANDEMIC

Primary Author: PHYLLIS J GOODMAN
SWOG Statistics and Data Management Center

Co-Author(s):
CATHRYN RANKIN
SWOG Statistics and Data Management Center
WENDY LAWTON
SWOG Group Chair’s Office
DANA B SPARKS
SWOG Operations Office
MONICA YEE
SWOG Statistics and Data Management Center
RODNEY SUTTER
SWOG Statistics and Data Management Center
ELAINE ARMSTRONG
SWOG Operations Office
DANI WEATHERBEE
SWOG Statistics and Data Management Center

SWOG Cancer Research Network, part of the National Cancer Institute’s National Clinical Trials Network and the National Cancer Institute Community Oncology Research Program, designs and conducts clinical trials to improve the lives of people with cancer. Trial designs are Phase 2, 2/3 and 3 as well as observational cohorts and focus on cancer treatment or on supportive care, symptom management, cancer care delivery, or cancer control. Approximately 4000 patients are put on SWOG trials annually by study sites across all 50 states as well as by international sites; currently, there are 90 active trials.

The COVID-19 pandemic forced SWOG to pivot to address the large range of issues facing the study sites that put patients on its trials, as well as required the group to track the impact of the pandemic on its trials. Information was flowing from a variety of sources and we needed to quickly respond, as well as efficiently distribute the information to study sites, who themselves were burdened with pandemic-related staffing and process issues.

National Cancer Institute guidance to sites addressed concerns regarding patient safety and minimizing patient risk. Procedures for obtaining informed consent remotely, the distribution of study drug, and the use of telehealth visits were among the issues addressed. SWOG provided protocol-specific guidance regarding timing of specified activities including allowable registration windows, timing of labs, and collection of specimens and patient-reported outcomes. SWOG created new reports to monitor accrual, developed, and deployed a form to collect National Cancer Institute-mandated COVID-specific data, provided study sites with information on the specimen repositories ability to receive and process biological trial samples, as well as a prioritized, study-specific list of specimen requirements. Other changes included a switch to remote Quality Assurance audits and adjustments to how sites were
monitored regarding submission of data and specimens. New internal data management processes for the remote work environment were implemented.

Because of the volume of information, SWOG rapidly launched a clearinghouse on SWOG’s website (swog.org) to centrally compile real-time COVID-19 resources, news, and information, including the general guidance documents from SWOG, the National Cancer Institute, its Central institutional review board, and the Food and Drug Administration. Protocol-specific memoranda were issued for 22 SWOG trials.

The semi-annual SWOG Group Meeting scheduled for April 2020 was switched to a remote meeting as were the subsequent two meetings and other trainings and seminars. Even with all the changes, 14 new trials were activated in 2020 and 107 journal articles published. Accrual was impacted differentially by disease area, study phase, and trial type.

SWOG staff across multiple offices worked together to modify and adjust processes so that site staff could continue to offer the highest quality clinical trials and care for cancer patients amid a public health emergency. In our presentation, we will discuss lessons learned and the impact of the pandemic on the future of clinical trials.

**Funding**

This study was supported by National Cancer Institute grant nos U10CA180888, U10CA180819, and UG1CA189974.

**CP-78**

PLATFORM TRIALS TAKE THE SPOTLIGHT IN COVID-19

**Primary Author:**

ALYSSA M VANDERBEEK

*Institute of Cancer Research*

**Co-Author(s):**

ZHULIN YIN

*The Institute of Cancer Research*

JUDITH M BLISS

*The Institute of Cancer Research*

CHRISTINA YAP

*The Institute of Cancer Research*

**Background:** Platform designs—master protocols that allow for new treatment arms to be added over time—have not previously been widely used outside of oncology. According to a review by Park et al. (2019), 83 master protocol trials were initiated between 2001 and 2019, of which 76 were in cancer. However, COVID-19 trials RECOVERY and Solidarity have effectively adopted this innovative design, demonstrating that platform designs have a place in non-oncology settings. The addition of these trials to the platform trial landscape will provide needed insight to the design, analysis, and conduct of platform trials. We present a rapid review to describe the implementation of platform trials in COVID-19.

**Methods:** We conducted searches in PubMed, ClinicalTrials.gov, and the Cytel COVID-19 Clinical Trials Tracker between 21 October and 4 November 2020. Platform trials were defined by their self-identification and/or the flexibility to add future arms, as explicitly stated or assumed in the trial registration page, trial website, or official study documents.

**Results:** Forty-five platform trials in COVID-19 were registered in the first 10 months of 2020, marking a substantial increase in the use of this design. Twenty-three trials (51%) have publicly shared their protocol, and three provide full statistical analysis plans. Fifteen trials (33%) have committed to sharing individual patient data. Fifteen (33%) clearly state a Bayesian approach. Forty trials (88%) incorporate adaptive features; the three most common are futility stopping (24, 53%), early efficacy stopping (19, 42%), and sample size reassessment (13, 28%). The adaptive features used in nine trials (20%) are unclear, either due to ambiguous language or insufficient information.

**Conclusion:** Although catastrophic for the world, COVID-19 has accelerated uptake of this innovative design. Platform designs have been efficiently implemented, and adaptive features and Bayesian methods are being used often—a deviation from conventional frequentist approaches and traditional fixed sample designs—and an encouraging number of trials have committed to share individual patient data. Challenges remain, but a significant barrier of using such complex designs has been broken that will greatly inform future clinical trial design, conduct, and regulatory proceedings.

**IN-CONFERENCE WORKSHOPS**

**WK-1**

BIOMARKER-DRIVEN TRIALS FOR PRECISION MEDICINE

**Organizer:**

SUSAN HALABI

*Duke University*

**Contributor(s):**

MARY REDMAN

*Fred Hutchinson Cancer Research Center*

Innovations in molecular and genetic laboratory assays have advanced our understanding of cancer biology by allowing identification of specific molecular and cellular characteristics of tumors (biomarkers). These advances in the understanding of cancer biology have additionally led to therapeutic advances and importantly advances in the type of trial designs used to
evaluate these “biomarker-driven” therapies in possibly biomarker-defined populations. Many investigators and statisticians are exposed to such trials without a deep understanding of the scientific and practical challenges in the design and conduct of biomarker-driven trials. Furthermore, it is evident now in designing contemporary trials with biomarkers that the one-size fits all approach does not work as the design heavily depends on the level of evidence of the biomarker. Even with careful planning, there can be logistical and operational challenges in the conduct of such studies and scientific challenges in the interpretation of the data generated from these trials. The main goal of the workshop is to discuss basic principles in the choice of design and to introduce participants to designs that include biomarkers to define the primary study population, or as a subset of the primary population, or to potentially identify a biomarker-defined population. This half-day workshop is offered to all statisticians and clinicians who are motivated to learn on how to plan and design of biomarker clinical trial. Clinicians with a basic understanding of hypothesis testing and multiple testing issues would benefit from this workshop. We will start by making a distinction between prognostic and predictive biomarkers, and then discuss the basic designs standard randomized controlled trial with biomarker analysis as secondary, targeted design, hybrid design (including biomarkers pre-defined and biomarker discovery), and strategy design. We will then examine some of the considerations for biomarker choice and inclusion (binary, categorical, and continuous) and then delve into master protocols (basket vs umbrella)—ongoing versus single use and discuss discovery-based versus confirmatory master protocols. The concepts will be demonstrated using real-life examples from successfully designed cancer trials. In addition, two cases studies will be presented in two breakout rooms with audience participation. Each instructor will lead the discussion on one of the case studies. Taught by statisticians who have extensive experience in biomarker-driven trials, including experience with stand-alone biomarker-based trials and biomarker-driven trials conducted within a biomarker-driven master protocol, participants will learn the best practices for designing enriched trials, stratified biomarker design, basket and umbrella trials among others. Here are our thoughts: (1) Introduction: 5 min (Halabi, Redman), (2) Lecture Prognostic, Predictors, Biomarker-driven designs: 30 min (Halabi), (3) Case study 1: Breakout; 20 min (Halabi, Redman), (4) 5-min break, (5) Lecture Master Protocols, Umbrella & Basket: 25 min (Redman), (6) Case study 2 Breakout session: 20 min (Halabi, Redman), and (7) Summary: 15 min (Halabi, Redman).

Goals of Tutorial: This workshop will provide a basic exposition of the methods involved in the design of biomarker-driven trials. At the end of the workshop, participants will: understand the different types of biomarkers, be able to choose the optimal design for different biomarkers, consider which population to study, identify which endpoint is reasonable, consider different schemes on how allocate the type 1 error rate and learn how to analyze and best present the results.

WK-2

DESIGN CLINICAL TRIALS WITH MULTIPLE ENDPOINTS: METHODS AND APPLICATIONS

Organizer:

SHENG S LUO

Duke University

Contributor(s):

YING YUAN

The University of Texas MD Anderson Cancer Center

Description of Tutorial: Considering multiple endpoints is an effective approach to address the intrinsic high uncertainty associated with drug development and maximize the success rate of drug development. An increasing number of pivotal trials have adopted this approach, including Impassion 130, IMbrave 150, and IMpower 133, resulting in drugs successfully approval by the Food and Drug Administration. Using multiple endpoints brings many design and statistical challenges, for example, the regulatory concern the multiplicity issues caused by testing multiple endpoints. This short course introduces the state-of-the-art methods to design clinical trials with multiple endpoints, with compliance to the regulatory guidance. The application of the methods will be illustrated using the pivotal trials. The sample size determination for multiple endpoints will be briefly illustrated using the available software (e.g. PASS).

Goals of Tutorial: The core goal is to introduce state-of-the-art methods to design clinical trials with multiple endpoints and their applications using high-profile pivotal trials (e.g. Impassion 130, IMbrave 150, and IMpower 133) published in The New England Journal of Medicine. By the end of the session, participants will be able to: (1) understand the features and challenges of multiple endpoints; (2) master statistical methods to address the challenges of multiple endpoints; and (3) apply the introduced methods to design phase 2 and 3 trials with multiple endpoints.

WK-3

DESIGNING CLINICAL RESEARCH TO INFORM DECISION-MAKING USING VALUE OF INFORMATION

Organizer:

ANNA HEATH

The Hospital for Sick Children
Contributor(s):

NATALIA KUNST  
Harvard Medical School and Harvard Pilgrim Health Care Institute

HAWRE JALAL  
University of Pittsburgh

NICKY WELTON  
University of Bristol

Objectives: This workshop introduces how four Value of Information measures can be used to determine research priorities and trial designs that support decision-making. This workshop will present the assumptions of a Value of Information analysis and methods to display results graphically.

Description: Clinical research should aim to support decision-making around the best treatment for use in clinical practice. Value of Information is a concept that determines what research would have the greatest potential to reduce uncertainty for decision-makers in healthcare. Specifically, a Value of Information analysis can be used to: (1) inform whether research is needed in a specific clinical area, (2) select what outcomes should be included in a trial by determining what outcomes would support decision-making, and (3) determine the optimal design for a research study, including sample size considerations, to reduce decision uncertainty as efficiently as possible. Despite this versatility, Value of Information has rarely been used in practice due to a lack of familiarity, computational challenges, difficulties interpreting these measures, and concerns about the assumptions underpinning them. Thus, the Collaborative Network on Value of Information Group was formed to improve the understanding and implementation of Value of Information in clinical research.

In this workshop, members of the Collaborative Network on Value of Information Group will address these issues by: (1) introducing the key concepts behind Value of Information, (2) presenting key Value of Information measures, and (3) highlighting how they can be useful in directing future research. We will demonstrate how to compute these measures using three alternative methods, using an Excel spreadsheet and online tools. We will also explore graphical presentations of these measures.

The workshop is a mixture of lectures and computer-based examples. Exercises using a web-based interface and Excel worksheets will be provided to calculate Value of Information measures. Participants will also be provided with an example Value of Information analysis from the literature.

Timetable Introduction to Value of Information Lecture (45 min): (1) methods for research prioritization and trial design, (2) decision-making in healthcare, (3) health economic decisions and monetary value, (4) uncertainty in decision-making, (5) expected Value of Perfect Information, (6) calculating expected Value of Perfect Information, (7) presenting expected Value of Perfect Information Exercise, (8) on computer (15 min)—determining trial outcomes using Value of Information Lecture (30 min), (9) Expected Value of Perfect Partial Information, (10) calculating Expected Value of Perfect Partial Information Exercise—on computer (15 min) designing clinical research using Value of Information Lecture (15 min): Expected Value of Sample Information, Expected Net Benefit of Sampling, Calculations for these measures, and Conclusions.

Goals of Tutorial: This course will introduce Value of Information measures and their use in research prioritization and trial design. By the end of the course, participants will be able to: (1) understand how Value of Information analysis can be used to determine research priorities and design clinical research, (2) interpret four key Value of Information measures, the Expected Value of Perfect Information, the Expected Value of Perfect Partial Information, the Expected Value of Sample Information, and the Expected Net Benefit of Sampling, and (3) explore the results of a Value of Information analysis using graphical displays.

WK-4  
ESSENTIALS OF CLINICAL TRIALS

Organizer:  
YVES ROSENBERG  
NHLBI, NIH

Contributor(s):  
DIXIE ECKLUND  
University of Iowa

JULIE QIDWAI  
Veristat

CHRISTOPHER COFFEY  
University of Iowa

This workshop will offer an overview of the some of the essential concepts related to the design and conduct of clinical trials. Presentations and discussion will cover the rationale for clinical trials, various design, and operational issues, as well as regulatory and basic ethical principles. Faculty with experience in conducting US and international, industry, and government–funded clinical trials, including COVID-related trials, will provide and contrast perspectives as necessary to be relevant to a US as well as an international audience. Each section will start with a brief overview of the topic and will be followed by an interactive discussion.

Specific topics to be addressed include the following:  
Introduction to the workshop: 5 min.
Rationale, phases and types of clinical trials: 30 min. This section will cover the rationale for clinical trials and why and when we need randomization; examine different study designs (parallel, cross-over, cluster, factorial, and other innovative designs); and phases of clinical trials from first in human to efficacy trials and effectiveness and comparative effectiveness trials; platform trials and other designs used in the COVID pandemic.

Principle of Hypothesis Testing, Objectives, and Endpoints: 30 min. This section will review how to state a hypothesis, define objectives, choose endpoints, choose your study population, and select a control group issues around placebo use and blinding. Examples of endpoints selected for vaccine development and treatment regimens during the COVID pandemic will be evaluated.

Basic protocol, data management, and quality control considerations: 25 min. This section will explain how to develop a well-written protocol, maintain data integrity by minimizing and documenting dropouts, monitoring and assessing protocol adherence, and monitoring safety. Examples of quality control with COVID protocols will be discussed.

Regulatory and ethical issues: 25 min. This section will cover the general rules and practices surrounding the informed consent process, changes that occurred as a result of the COVID pandemic, Ethics Committees and Institutional Review Board rules and reporting requirements, and Good Clinical Practices. Examples of ethical dilemmas and racial disparities that occurred as a result of the COVID pandemic will be discussed.

Wrap up Discussion: 5 min.

Goals of Tutorial: Attendees should be able to describe the rationale and key design elements of a clinical trial, the essential principles of data management and quality control, as well as the essential ethical concepts and regulatory issues related to the design and conduct of clinical trials, including trials performed during a pandemic, and use this knowledge to successfully contribute as a researcher or collaborator in one or more stage of a clinical trial.

**HOW THE ESTIMAND ADDENDUM TO THE ICH E9 GUIDELINE HELPS STRUCTURE CLINICAL OBJECTIVES, ANALYSES, AND CONCLUSIONS—A SERIES OF ONCOLOGY CASE STUDIES**

**Organizer:**
STEFAN ENGLERT
AbbVie Inc.

**Contributor(s):**
JULIANE MANITZ
EMD Serono Inc./Merck KGaA
FENG LIU
AstraZeneca
STEVEN SUN
Johnson & Johnson

**Background:** The estimand framework aims to increase the dialogue between functional areas working on proper alignment of trial objectives formulation, design, conduct, statistical analyses, and conclusions. The draft addendum on Estimands and Sensitivity Analysis in Clinical Trials to the ICH E9 guideline on Statistical Principles for Clinical Trials was released in August 2017. In December 2019, the final version of the ICH E9 estimand addendum was published. The new framework requires clarity and precision in description of the treatment effect, in particular, explicitly accounting for events which occur after randomization/treatment start and either preclude observation of the variable of interest or affect its interpretation (“intercurrent events”). It highlights the need for a discussion among key stakeholders during the design phase, resulting in more precise clinical trials objective. The estimand framework is anticipated to have a major impact on drug development. The estimand that reflects the trial objectives will determine the trial design, data collection, trial conduct, analysis, and interpretation. This work is a result of a cross-industry working group of statisticians and clinicians working on connecting the ICH E9 addendum concepts to applications in oncology. The working group was formed in February 2018 and currently has 34 members (14 EU + 20 US) from 21 companies to ensure common understanding and consistent definitions for time-to-event estimands in oncology.

**Methods:** Followed by a general introduction to the estimand framework, we illustrate the impact of the addendum by applying it to a series of oncology case studies:

- Censoring mechanisms: evaluate the use of censoring to handle intercurrent events, related assumptions, and interpretation, discussing the often performed sensitivity analyses and possible alternatives in view of the estimand framework.
- Treatment switching: describe how the estimand framework allows to explicitly account for different types of treatment switching and offers a systematic and transparent approach for assessment.
- Solid and hematologic tumors: focus on relevant estimands, intercurrent events, and sensitivity analyses and demonstrate how the estimand framework seeks to increase transparency on the treatment effect of interest and facilitates communications between stakeholder.
COVID-19: assess the impact of COVID-19 on the clinical trial objective, propose strategies to handle COVID-19-related intercurrent events, and show how the estimand framework provides a common language to discuss the impact of COVID-19 in a structured and transparent manner.

Results: Key findings from this exercise are that the estimand framework: (1) makes implicit assumptions transparent, (2) facilitates the discussions about patients’ journeys, (3) seeks to increase transparency on the clinical question of interest and facilitates a precise definition of the treatment effect, (4) prospectively plans the handling of certain intercurrent events, potentially leading to a different data collection strategy, and (5) is useful for structuring discussions about the impact of pandemics and mitigating measures one can take.

Conclusion: Recommendations for design, data collection, analysis, and reporting for clinical trials planned post-addendum will be given. Key clinical implications of this work are that the treatment effect reflecting the clinical question posed by a given clinical trial objective will need to be more precisely defined in study protocols, guidelines, and publications. If the estimate is likely to be biased in light of an unforeseen impact like COVID-19, the estimand framework provides various stakeholders a common language to discuss the impact in a structured and transparent manner.

Goals of Tutorial: The goal of this session is to bring all disciplines together and maximize awareness of the ICH E9 addendum as well as demonstrate how it helps interdisciplinary teams to formulate clinical trial objectives, design, conduct, primary, secondary, and sensitivity analyses, as well as conclusions. There will be time allocated after each methods section to interactively deepen the knowledge and gain firsthand insights from a cross-industry international working group.

The concept of applying Quality by Design to clinical trials is increasingly accepted and is becoming common clinical trials vocabulary. In addition to ongoing incorporation in global regulatory frameworks (ICH E8), the approach is being successfully applied in a range of both academic and industry studies to streamline designs, and to proactively identify and address factors that are likely to significantly impede the conduct of the study, place trial participants at unnecessary risk, or impede usability of the resulting data (in other words, to become “errors that matter”). This panel discussion will provide attendees with the information and resources they need to begin or enhance their implementation of Quality by Design, and will: (1) share case studies highlighting diverse models of implementation and lessons learned; (2) discuss implementation maturity and approaches for planning and tracking progress over time; and (3) engage the audience in discussion around understanding and applying key concepts, including identifying critical-to-quality factors, proactively addressing important risks to study quality, and engaging the broad range of stakeholders in study design.

Discussion will include design-related lessons learned and best practices identified during the COVID-19 pandemic. Attendees will walk away with a better understanding of the importance and benefits of a Quality by Design approach, and be armed with new tools and insights to support rigorous, high-quality clinical trials that meaningfully advance public health.

Speakers will be drawn from Clinical Trials Transformation Initiative’s multi-stakeholder project team, and we anticipate including one or more principal investigators with direct experience implementing Quality by Design at academic health centers (likely drawn from Duke, University of California Irvine, Georgetown, Kansas University Medical Center, and/or Oxford, depending on global travel conditions), as well as regulatory and/or ICH E8 working group, patient, and industry perspectives.

Goals of Tutorial: Bring forward a new way of designing better protocols.

POSTER PRESENTATIONS

P-1

DEVELOPMENT AND IMPLEMENTATION OF WORK ENGAGEMENT STRATEGIES IN A CLINICAL RESEARCH CONSORTIUM DURING THE CORONAVIRUS DISEASE 2019 (COVID-19) PANDEMIC

Primary Author: MARCUS R JOHNSON
Durham VA Healthcare System
Co-Author(s):
ALIYA ASGHAR
VA Cooperative Studies Program
KANDI VERLARDE
VA Cooperative Studies Program
MARTI DONAIRE
VA Cooperative Studies Program
KAREN BRATCHER
VA Cooperative Studies Program

Work engagement is defined as a positive work-related state of mind that is characterized by vigor, dedication, and absorption. The engagement of staff has been associated with their performance and efficiency, productivity, safety, attendance and retention, customer service and satisfaction, and several other organizational success factors. The Coronavirus Disease 2019 (COVID-19) is the infectious disease caused by the most recently discovered coronavirus and is now a pandemic that is affecting many countries globally. The literature surrounding the employment of measures and strategies to increase work engagement among clinical research staff during pandemics is scarce, and to date, focuses primarily on healthcare and community health workers. The Cooperative Studies Program Network of Dedicated Enrollment Sites is a clinical research consortium of 10 medical centers that are embedded within the Department of Veterans Affairs Health Care System. The consortium recently developed and implemented strategies that were intended to maintain work engagement among clinical research staff at each of the sites within the consortium.

In this poster, we describe the development and deployment of these strategies to clinical research study teams in our clinical research consortium. It is our hope that the opportunities, successes, and challenges described here will serve as a useful resource for other clinical research consortia that are working to identify approaches to keep their staff members engaged during the current pandemic, as well as in other potential future situations in which their primary operations may be altered during other times of crises.

P-2
LESSONS LEARNED FROM THE RESTORING INSULIN SECRETION STUDY: A MULTI-PROTOCOL, MULTI-CENTER CLINICAL TRIAL IN ADULTS AND YOUTH WITH PREDIABETES AND EARLY TYPE 2 DIABETES

Primary Author:
SHARONE L EDELSTEIN
George Washington University

Co-Author(s):
ASHLEY H TJADEN
George Washington University

A few clinical trials recruit both youth and adults to protocols with common interventions and outcomes. The challenges of recruiting and retaining youth and adult participants can vary widely, as can maintaining protocol adherence and retention. The Restoring Insulin Secretion Study provides an excellent example of collaboration among four adult and four pediatric academic clinical centers in several aspects of clinical trials. From 2013 to 2019, the Restoring Insulin Secretion Consortium conducted three parallel clinical trials in youth and adults with prediabetes or early type 2 diabetes: a four-site, two-arm Pediatric Medication Study (91 youth and 10–19 years), a three-site, four-arm Adult Medication Study (267 adults and 20–65 years), and a single-center, two-arm Adult Surgical Study (88 adults and 22–65 years). Medication study participants received 12 months of pharmacologic treatment followed by 9 months of treatment, while surgical study participants received 24 months of treatment. Primary study outcomes included complex measures of insulin secretion from the pancreatic beta-cell. Clinical investigators included adult and pediatric endocrinologists with little experience in working across age groups. A central laboratory performed all assays. A common coordinating center managed communications, documents, data entry systems, data analysis, and publication. The three trials successfully completed common protocol elements including (1) common eligibility criteria (except age); (2) common treatment arms: all studies included a randomized treatment arm of metformin alone, and the medication studies randomized youth and adults to a second common intervention: 3 months of insulin glargine followed by 9 months of metformin; (3) Common procedures: All 3 studies completed identical 2-step intravenous hyperglycemic clamps and 3 hour oral glucose tolerance tests at baseline as well as twice during follow-up; (4) Common training: All investigators and study staff completed a joint 3-day initial study training, and web-based annual refreshers; and (5) Common timing: the two medication studies had identical visit schedules and procedures completed at each visit. Recruitment among youth was accomplished directly through pediatric endocrinology and diabetes centers; participants remained engaged and active throughout the study, with youth and their parents eager to work with the study investigators to improve their health. In contrast, many approaches were taken to identify and engage potential adult participants, including employing consultants and hiring a recruitment company to revise recruitment materials. Retention among adults was more challenging with issues related to treatment side effects of a medication
only used in adults, inclusion of a placebo group, work schedules, and life events hindering completion. Despite these challenges, investigators and coordinators were successful in maintaining a high level of protocol adherence in all studies. Results across the cohort revealed large differences in the pathogenesis of type 2 diabetes in youth versus adults, as well as important differences in beta-cell function in response to treatment over time, results which were not otherwise possible to identify using studies without identical elements. In summary, although challenging, a common and intensive protocol across a wide range of ages is possible, and allowed both intervention and age-related comparisons of dysglycemia and beta-cell function. Restoring Insulin Secretion allowed investigators to cross-fertilize ideas and test assumptions about disease similarities or differences across populations.

**P-3**

**DEVELOPMENT OF A COVID-19 IMPACT ASSESSMENT FOR NATIONAL DRUG ABUSE TREATMENT CLINICAL TRIALS NETWORK**

**Primary Author:** KATHRYN R HEFNER  
The Emmes Company

**Co-Author(s):**  
ASHLEY CASE  
The Emmes Company  
ABIGAIL MATTHEWS  
The Emmes Company  
LANDHING MORAN  
NIDA Center for Clinical Trials Network

The COVID-19 Impact Assessment was developed by Emmes’ National Drug Abuse Treatment Clinical Trials Network’s Data and Statistics Center along with National Institute on Drug Abuse’s Center for Clinical Trials Network. With numerous clinical trials investigating treatments for substance use disorder set to initiate data collection in 2020 and 2021, the COVID-19 pandemic presented novel and unanticipated challenges to study implementation. Specifically, it presented potential for disrupting a wide range of factors that could impact trial operations (study visit planning, clinic closures, public health measures, transportation and childcare issues, changes in clinical practice, and economic factors) and outcomes (social isolation, changes in medical and/or psychiatric functioning, changes in substance use patterns including substance used, quantity and frequency).

Several National Drug Abuse Treatment Clinical Trials Network clinical trials were in protocol development stage when the pandemic hit. Initially, different protocol teams began compiling questions and assessments within their respective studies in anticipation of the need to assess the prevalence of COVID-19 within the study population as well as the potential impacts of the pandemic. Center for Clinical Trials Network and National Institute on Drug Abuse Data and Statistics Center identified the opportunity to streamline assessment by creating a standardized form that could be used across all trials, which would allow for cross-study harmonization and data analysis. Items for the assessment were drawn from publicly available survey repositories including the CDC Community Survey Online Question Bank, PhenX Toolkit, and the National Institutes of Health Office of Behavioral and Social Sciences Research, and were identified in reference to five primary domains of particular relevance to National Institute on Drug Abuse’s vulnerable population of individuals with substance use disorder: (1) personal exposure and illness related to COVID-19 (10 items); (2) mental health and healthcare impact (32 items); (3) knowledge and beliefs about COVID-19 (7 items); (4) social distancing regulations, behavior, and beliefs (37 items); and (5) employment and economic impact and housing stability (9 items).

Through an iterative process, stakeholders from Data and Statistics Center, Center for Clinical Trials Network, and Clinical Trials Network determined which items should be retained for inclusion, removed, or added, or considered priority for inclusion during the COVID-19 pandemic, until consensus was reached. While the COVID-19 illness and exposure and Mental Health and treatment impact were considered higher priority, the inclusion of modules or (select items) in trial assessments should be retained for inclusion, removed, or added, or considered priority for inclusion during the COVID-19 pandemic, until consensus was reached. While the COVID-19 illness and exposure and Mental Health and treatment impact were considered higher priority, the inclusion of modules or (select items) in trial assessments is at the discretion of the lead investigative team. These may be considered a menu of options for investigators to choose from, so that constructs included are assessed in a harmonized manner across trials.

Data from the COVID-19 impact assessment can inform how COVID-19 related illness, changes in mental health functioning and/or treatment access (including substance use treatment medications and psychosocial supports) may impact the operations and outcomes of treatments for substance use disorder in Clinical Trials Network trials that are conducted during the COVID-19 era. Since the development of the form, five National Drug Abuse Treatment Clinical Trials Network trials that are in pre-implementation stage (beginning recruitment as early as January 2021) have elected to include the form either in its entirety or select domains, predominantly choosing to include the mental health and healthcare impact and personal exposure and illness modules.

**P-4**

**DEVELOPING AN E-CONSENT SYSTEM**

**Primary Author:** BRIAN K BURKE  
The George Washington University
With the COVID-19 pandemic, the ability to coordinate and manage research studies remotely has become increasingly important. Most systems offer a variation of a mobile interface for study participants to complete self-administered questionnaires outside of the clinical setting. However, there was a need for functionality to allow a potential participant to virtually and electronically complete a screening questionnaire and provide consent. The web development team and research staff at the George Washington University Biostatistics Center collaborated to create a web-based public form and electronic informed consent system. This system allows potential participants to be screened and join studies without the need to be physically present to sign regulatory documents. The e-consent system is based on the Biostatistics Center’s existing electronic patient report outcome system. Users are able to access the system on a variety of devices, as the display is tailored to the size of the screen. To assure data quality and security, the system incorporates reCAPTCHA verification, email verification, tailored in-system messaging, personal links and codes, link expiration, electronic signature, and encryption. Existing features from the electronically patient report outcome system—such as skip patterns, range checks, lookup tables, and partial saving—were utilized to minimize data quality issues. In describing the design, implementation, successes, and challenges of this system, the Biostatistics Center team hopes to inform other coordinating centers and research studies interested in utilizing virtual enrollment systems for remote research.

P-5

PATIENT AND PUBLIC INVOLVEMENT IN CLINICAL TRIALS: A MIXED-METHODS STUDY IN A CLINICAL TRIALS UNIT TO IDENTIFY GOOD PRACTICE, BARRIERS, AND FACILITATORS

Primary Author:
ATHENE J LANE
University of Bristol

Introduction/objective: Trial success relies on adequate and timely recruitment. We investigated how patient and public involvement or research partners/stakeholders is implemented within a UK Clinical Trials Unit’s portfolio of trials, perceived barriers to, and facilitators of, its successful patient and public involvement implementation, and perspectives on the role of Clinical Trials Units in patient and public involvement.

Background: Only a third of trials achieve their recruitment targets and patient and public involvement can improve how trials are designed, conducted, and disseminated. The inclusion of patient and public involvement has been advocated for many years, with formal guidance from INVOLVE in the United Kingdom and Patient-Centered Outcomes Research Institute in the United States on implementing patient and public involvement in clinical research. However, challenges to successful patient and public involvement have frequently been reported and the uptake and utilization of patient and public involvement in trials are unclear.

Methods: A mixed-methods study design, involving: (1) an online survey of 26 Trial Managers, to determine how trials include patient and public involvement and the support required from Clinical Trials Units and (2) interviews with Trial Management Group members and patient and public involvement representatives from eight purposively selected case study clinical trials. Quantitative survey data were summarized using descriptive statistics and interview transcripts analyzed thematically. Two public contributors advised on interview topic guides and provided feedback on findings.

Results: The 21 Trial Managers (81%) who completed the survey had a mean of 6.7 years trial management experience. Fifteen Trial Manager (71%) reported that patient and public involvement contributors were on their Trial Steering Committee (n = 8), Trial Management group (n = 6), or both (n = 1) with a mean of 4.8 contributors per trial (range = 0–15). The trials of four Trial Managers consulted patients through other methods, such as a separate patient and public involvement group. One Trial Manager was unaware of any patient and public involvement group. One Trial Manager was unaware of any patient and public involvement group. One Trial Manager was unaware of any patient and public involvement group. One Trial Manager was unaware of any patient and public involvement group.
Managers reported no changes had resulted from patient and public involvement consultation. Twelve Trial Managers reported that patient and public involvement representatives were paid for their activities range (£10–50/h). Only five Trial Managers reported that training was provided for patient and public involvement representatives; but it was valued by the patient and public involvement contributors. A few trial staff had received patient and public involvement training but, again, where they had it was found to be useful.

Across the eight Trial Management groups, 19 interviews were conducted with public contributors (n = 8), Trial Manager (n = 5), Chief Investigators (n = 3), patient and public involvement-coordinators (n = 2) and a researcher. Public contributors wanted and valued feedback on changes from their inputs, but this was not always provided. Trial Managers reported barriers to successful patient and public involvement: namely, recruitment challenges, the representativeness and availability of patient and public involvement members, managing group dynamics, maintaining professional boundaries, negative attitudes to patient and public involvement among some researchers, a lack of continuity of trial staff, and the complex academic environment. Successful patient and public involvement required early and explicit planning, sharing of power with public contributors, building and maintaining relationships, and joint understanding and clarity about expectations/roles. Clinical Trials Units have an important role to play in supporting patient and public involvement recruitment, signposting, and coordinating patient and public involvement within trials.

**Conclusion:** Patient and public involvement in clinical trials is highly variable although can be impactful. Planned patient and public involvement supported by training and Clinical Trials Unit coordination can facilitate patient and public involvement and ensure public contributors’ inputs are optimized.

**P-6**

**IDEAL: AN INTEGRATED EVALUATION PATHWAY FOR COMPLEX NON-PHARMA INTERVENTIONS**

**Primary Author:**
PETER MCCULLOCH  
University of Oxford

**Co-Author(s):**
ALLISON HIRST  
University of Oxford

ARSENIO PAEZ  
University of Oxford

MUQATHIR IBRAHIM  
University of Oxford

This talk will introduce the IDEAL Framework and Recommendations to those not familiar with it. IDEAL is an integrated evaluation pathway for surgery, therapeutic devices, and other complex invasive therapies. It is particularly suited to the pandemic era as it gives ethically and methodologically based recommendations on how to evaluate very new innovations urgently, and how to quickly progress to wider scale use while maintaining appropriate evaluation up to and beyond the point of a randomized controlled trial. It deals with the question of when a randomized clinical trial is inappropriate and with the specific problems of randomized clinical trials of complex interventions.

**P-7**

**THE EFFECTS OF PRIOR PROSPECTIVE COLLABORATIVE STUDIES (IDEAL STAGE 2B) ON THE QUALITY AND HIGH-IMPACT JOURNAL PUBLICATION OF RANDOMIZED CONTROLLED TRIALS (IDEAL STAGE 3) EVALUATING SURGICAL INNOVATIONS: A CASE–CONTROL SYSTEMATIC REVIEW OF THE LITERATURE**

**Primary Author:**
MUQATHIR IBRAHIM  
University of Oxford

**Co-Author(s):**
JIAYIE YU  
University of Oxford

BAPTISTE VASEY  
University of Oxford

ARSENIO PAEZ  
University of Oxford

ALLISON HIRST  
University of Oxford

PETER MCCULLOCH  
University of Oxford

**Introduction:** Randomized controlled trials in surgery often face methodological challenges, resulting in low-quality or sometimes, failed trials. The IDEAL framework was developed to address these challenges. The exploration stage (IDEAL 2b) in particular proposes that carefully planned prospective collaborative multicenter studies prior to the development of a randomized clinical trial (Assessment stage—IDEAL 3) will increase the quality and high-impact journal publication of surgical randomized clinical trials. Little empirical evidence exists to support this proposal.

**Objective:** To assess the effect of IDEAL stage 2b on the quality and high-impact journal publication of randomized clinical trials evaluating surgical innovation.
Methods: We conducted a case–control systematic review with two objectives. We developed a search strategy for Ovid Medline to identify randomized clinical trials evaluating surgical innovations, published between 2015 and 2019. Randomized clinical trials in journals with an impact factor of 5 or higher were classified as “cases” while those with impact factor less than 5 were “controls.” We used Google Scholar to search for any prospective collaborative studies (IDEAL 2b-like studies) which had preceded the included randomized clinical trials. We also assessed the quality of all randomized clinical trials with the Cochrane Risk of Bias tool 2 and categorized them as High or Low quality. We tested the odds of exposure to IDEAL 2b-like studies given publication in a high-impact journal and the probability of a high-quality randomized clinical trial if IDEAL 2b-like study was present.

Results: Preliminary data suggest a higher proportion of IDEAL 2b-like studies and high-quality randomized clinical trials in the cases compared to controls.

Conclusion: This study will determine whether the proposition that preparatory IDEAL2b-like studies can influence the quality and impact of subsequent randomized clinical trials is supported by current literature.

P-8
THERE ARE TOO MANY P-VALUES SMALLER 0.01 FOR CLINICALLY RELEVANT PRIMARY ENDPOINTS REPORTED IN RANDOMIZED TRIALS—RESULTS FROM A SYSTEMATIC ANALYSIS OF HIGH-IMPACT PUBLICATIONS

Primary Author:
RENKE MAAS
Institut für Experimentelle und Klinische Pharmakologie und Toxikologie, FAU Erlangen

Co-Author(s):
BIRGIT DEUTSCH
Institut für Experimentelle und Klinische Pharmakologie und Toxikologie, FAU Erlangen

ANNA ROGGENHOFER
Institut für Experimentelle und Klinische Pharmakologie und Toxikologie, FAU Erlangen

JOERG HASFORD
Ludwigs-Maximilian-University Munich

There are too many p-values smaller 0.01 for clinically relevant primary endpoints reported in randomized trials—results from a systematic analysis of high-impact publications.

Publications of randomized clinical trials reporting highly significant differences in clinically relevant outcomes are usually greeted with due and positive enthusiasm. However, serious ethical questions arise when more patients have been allocated to the inferior treatment than necessary to demonstrate statistical and clinical superiority of a treatment. This prompted us to systematically assess the balance of statistical certainty and ethics in publications in high-impact medical journals.

Methods: All original reports of randomized clinical trials reporting clinically relevant outcome(s) (overall survival, disease specific mortality, and progression of or protection from serious disease) published 2018 in five major medical journals (British Medical Journal, New England Journal of Medicine, Journal of Clinical Oncology, Journal of the American Medical Association, and The Lancet) were identified and analyzed.

Results: Of the 384 trials analyzed, 152 randomized clinical trials (39.6%) reported a p-value < 0.01 indicating a possibly larger than necessary difference between treatment groups for total mortality or a clinically relevant endpoint, respectively. None of these publications specifically addressed ethical aspects related to the highly significant findings. Moreover, for 133 publications (88%) reporting highly significant differences for total mortality or a clinically relevant endpoint, respectively, we could not identify an ethically acceptable explanation for these highly significant findings.

Conclusion: When reporting highly significant results it is advisable, in all but the most obvious cases, to provide explanations that these results were obtained without violating ethical norms. Furthermore, the data indicate that there is a need to find new means of balancing research interests and the public demand for maximized certainty of clinical benefit with ethical principles protecting patients participating in clinical trials.

P-9
AN EXAMINATION OF METHODS FOR PARAMETER ESTIMATION FOLLOWING ADAPTIVE GROUP SEQUENTIAL CLINICAL TRIALS

Primary Author:
BRYAN S NELSON
Harvard T.H. Chan School of Public Health

Co-Author(s):
LINGYUN LIU
Cytel/Vertex Pharmaceuticals

CYRUS MEHTA
Cytel/Harvard T.H. Chan School of Public Health

Statistical methods for controlling the type 1 error of hypothesis tests in adaptive group sequential clinical
P-10
COST-EFFECTIVENESS ANALYSIS OF CARDIAC VERSUS NON-CARDIAC SURGERY ON CARDIOVASCULAR-DISEASE POPULATION

Primary Author:
FEI YUAN
McMaster University

Co-Author(s):
ANDREW LAMY
McMaster University
WESLEY TONG
McMaster University

Background: Cardiovascular diseases are becoming a bigger and bigger burden for an aging population. As the diseases progress, this population requires various surgical services. These services are increasingly expensive in a healthcare system. Therefore, to help with optimally allocating resources and government policymaking, it is desired to analyze the cost-effectiveness of various surgical services. This is especially meaningful to “advance rigorous and ethical trials in this field in the Pandemic Era.”

Objectives: To understand the long-term economic implication of various surgical services on cardiovascular-disease population through a life-time cost-effectiveness analysis.

Methods: This case-control study investigated cardiac and non-cardiac surgeries based on participants from three international surgical trials. These trials originally studied various interventions on patients who had cardiovascular diseases and required various surgeries. There were 12,252 participants undertook cardiac surgery from the Coronary Artery Bypass Surgery Off or On Pump Revascularization Study and Steroids In caRdiac Surgery Trial trials. There were 10,010 non-cardiac surgery participants from the Perioperative Ischemic Evaluation 2 trial. The evaluation was based on incremental cost-effectiveness ratio. These costs were included cardiac surgeries, non-cardiac surgeries and three health states (myocardial infarct, stroke, and new renal failure). Canadian public payer’s perspective was adopted and translated into US dollars. A time horizon of life-time was applied. A cohort-level aggregated Markov model was used with a cycle length of 30 days in the analysis. There were five health states included event-free, myocardial infarct, stroke, new renal failure, and death. Death was an absorbing state. For other states, participants could stay in their current state or transmit to any other states.

Results: The mean ages were 67.6 years old for participants undertaken cardiac surgery and 68.5 years old for those undertaken non-cardiac surgery. The Markov model started from randomization and ran exhaustively till every participant was projected to die in the model. The mean cost per non-cardiac surgery patient was estimated to be US$9206 and mean survival was 2.18 quality-adjusted life year gained. The mean cost per cardiac surgery participants was estimated as US$22,245 and such patients survived 3.09 quality-adjusted-life year on average. Compared with a non-cardiac surgery patient, a cardiac surgery one spent US$13,038 more and survived 0.91 quality-adjusted-life year longer on average. An incremental cost-effectiveness ratio was estimated as US$14,274 per quality-adjusted-life year gained. Being lower than US$50,000 per quality-adjusted-life year gained, this qualified the cardiac
surgeries as treatments with “better outcomes at lower cost” according to the American College of Cardiology/American Heart Association guideline.

**Conclusion:** For participants who have a medical history of cardiovascular diseases and a Canadian perspective, cardiac surgeries have high value to improve patients’ cardiovascular health. This analysis needed to adjust for potential confounding factors.

**P-11**

**AUTOMATED, CUSTOMIZED, PATIENT-CENTRIC PERFORMANCE FEEDBACK IN CLINICAL TRIALS TO IMPROVE DATA ACCURACY AND PATIENT COMPLIANCE**

**Primary Author:**
ANDREA MARRAFFINO
WCG Analgesic Solutions

**Co-Author(s):**
ARTURO MORALES
WCG Analgesic Solutions

Ensuring the accuracy and reliability of clinical trial data is a cornerstone of producing reliable study results. Variability in key performance tasks at a subject level can introduce noise and increase the risk of trial failure. In the age of COVID-19, these issues are even more important as many patients are being monitored remotely. An automated, customizable, patient-centric feedback system was created to allow patients to be more involved in their own performance and progress throughout the clinical trial. This system provides feedback to patients on key performance tasks such as adherence to study medication, compliance with eDiary completion, and accuracy of symptom reporting. The reports are generated on a weekly or bi-weekly period throughout the study and provided to subjects on a tablet or hand-held device. In a 30-min usability session, four patients assessed the usability and understanding of the reports. Overall, patients reported with enthusiasm for the design of the report as well as appreciation for the effort to connect more with patients by providing them with feedback on their performance. Participants were generally able to use the report to find details of their performance and determine if it was in an acceptable range or needed improvement. This performance feedback system is currently being implemented in an ongoing clinical trial. Future research will provide information on the effectiveness of the performance system in improving data accuracy and patient compliance.

**P-12**

**ELASTIC PRIORS TO DYNAMICALLY BORROW INFORMATION FROM HISTORICAL DATA IN CLINICAL TRIALS**

**Primary Author:**
LIYUN JIANG
China Pharmaceutical University

**Co-Author(s):**
LEI NIE
Center for Drug Evaluation and Research, US Food and Drug Administration
YING YUAN
The University of Texas MD Anderson Cancer Center

Use of historical data and real-world evidence holds great potential to improve the efficiency of clinical trials. One major challenge is how to effectively borrow information from historical data while maintaining a reasonable type 1 error. We propose the elastic prior approach to address this challenge and achieve dynamic information borrowing. Unlike existing approaches, this method proactively controls the behavior of dynamic information borrowing and type 1 errors by incorporating a well-known concept of clinically significant difference through an elastic function, defined as a monotonic function of a congruence measure between historical data and trial data. The elastic function is constructed to satisfy a set of information-borrowing constraints prespecified by researchers or regulatory agencies, such that the prior will borrow information when historical and trial data are congruent, but refrain from information borrowing when historical and trial data are incongruent. In doing so, the elastic prior improves power and reduces the risk of data dredging and bias. The elastic prior is information-borrowing consistent, that is, asymptotically controls type 1 and 2 errors at the nominal values when historical data and trial data are not congruent, a unique characteristics of the elastic prior approach. Our simulation study that evaluates the finite- sample characteristic confirms that, compared to existing methods, the elastic prior has better type 1 error control and yields competitive or higher power.

**P-13**

**SUMMARIZING HISTORICAL INFORMATION FOR MAKING GO/NO GO DECISIONS IN PHASE 2 CLINICAL TRIALS**

**Primary Author:**
WEI WEI
Yale University
Co-Author(s):
YILIN LIU
Yale University

In phase 2 clinical trials of cancer or other serious diseases, a novel treatment is considered promising if the observed response rate is significantly higher than the historical response rate, which is the response rate to the standard of care or the best treatment available. The historical response rate is commonly pre-specified by clinical investigators based on past experience and is often considered as a fixed quantity. To take into account the between-trial heterogeneity in historical response rate, we propose a novel approach to synthesize evidence from historical controls by clustering previous trials into non-exchangeable subgroups and averaging over all the possible subgroup models. Based on the synthesized evidence on historical response rate, we construct a statistical framework for making better informed go/no go decisions by minimizing the total misclassification errors associated with the model for historical response rate and the model representing the promising response rate.

P-14
LEVERAGING PATIENT PROFILES FOR ROUTINE PHARMACOVIGILANCE IN CLINICAL TRIALS

Primary Author:
RADHIKA KONDAPAKA
The Emmes Company, LLC

Co-Author(s):
AMARNATH VIJAYARANGAN
The Emmes Company, LLC

ALICE HENNING
The Emmes Company, LLC

PRATAP KUNWAR
The Emmes Company, LLC

KELLY SCHLESSMAN
The Emmes Company, LLC

Patient profiles are PDF files that include all of the study data collected for an individual patient organized by time. Historically, the Food and Drug Administration used to require companies to submit patient profiles in full, which served as a complement to the electronic data sets. While those guidelines have since changed to instead require the submission of data in the Study Data Tabulation Model format and patient profiles are no longer a necessity for the Food and Drug Administration, a pared down version of the patient profile outline can be utilized to efficiently organize data for pharmacovigilance. While standard patient profiles contain more detailed displays in form of Tables, Listings, and Figures, the new proposed method involves transforming that data into a graphical form for more efficient space use and easier reading.

All of the relevant data from the traditionally bulky patient profiles is still captured, but on one page and in a way that makes it more streamlined for use by the pharmacovigilance team during ongoing safety monitoring in a clinical trial. The new patient profile format depicts the patient information in a graphical form where key data such as the timing of adverse events, concomitant medication administration, or laboratory measurements may be visualized in relation to the timing of the drug administration for analysis. Other data such as demographics, medical history, physical exam results, vital signs, protocol deviations, and subject disposition are also succinctly displayed.

Before the study enrollment, the pharmacovigilance team can meet with the study statistician or SAS programmer to customize the patient profile per the study indication and safety endpoints. The pharmacovigilance team can use this program anytime to run the patient profiles on events of interests, serious adverse events, deaths, Suspected Unexpected Serious Adverse Reaction, premature study product discontinuations, and study participation withdrawals. These concise patient profiles will be a useful tool to present cases at data monitoring committee meetings. The graphic presentation gives a snapshot of the data and can be combined with study-specific safety data monitoring like electrolytes, liver function tests, blood pressure, oxygen saturation, and other continuous measurements. The program should be readily accessible to run on live data and will help make patient safety decisions.

The one-page patient profiles enable efficient use of medical monitors’ and data monitoring committee review time. Instead of traditional patient profiles, the new format provides a succinct presentation of patient data for routine safety monitoring in clinical trials.

P-15
PHASE I CHIMERIC ANTIGEN RECEPTOR T CELL CLINICAL TRIALS REVIEW

Primary Author:
TAO WANG
Baylor College of Medicine

Co-Author(s):
SHAUN BULSARA
Baylor College of Medicine

MENGFEN WU
Baylor College of Medicine

Chimeric antigen receptor T cells with a tumor specificity have been increasingly investigated as revolutionary cancer immunotherapy. Since the safety and
efficacy of CD19 chimeric antigen receptor T cell used in lymphoma initially reported in 2010, the development of new targets used for generating the chimeric antigen receptor T cells and their clinical applications, along with the number of early phase clinical trial of chimeric antigen receptor T therapies, has increased dramatically. Phase 1 trials, also called dose-finding studies, are usually the first step with the goal of testing for safety of novel chimeric antigen receptor T therapy to determine the maximum tolerated dose. Several dose escalation methods have been developed over time including rule-based designs, model-based designs, and the relatively new class of model-assisted designs. Our goal of this project is to overview the phase 1 designs used in current chimeric antigen receptor T trials. We searched PubMed for peer-reviewed literatures published between 1 January 2015 and 30 September 2020. The search was limited to human studies in the English language using the keywords “CAR-T phase 1,” “clinical trials,” and “full text.” Eighty-one papers were retrieved and two were excluded due to pre-defined exclusion criteria.

Data were summarized using descriptive statistics. Seventy-nine papers were included for analysis, and 72 (91.1%) of these papers had at least partial phase 1 components. Excluding 2020, since the review only partially covers that year, the number of chimeric antigen receptor T publications gradually increased over the specified years. About 64.1% of the phase 1 manuscripts centered on either leukemia, lymphoma, or a combination of the two as the disease of interest. Of the phase 1 manuscripts, 56.9% did not mention the dose escalation design, and 34.7% used the traditional 3 + 3 or a variation of said design. Almost 56% of the phase 1 trials had sample sizes of 30 or less. A majority of the phase 1 manuscripts (55.6%) did not report cohort size, and those that did report had sizes of 2 (4.2%) or 3 (40.3%). 68.1% of the trials had safety evaluation period within 6 weeks while 23.6% did not specify the timing of evaluation. 72.2% of the phase 1 chimeric antigen receptor T studies had lymphodepletion with either single (72.2%) or multiple (23.6%) infusions at one to three dose levels (62.5%). The vast majority of the trials (94.4%) had safety as the primary endpoint and 66.7% had response as the secondary endpoint. 28.2% of the trials reported at least one DLT, 42.3% had ≥ grade 3 CRS and 40.9% had ≥ grade 3 neurotoxicity. Although a majority of the phase 1 studies were registered with CT.gov, only 13.9% had any results submitted or posted to CT.gov. Standardizing the criteria and basic elements of publications are critical to ensure high quality of phase 1 clinical trials. Rule-based designs, such as the 3 + 3 and its variations, are still dominant. However, these designs are deemed less accurate and allocate more patients to suboptimal dose levels. With the quick development and high costs of chimeric antigen receptor T cell therapy, adoption of advanced designs such as model-based and model-assisted should increase to improve efficiency of clinical trials.

P-16
CONSEQUENCES OF PHASE 1 DOSE SELECTION ON GO/NO-GO DECISIONS IN ONCOLOGY DRUG DEVELOPMENT
Primary Author:
ALYSSA M VANDERBEEK
Institute of Cancer Research

Motivation: There are two critical decision points in clinical drug development: dose selection in phase 1, and at the end of phase 2, whether to proceed or not to a confirmatory phase 3 trial (go/no-go). It is therefore important to optimize decision outcomes; selecting a safe and effective dose, and accurately discriminating between effective and ineffective therapies. Novel phase 1 and phase 2 designs seek to improve end-of-trial decisions, and are often developed, compared, and assessed within each stage of development. However, less is known about the impact of early stage development decisions—particularly in phase 1—on later-stage outcomes. This study examines such effects in the phase 1 and phase 2 pipeline.

Methods: Using simulations, a partial development pipeline—a phase 1 trial followed by a phase 2 trial—is generated for five fixed dose–response scenarios. Phase 1 considers the 3 + 3, continual reassessment method, and EffTox designs; a single-arm phase 2 trial with binary endpoint follows. The probability that the drug itself is effective is also considered. Pipelines are compared using the probability of best dose selection and the probability of acceptable dose selection after phase 1, and the overall F-score (trade-off between true and false positives and negatives) of the phase 2 results.

Results: No single phase 1 design is superior in all dose–response scenarios, but the EffTox pipeline performs best for non-monotonically increasing dose–response curves. Still, false positive rates can be much higher than expected in phase 2, due in part to the drug’s effectiveness probability. There is also evidence to suggest that the number of effective doses and size of their clinical effect are additional factors.

Conclusion: This preliminary simulation study suggests that phase 2 results are sensitive to both the preceding phase 1 design and the underlying shape of the dose–response curve. True and false positive and negative rates can be much higher than expected. Consideration of (1) the experimental drug’s potential dose–response profile and (2) the impact of present trial design on future trial outcomes has the potential to
improve decision-making—in this study, phase 2 go/no-go—and offer important insight during trial planning.

P-17
CURE SICKLE CELL INITIATIVE COMMON DATA ELEMENTS RECOMMENDATIONS VERSION 1.0 DEVELOPMENT

Primary Author:
SHERITA ALAI
Emmes

Co-Author(s):
SOPHIE LANZKRON
Johns Hopkins School of Medicine

ALEXIS THOMPSON
Northwestern University

PATRICK CARROLL
Johns Hopkins School of Medicine

MICHAEL DEBAUN
Vanderbilt University Medical Center

JULIE KANTER
Medical University of South Carolina

PUNAM MALIK
Cincinnati Children’s Hospital

DEEPA MANWANI
Albert Einstein College of Medicine

JOHN PIERCIEY
Bluebird Bio

MARK WALTERS
University of California San Francisco

TRACI MONDORO
National Heart, Lung and Blood Institute

VICTORIA COLEMAN-COWGER EMMES

The Cure Sickle Cell Initiative was created in 2017 by the National Heart, Lung, and Blood Institute, part of the National Institutes of Health, to accelerate the development of safe and effective genetic therapies to improve the lives of individuals with sickle cell disease. The Cure Sickle Cell Initiative has actively engaged the sickle cell disease community of affected individuals, family members, clinicians, and advocates to work together on a path to cures, while also encouraging collaboration among researchers, industry, non-profit organizations, and policy-making agencies.

Common data elements have been developed for other diseases and it is overdue to establish a set of consensus data elements on sickle cell disease. To remedy this, Cure Sickle Cell Initiative common data element recommendations have been developed to build upon current consensus and facilitate the start-up of multi-center clinical research efforts. A standardized set of clinical research recommendations will increase efficiency and effectiveness of clinical research studies, increase data quality, and help educate new clinical investigators in the field of sickle cell genetic therapies research.

This process was identified, developed, and vetted by experts in the scientific community through a transparent and inclusive process. Over the past 9 months, the Cure Sickle Cell Initiative common data element Working Groups comprised of patient advocates, clinicians, and researchers reviewed current recommendations for sickle cell clinical studies from American Society of Hematology, US Food and Drug Administration, consensus measures for Phenotypes and eXposures (PhenX) catalog, Center for International Blood and Marrow Transplant Research, and various other research institutions. The Working Groups then compiled a hierarchical designation to describe both a minimum (core common data elements) and a comprehensive data set. By incorporating the new recommendations from subject matter experts, the Working Groups have worked to avoid future redundancy and provide collection of data elements through template case report forms and instrument recommendations along with any needed guidelines.

The following five (5) working groups were tasked over 6 months with drafting common data element recommendations: Genetics/Assays; Physical Examination/Medical History; Cardiopulmonary and Renal Function; Outcomes; and Monitoring Side Effects. The group members reviewed the data forms, discussed, and created consensus regarding the form content. The product of the working groups was then reviewed through a Public Review Period via the Cure Sickle Cell Initiative website (curesickle.org). After Public Review and follow-up with Working Groups, version 1.0 of the standardized data forms will be released on the Cure Sickle Cell Initiative website and the National Library of Medicine website (https://cde.nlm.nih.gov/).

The National Heart, Lung, and Blood Institute intends to require Cure Sickle Cell Initiative awardees to use the Core common data elements to expedite study start-up, standardize data collection, and allow for future data sharing. This first iteration of the common data element recommendations will be updated through an oversight committee on an annual basis. Adoption of data standards for clinical research is a shared goal for pharmaceutical companies, regulatory agencies such as the Food and Drug Administration, academic and government-based clinical researchers and government agencies, such as the National Heart, Lung, and Blood Institute. This important step for sickle cell disease and genetic therapies is unique and vital for future research in these areas.
P-18

RESPONSIBLE TRACKING AND ACCOUNTING OF PARTICIPANT MENTAL HEALTH URGENCY/EMERGENCY RESPONSES IN SUBSTANCE USE TREATMENT CLINICAL TRIALS

Primary Author: LESLIE REVOREDO
The Emmes Company, LLC

Co-Author(s):
DAGMAR SALAZAR
The Emmes Company, LLC
ASHLEY CASE
The Emmes Company, LLC
BRETT HART
The Emmes Company, LLC

The Emmes Company, LLC (Emmes), acting as the Clinical Coordinating Center and the Data and Statistics Center for the National Institute on Drug Abuse National Drug Abuse Treatment Clinical Trials Network supports multi-site clinical trials involving behavioral and pharmacological interventions for substance use disorder treatment. These studies often recruit inherently vulnerable populations at increased risk for mental health disorders, including suicidal and homicidal ideation. We recently undertook a collaborative initiative to improve tracking and accounting of study site mental health urgency/emergency responses, in order to ensure that Standard Operating Procedures were followed for these instances, through use of a standardized electronic Case Report Form. This presentation details the development and initial implementation of that electronic Case Report Form. In response to previously identified documentation and tracking process gaps, Emmes Clinical Coordinating Center Safety Monitors, Clinical Study Managers, and Data and Statistics Center Data Managers developed an electronic Case Report Form dedicated to ensuring timely, uniform, and complete accounting of study site responses to participant mental health urgencies/emergencies. These gaps included previous studies in which participants responded to suicidality assessments in a way that necessitated that they be seen by a clinician for further evaluation but was not done. The Mental Health Follow-Up Assessment was designed to track site responses to suicidal ideation and ensure that participants requiring follow-up were provided with that care. It was later revised to distinguish between “in person” and “remote” research–participant encounters, (e.g. in non-clinical facility or community settings) or virtual (e.g. video or telephonic), with in-person encounters less prevalent during the COVID-19 pandemic. Unique metrics for documenting site adherence to mental health urgency/emergency response Standard Operating Procedures were established, focusing on documenting timely clinician notification and/or assessment of participants during “in person” encounters and provision of local and/or national mental health “hot-line” referral resources to participants during “remote” encounters.

To date, the Mental Health Follow-Up Assessment electronic Case Report Form has been successfully incorporated into the data collection and monitoring frameworks for nine Clinical Trials Network trials. Although Mental Health Follow-Up Assessment deployment outcome results are in the early stages of collection, successes have already been realized in terms of (1) enhancing “real-time” monitoring of site responses to participant mental health urgencies/emergencies, (2) standardizing the documentation of study site responses, (3) encouraging adherence to site local mental health response Standard Operating Procedures, (4) improving research staff understanding of the differential responses required during “in person” versus “remote” participant encounters, (5) facilitating collection, collation, and reporting of mental health safety monitoring data, and (6) reducing the time burdens on site research staff and Emmes safety and data monitoring teams in accomplishing these tasks.

The newly developed Mental Health Follow-Up Assessment has already returned a number of the originally intended benefits. Accordingly, the Clinical Trials Network investigators and Emmes protocol teams look forward to expanding upon these initial successes as the conduct of additional Clinical Trials Network studies unfold. This fundamental process of preparing a tracking program for a required on-site task can easily be applied to other situations and instances across clinical trials and disease areas.

P-19

TRANSITIONING TO THE REMOTE COLLECTION OF URINE SPECIMENS FOR URINE DRUG SCREENING DURING COVID-19

Primary Author: LESLIE REVOREDO
The Emmes Company, LLC

Co-Author(s):
MITRA LEWIS
The Emmes Company, LLC
DIKLA SHMUELI-BLUMBERG
The Emmes Company, LLC
EVE JELSTROM
The Emmes Company, LLC

The Emmes Company, LLC (Emmes), acting as the Clinical Coordinating Center and the Data and Statistics Center for the National Institute on Drug Abuse National Drug Abuse Treatment Clinical Trials Network supports multi-site clinical trials involving behavioral and pharmacological interventions for substance use disorder treatment. These studies often recruit inherently vulnerable populations at increased risk for mental health disorders, including suicidal and homicidal ideation. We recently undertook a collaborative initiative to improve tracking and accounting of study site mental health urgency/emergency responses, in order to ensure that Standard Operating Procedures were followed for these instances, through use of a standardized electronic Case Report Form. This presentation details the development and initial implementation of that electronic Case Report Form. In response to previously identified documentation and tracking process gaps, Emmes Clinical Coordinating Center Safety Monitors, Clinical Study Managers, and Data and Statistics Center Data Managers developed an electronic Case Report Form dedicated to ensuring timely, uniform, and complete accounting of study site responses to participant mental health urgencies/emergencies. These gaps included previous studies in which participants responded to suicidality assessments in a way that necessitated that they be seen by a clinician for further evaluation but was not done. The Mental Health Follow-Up Assessment was designed to track site responses to suicidal ideation and ensure that participants requiring follow-up were provided with that care. It was later revised to distinguish between “in person” and “remote” research–participant encounters, (e.g. in non-clinical facility or community settings) or virtual (e.g. video or telephonic), with in-person encounters less prevalent during the COVID-19 pandemic. Unique metrics for documenting site adherence to mental health urgency/emergency response Standard Operating Procedures were established, focusing on documenting timely clinician notification and/or assessment of participants during “in person” encounters and provision of local and/or national mental health “hot-line” referral resources to participants during “remote” encounters.

To date, the Mental Health Follow-Up Assessment electronic Case Report Form has been successfully incorporated into the data collection and monitoring frameworks for nine Clinical Trials Network trials. Although Mental Health Follow-Up Assessment deployment outcome results are in the early stages of collection, successes have already been realized in terms of (1) enhancing “real-time” monitoring of site responses to participant mental health urgencies/emergencies, (2) standardizing the documentation of study site responses, (3) encouraging adherence to site local mental health response Standard Operating Procedures, (4) improving research staff understanding of the differential responses required during “in person” versus “remote” participant encounters, (5) facilitating collection, collation, and reporting of mental health safety monitoring data, and (6) reducing the time burdens on site research staff and Emmes safety and data monitoring teams in accomplishing these tasks.

The newly developed Mental Health Follow-Up Assessment has already returned a number of the originally intended benefits. Accordingly, the Clinical Trials Network investigators and Emmes protocol teams look forward to expanding upon these initial successes as the conduct of additional Clinical Trials Network studies unfold. This fundamental process of preparing a tracking program for a required on-site task can easily be applied to other situations and instances across clinical trials and disease areas.
The National Drug Abuse Treatment Clinical Trials Network routinely conducts multi-site clinical trials involving behavioral and pharmacological interventions for the treatment of substance use disorders. This commonly involves trials that test for substance use using urine drug screening procedures. Recent substance use is assessed by measuring the concentration of the various substances or their metabolites in urine specimens obtained from study participants once they provide consent and are enrolled in the study. This is an important assessment used in Clinical Trials Network trials, as these results are used for study outcomes and/or clinical care. The Emmes Company, LLC, acting as the Clinical Coordinating Center is responsible for providing and managing the urine drug screen supplies for these trials. Traditionally, this testing would occur when a participant is on-site for a study visit. Due to the COVID-19 pandemic and the stay at home orders set forth, this has disrupted the process of performing substance use testing at research sites. The Clinical Coordinating Center has developed a process to allow the studies to continue urine drug screening by obtaining urine specimens from trial participants remotely. This process is achieved with sites sending participants a “urine collection kit” which contains the necessary supplies for the collection and transportation of urine. The urine collection kits are easy to use and contain detailed instructions. Participants will collect their urine and ship it in the kit provided back to the site for screening. Once the kits are received on-site, research staff will first test the urine specimen for adulteration then substance screening.

Since this is a new process within the Clinical Trials Network, the Clinical Coordinating Center worked closely with research staff to deliver specific training to enable consistent collection and testing of urine samples from participants. This training included the International Air Transport Association regulations and packing requirements for exempt human specimens and instructions for the preparation, shipping, and receipt of the urine collection kits. The Clinical Coordinating Center also provided additional resources for sites, such as a Remote Urine Specimen Collection and Shipping Manual which details the entire process.

While this process has enabled studies to continue, one limitation is that study staff are not able to determine that the urine specimen is definitively from the study participant. When urine specimens are collected in person, research staff will do a temperature check on the specimen to determine that it is from the study participant. This reduces the likelihood of substitution. For urine samples collected remotely, the temperature check cannot be performed. Therefore, Principal Investigators must be aware of the potential for urine specimens to be from another source when considering this option for their studies.

All options considered, utilizing this different approach for remote urine collection has allowed for studies to continue capturing important study data during the pandemic when participants cannot access the research site. Although this process was developed to solve a need during the COVID-19 pandemic, it can potentially be a primary method for the collection of specimens for other studies in the future.

P-20
THE RECORD OF STUDY CONSULTATION: A TOOL TO CONTROL, DOCUMENT, and STREAMLINE SPONSOR GUIDANCE

Primary Author:
CHRISTOPHER A SCHROTH
Hines VA

Co-Author(s):
DOUGLAS LAMMIE
Hines VA
KIMBERLY CARLSON
Hines VA
CAITLIN AUTHIER
Hines VA

Protocol Consultations play an integral role in ensuring study sites do not deviate from protocol by providing sponsor-approved guidance on complex protocol questions. However, the Protocol Consultation process is often unstructured and can take the form of a site informally reaching out to a sponsor contact via email, phone call, and so on. Remote telework presents additional challenges to the collaborative nature of the sponsor-level discussions that involve a wide variety of personnel: Project Managers, Biostatisticians, Quality Assurance RNs, Pharmacists, Regulatory Monitors, and Principal Investigators. Involvement of numerous multi-disciplinary players necessitates the need for a streamlined communications process that can help avoid delays in providing guidance, miscommunications, and oversights. Furthermore, details and justifications for sponsor guidance on protocol consultations can be lost, especially when such decisions are made in the context of email, conference calls, and so on. To combat these challenges, the Cooperative Studies Program 2008 (CSP 2008) team at Hines Veterans Affairs Cooperative Studies Program Coordinating Center developed a Record of Study Consultation process which outlines the Sponsor-level workflow and utilizes Adobe Acrobat forms, secured SharePoint sites, and MS Teams to further streamline the Protocol Consultation process. The fillable Record of Study Consultation form is readily available to study sites on the secure Cooperative Studies Program 2008
Sharepoint Site, pre-populated to be sent to the correct sponsor personnel and prompts for key information that is needed to provide guidance. Upon submission to the sponsor, internal workflow ensures that Record of Study Consultations are assigned to appropriate Subject Matter Experts for completion and approval. MS Teams is used for comprehensive task management workflow, facilitating discussion and providing the internal team with notification/documentation of status updates, content updates, feedback, verification, and task completion. Sponsor-approved Record of Study Consultations are signed off on by the Subject Matter Expert and returned to the site. The Record of Study Consultation process streamlined communication processes, enhanced decision-making, documentation, and increased efficiency regarding providing sponsor guidance.

**P-21**

**PTXRX STUDY: HOW PRAGMATIC TRIAL DESIGN CAN INOCULATE AGAINST RISKS TO STUDY INTEGRITY IN A GLOBAL PANDEMIC**

**Primary Author:**
KIMBERLY M CARLSON  
*Department of Veterans Affairs*

**Co-Author(s):**
CHRISTOPHER SCHROTH  
*Department of Veterans Affairs*

DOUGLAS LAMMIE  
*Department of Veterans Affairs*

DAVID LEEHEY  
*Department of Veterans Affairs*

THOMAS KOPPES  
*Department of Veterans Affairs*

LINDA POLZIN  
*Department of Veterans Affairs*

TAMMY RHODA  
*Department of Veterans Affairs*

CHRISTINA CLISE  
*Department of Veterans Affairs*

JEFFERY HUMINIK  
*Department of Veterans Affairs*

YONGLIANG WEI  
*Department of Veterans Affairs*

Important safety measures limiting in-person contact to curb COVID-19 transmission make it more difficult for patients to access clinical trials and for sponsors to conduct trial management. These measures may lead to pausing or delaying study activities, to the detriment of study participants and the study’s integrity. The COVID-19 pandemic highlights the importance of efficient, innovative clinical trials designed with the capacity to be rapidly responsive to unique challenges. The US Department of Veterans Affairs Pentoxifylline in Diabetic Kidney Disease study is a multi-site, pragmatically designed randomized controlled trial that tests the hypothesis that pentoxifylline, when added to standard of care, leads to a reduction in the incidence of End Stage Renal Disease and mortality. The study opened for recruitment at six Veterans Affairs medical centers in December 2019, months before the COVID-19 pandemic disrupted all aspects of clinical care and halted all non-essential, in-person research activities. The study’s protocol was designed to accommodate either in-person or remote participant follow-up and data collection for all visits after baseline. In addition, participant study visit schedules were built with the flexibility to align with existing clinic visits. The ability to collect data remotely resulted in a minimal amount of missing data. The study’s investigational product is maintained and distributed centrally by the Albuquerque Cooperative Studies Program Pharmacy Center. This allowed for the continuation of distribution of the study investigational product without the need for an in-person visit to a Veterans Affairs medical center. Pentoxifylline in Diabetic Kidney Disease’s trial design and protocol leverage the Veterans Affairs’s research infrastructure, remote platforms, and a centralized mail-order pharmacy, and allowed the study to safely continue during a uniquely challenging global pandemic.

**P-22**

**A PARTNERSHIP BETWEEN CLINICAL SAFETY AND STATISTICS FOR AGGREGATE SAFETY ASSESSMENT**

**Primary Author:**
GABRIELLE MURASHOVA  
*Merck*

**Co-Author(s):**
GREG BALL  
*Merck*

SANDRA FARRELL-SHEIN  
*Merck*

In 1937, after the tragic Elixir Sulfanilamide incident, the Bureau of Chemistry was re-branded as the Food and Drug Administration. The push for drug regulation initiated a shift in efficacy analyses that focused on statistical evidence with scant input from experienced clinicians and physicians. More recently, the pharmaceutical industry and regulatory agencies have begun to evaluate the safety of pharmaceuticals with the same level of emphasis as for efficacy. Today, we view the safety profile of a pharmaceutical drug in a more
pragmatic way, one that harnesses the practical experience of clinical safety professionals with the quantitative approaches of statisticians to provide a more complete and impactful interpretation of safety. Nevertheless, the best way to view and organize large quantities of data is still under debate. For example, according to the Food and Drug Administration: “MedDRA classification is highly granular, with more than 23,000 Preferred Terms. When related PTs are not grouped together, it is possible to miss important safety signals.” As part of a 3-year initiative, we plan to leverage both clinical and statistical approaches, within a cross-disciplinary team, to develop optimal ways to interrogate, view, and present safety data to a diverse audience. The 3-year initiative will have these overarching themes: Clustering: to help with signal detection and characterizing the signals in safety data, Visual Analytics: To help see the signals and correlations better and faster, and Identify Appropriate Benchmarks for Aggregate Analysis: to establish the proper benchmarks that add to the efficiency, reliability, and robustness of the process. The methods we present here were developed following a learning and decision-making approach and will be fully adaptable in order to account for a variety of unprecedented factors that would ordinarily have a negative effect on the safety evaluation process, such as a pandemic. What we present today is the initial framework, objectives, and examples that will drive this initiative, next year, we will discuss the ideas we have developed and the following year, we will present on the implementation of those ideas.

P-23
THE STATE OF THE ART OF THE PRIOR ELICITATION METHOD IN CLINICAL TRIAL DESIGN AND ANALYSIS: A LITERATURE REVIEW

Primary Author:
DARIO GREGORI
University of Padua

Co-Author(s):
ILEANA BALDI
University of Padova
PAOLA BERRIALLA
University of Torino
DANILA AZZOLINA
University of Padova

In clinical trial design and analysis, the Bayesian inference is increasingly common. The subjective information obtained from an expert elicitation technique can be useful to define a prior probability distribution where there is no or minimal empirical evidence available. This research aims to explore the state-of-the-art Bayesian methods of prior elicitation with an emphasis on studies published in the clinical trials field. On 1 November 2020, a literature search was performed on the Current Index to Statistics, PubMed, and Web of Science databases, considering “prior elicitation” as a search string. The title and the abstract were manually screened to identify the prior elicitation pertinent articles. Summary statistics and publishing patterns over time have been represented. Finally, a Latent Dirichlet Allocation model was developed to identify latent topics among the pertinent papers retrieved. The algorithm was validated on the clinical trial prior elicitation pertinent articles. In addition, the overall accuracy was measured by comparing the manual and the automatic classification. Four hundred sixty documents pertinent to the Bayesian prior elicitation were identified. Of these, 213 (45.4%) were published in the “Probability and Statistics” area. Forty-two articles pertain to clinical trial and the majority of them (81%) reports parametric techniques as an elicitation method. Theoretical and Applied latent topics have been identified by the Latent Dirichlet Allocation algorithm. Among the 42 clinical trial papers, most of the literature on prior elicitation is characterized by an Applied (16%) rather than a Theoretical (4%) topic. However, the distribution of the Latent Dirichlet Allocation predictions on the overall 460 prior elicitation articles, according to publication years, revealed that the elicitation procedure is prevalently addressed in theoretical topic literature until 2010. The pattern is reversed in recent years evidencing an increasing interest in the elicitations methods also in the generally applied research. An increased interest in prior elicitation has been seen in the last decade and the distance between theory and its implementation is becoming narrower and narrower. Study results suggest that the diffusion of prior elicitation methods can be found in research fields other than theoretical may be linked to the recent rise in popularity of the Bayesian approaches in the general literature and clinical trial studies. However, the parametric elicitation solutions are the most widely applied especially in the clinical trial research. Given the promising versatility of non-parametric approaches to the elicitation of experts’ opinions, further efforts are required to ensure their diffusion, both at the design and analysis stage of a clinical trial.

P-24
WHAT IS THE (END) POINT?

Primary Author:
AMY H SULLIVAN
Columbia Mailman School of Public Health
Co-Author(s):
ALYSSA M VANDERBEEK
Institute of Cancer Research
BRIAN S HENICK
Columbia University Irving Medical Center
CODRUTA CHIUZAN
Columbia Mailman School of Public Health

Introduction: As of 1 January 2020, the Food and Drug Administration has approved 46 tyrosine kinase inhibitors for the treatment of neoplasms with hundreds more in clinical trial pipelines. Despite advances in precision medicine, responses to tyrosine kinase inhibitors can vary widely by patient population and challenges remain in characterizing drug efficacy, toxicity, and optimizing incorporation of these endpoints in clinical trials. Our comparative review of phase 1 and 2 trials conducted in this setting focuses on design characteristics that can affect the trial outcomes.

Methods: We conducted a PubMed search for publications of phase 2 clinical trials investigating tyrosine kinase inhibitors in oncology published between 1 January 2014 and 20 November 2019. We classified studies as meeting (successful) or not (failed) the primary endpoint and extracted information about trial aims, design, and outcomes. We further collected data on phase 1 studies that were cited in these phase 2 publications.

Results: Among 154 phase 2 trials, 114 were completed and 40 terminated early. The most common cancer types were thoracic (N = 63), genitourinary (N = 18), gastrointestinal (N = 18), gynecologic (N = 18), and breast (N = 15). Half of the completed phase 2 trials met their endpoint: 46 (40.4%) had a binary (objective response rate) endpoint versus 11 (9.6%) with a time-to-event (PFS) outcome. A total of 6968 patients were enrolled to failed phase 2 trials, compared to 4016 patients on successful trials. The majority (75%) of successful phase 2 trials were non-randomized (single or two-stage). Of the phase 1 trials cited by 44 completed phase 2 trials, 10 tested a different treatment regimen than was used in phase 2. 43.2% of the trials carried the identified maximum tolerated dose to phase 2. The remaining trials used either a lower dose (22.7%), higher (6.8%) or the maximum tolerated dose was not determined (27.3%).

Conclusion: The success rate of phase 2 tyrosine kinase inhibitor trials is near-akin to flipping a coin, with most patients enrolled on trials that do not meet their efficacy endpoint. Non-randomized and biomarker-restricted trials were more likely to meet their endpoints, with higher success rates for binary outcomes. Citing a previous phase 1 study had no obvious relation to phase 2 outcomes, mostly because of discrepancies in clinical setting application, dose usage, and patient heterogeneity. Further research is needed to optimize tumor types for phase 2 trial advancement. Novel endpoints incorporating biomarkers may provide opportunities for optimal dose selection and improved trial outcomes in later stages. Additional results and novel design suggestions will be presented.

P-25
PREDICTION OF PARTICIPANT SATISFACTION IN CANCER CHEMOPREVENTION CLINICAL TRIALS USING RANDOM FORESTS

Primary Author:
DAVID ZAHRIEH
Mayo Clinic

Co-Author(s):
RYAN P MCMURRAY
Mayo Clinic
PAUL J LIMBURG
Mayo Clinic
SUMITHRA J MANDREKAR
Mayo Clinic

Introduction: We used the “Was It Worth It” questionnaire to assess whether participation in early phase (0, 1, and 2) chemoprevention trials conducted by the National Cancer Institute-funded Cancer Prevention Network was satisfactory for generally healthy, cancer-free participants. Trial features and participant characteristics may be predictive of participant satisfaction. Knowing the set of features that predict satisfaction may inform design of subsequent trials and develop strategies to improve accrual, retention, adherence, and diversity.

Methods: The Was It Worth It questionnaire was administered at the end of each trial’s intervention period. The target population was intention-to-collect, defined as all enrolled participants. Satisfaction was defined as a participant who answered yes to the first three questions: Q1 = “Was it worthwhile for you to participate in this research study”; Q2 = “If you had to do it over, would you participate in this research study again”; and Q3 = “Would you recommend participating in this research study to others.” Participants who answered no or uncertain to any of these three questions were categorized as not satisfied; participants who did not complete the questionnaire were also categorized as not satisfied. Eighteen trial- and participant-level predictors were interrogated with the random forest algorithm using complete-case data to identify features predictive of satisfaction. The predictive model was trained on 75% of the sample and tested on the remaining 25%. Up-sampling and down-sampling methods were applied to address potential issues with class imbalance. Performance measures based on the
confusion matrix, out-of-bag error rates, and receiver operator curves were used to evaluate the two sampling methods with the original model prior to assessing the relative importance of each feature; importance was assessed by mean decrease in accuracy.

Results: Eleven trials targeting four disease sites (colorectum; esophagus; liver; lung) completed the defined study intervention period (median (range) = 3 (0.25, 13.25) months), comprising an intention-to-collect sample from 606 participants. One hundred twenty-three (25.5%) were of a racial or ethnic minority, 161 (26.6%) were female, and 69 (11.4%) were ≥65 years old. 473 (78.1%) participants indicated satisfaction, while 133 (21.9%) were not satisfied (46 did not complete the questionnaire). Five hundred forty (89.1%) participants had complete data on all 18 predictors and the distribution of each predictor was similar to the 66 (10.9%) participants who were excluded. Sampling methods did not meaningfully improve model performance. Model accuracy was 84.4% (95% confidence interval = 77.2%, 90.1%); agreement between observed and predicted classes, measured by the Kappa statistic, was 0.36; and area under the receiver operator curve was 0.66 (95% confidence interval = 0.54, 0.77) for the original model. The top three features predicting satisfaction were off intervention reason, participating site, and duration of intervention, while the bottom three features were age, sex, and number of comorbidities.

Conclusion: A novel application of random forests to a largely qualitative survey indicated that age, sex, and number of comorbidities were relatively unimportant in predicting satisfaction, while early termination, certain participating sites, and interventions of longer duration were associated with relatively lower participant satisfaction, corresponding to the top three features predicting satisfaction in our early phase cancer chemoprevention trials.

P-26
ANALYSIS FOR CLUSTER-RANDOMIZED CROSSOVER TRIALS WITH MULTIPLE PERIODS
Primary Author: SHUN FU F LEE
McMaster University
Co-Author(s): JESSICA SPENCE
McMaster University
STUART CONNOLLY
McMaster University
JIA WANG
McMaster University

SHRIKANT I BANGDIWALA
McMaster University

Cluster-randomized crossover trials are frequently used to evaluate healthcare policy, health system delivery, and other interventions only applicable in group settings. When the number of clusters is limited, one way to increase efficiency is to incorporate more crossovers (i.e. increasing number of periods) in the design by alternating between intervention groups within clusters. Little is known on how to analyze a cluster-randomized trials with multiple crossovers.

The aim is to determine the appropriate analytic approach for a cluster-randomized crossover trials with more than one crossover (i.e. >2 periods) using a simulation and to compare the empirical power obtained from the simulation with the estimated statistical power using the design effect approach extended to studies more than two periods. The PADIT study was used as an illustrative example. The simulation results show that the generalized linear mixed model with periods nested within clusters as the random effects is a robust approach to analyze this design even with multiple periods. The most optimal efficiency gain can be obtained by increasing number of periods from 2 to 8 with trials of large cluster size and substantial intraclass correlation.

P-27
WINDOW MEAN SURVIVAL TIME
Primary Author: MITCHELL PAUKNER
University of Wisconsin–Madison
Co-Author(s): RICHARD CHAPPELL
University of Wisconsin–Madison

We propose a class of alternative estimates and tests to restricted mean survival time which improves power in numerous survival scenarios while maintaining a level of interpretability. The industry standards for interpretable hypothesis tests in survival analysis, restricted mean survival time and logrank tests, can suffer from low power in cases where the proportional hazards assumption fails. In particular, when late differences occur between survival curves, our proposed estimate and class of tests, window mean survival time, outperform both restricted mean survival time and LR without sacrificing interpretability, unlike weighted rank tests. Window mean survival time has the added advantage of maintaining high power when the proportional hazards assumption is met, while weighted rank tests do not. With testing methods often being chosen in advance of data collection, window mean survival time can ensure adequate power without distributional assumptions.