ACCOUNTING FOR BASELINE PROGNOSTIC VARIABLES AND PATIENT DROP-OUT IN THE ANALYSIS OF LONGITUDINAL OUTCOMES WITHIN RANDOMIZED TRIALS FOR ALZHEIMER'S DISEASE

ELIZABETH COLANTUONI

BLOOMBERG SCHOOL OF PUBLIC HEALTH, JOHNS HOPKINS UNIVERSITY

Introduction: Consider a two arm regulatory trial where the primary outcome is measured at baseline and several fixed follow-up times. The primary endpoint is the change in the primary outcome from baseline to the final follow-up and the average treatment effect is the difference in the mean change comparing the treatment and control arm. Assume that a set of potentially prognostic baseline variables are collected and patient drop-out is expected. To estimate the average treatment effect, the mixed model for repeated measurement (MMRM) is the standard statistical approach. However, novel targeted minimum loss estimators (TMLE) proposed by Van der Laan and Gruber in 2012 can be applied to this setting and may offer reductions in bias and gains in efficiency relative to MMRM.

Objective: We compare the key statistical properties (such as bias, variance, mean square error and Type I error control) of MMRM and TMLE for estimating the average treatment effect when varying the prognostic ability of the baseline variables and the models generating patient drop-out.

Methods: We use data from a completed trial comparing donepezil to placebo to simulate hypothetical clinical trials for a drug that reduces the decline in cognitive impairment among persons with amnestic subtype mild cognitive impairment.

Results: It is customary to include the baseline value of the primary outcome and any stratification variables used for randomization in the MMRM model. However, including additional a priori selected baseline variables that are prognostic for the primary outcome can provide substantial precisions gains for estimation of the average treatment effect. Under certain models that generate patient drop-out, the MMRM and TMLE have similar statistical properties. When drop-out is differential across treatment arms, the TMLE approach provides a reduction in bias compared to MMRM.

Conclusion: In future trials with similar patients as those in the donepezil trial, adjusting for prognostic baseline variables will substantially improve the precision to estimate the average treatment effect. Specification of the statistical approach should include careful consideration of the possible mechanisms that generate patient drop-out.

Contributors

Michael Rosenblum
Bloomberg School of Public Health, Johns Hopkins University

Jon Steingrimsson
Brown University

Aidan McDermott
Bloomberg School of Public Health, Johns Hopkins University

Arnold Bakker
Johns Hopkins School of Medicine

Michela Gallagher
AgeneBio, Inc. and Department of Psychological and Brain Sciences, Johns Hopkins University

ASSESSING THE EXTERNAL VALIDITY OF RANDOMIZED CLINICAL TRIALS: INVERSE PROBABILITY WEIGHTING APPROACHES ON TYPE-2 DIABETES DATA

DANIELE BOTTIGLIENGO

UNIVERSITY OF PADUA

Introduction

Randomized Clinical Trials (RCTs) are the most accepted method to conduct causal inference in medical setting, since they can provide unbiased estimate of treatment effect for subjects included in the experiments. Even if RCTs play a key role to assess intervention efficacy, there's often uncertainty surrounding the true effectiveness of the treatment in a target population and concepts as generalizability and external validation of trial results have been overlooked most of the times.

This study focuses on the generalizability of RCT effectiveness on target population, following the ideas proposed by Cole and Stuart (2010) [1] and Stuart et al. (2010) [2]. The goal is to provide an application of a statistical methods able to detect representativeness of clinical trial sample of the general population.

Materials

Clinical trials data were acquired from a Prospective Randomized Open label Blinded Endpoint (PROBE) study [3] conducted to detect the effect of DPP-4 inhibitor sitagliptin with respect to conventional therapy in type-2 diabetes. Information about type-2 diabetes population in Japan were acquired from the report of National Health and Nutrition Survey (NHNS) of 2012 [4].

Methods

Generalizability of clinical trial is assessed with two methods. Both approaches allow for the estimation of treatment effects in the target population by weighting the effect observed in clinical trial for the inverse probability of being selected in the experiment. Difference between observed result in the trial and estimated result in the population gives a measure of external validity: the lesser the difference, the higher the generalizability of clinical trial results. First method uses summary statistics to compute individual weights, whereas second method calculates individual weights with a Propensity Score based approach.

Since summary statistics were the only available information from target population, individual data were simulated using a Monte Carlo based approach for joint data simulation proposed by Demirtas et al. (2016) [5], which allows for the creation of datasets when the marginal distributions and the correlation structures of individual characteristics are known.

Results

Percentage level of Hb1Ac after 2-years follow-up was the clinical endpoint considered. Clinical trial results show that sitagliptin decreases Hb1Ac % level (-0.161; 95% CI, -0.293; -0.030). Using weights computed with the first method stratifying by age and sex, treatment effect has the same direction but a different magnitude: (-0.165; 95% CI, -0.300; -0.030) weighting for age, (-0.224;

95% CI, -0.358; -0.091) weighting for sex and (-0.228; 95% CI, -0.293; -0.091) weighting for both age and sex. With Propesity Score based method, sitagliptin lower the percentage level of Hb1Ac with an increased magnitude (-0.376; 95% CI, -0.527; -0.224).

Conclusions

According to both first and second methods, clinical trial results underestimate the impact of sitagliptin on Hb1Ac % level in the target population. The main limitation of the study lies in the limited amount of available information on target population, which could lead to simulation of individual data and a calculation of the weights not fully reliable. In the future more information on the target population should be retrieved through literature review and access to survey data.

Contributors

Corrado Lanera
University of Padua

Clara Minto
University of Padua

Dario Gregori
University of Padua

Paola Berchialla
University of Turin

Ileana Baldi
University of Padua

BEHAVIORAL RCT USING FREQUENT SOCIAL INTERACTION TO ENHANCE COGNITIVE RESERVE

HIROKO DODGE

OREGON HEALTH & SCIENCE UNIVERSITY/UNIVERSITY OF MICHIGAN

Background: Past epidemiological studies have demonstrated that larger social networks or more frequent social interactions may have protective effects against cognitive decline and incident dementia. Therefore, increasing social interaction could be a promising intervention for improving the cognitive well-being of the elderly. However, only a few intervention studies have attempted to increase social interaction in an older population. This is largely due to several challenges in developing behavioral interventions of this kind. The challenges include: 1) selection bias in RCT participants (hard to recruit those with social isolation), 2) how to standardize and deliver social interaction reliably, 3) selecting or constructing outcomes which are sensitive enough to capture efficacy when delivered as a prevention study to those who are not yet demented (pre-symptomatic), for whom the expected decline or changes are likely to be limited, and 4) showing ecologically valid translational effects (not limited to efficacy in practiced areas).

Objectives: Based on our previous pilot study, we developed a randomized controlled behavioral clinical trial to examine whether conversation-based cognitive stimulation as delivered through frequent Internet enabled video conversations has a positive effect on general and domain-specific cognitive functions and translational effects among socially isolated older adults with either normal cognition or mild cognitive impairment (MCI). We will introduce our protocol in the conference, delineating how we addressed each challenge noted above.

Methods: Socially isolated participants aged 80 and older (Lubben's Social network scale ? 12 and other criteria) will be recruited from a local Meals on Wheels program in Portland, Oregon and Detroit, Michigan, the latter targeting African American participants. Face-to-face communications with trained interviewers using an internet enabled link to a study-dedicated home computer will be conducted 4 times per week for 6 months and twice per week for additional 6 months with a parallel control group. Sustainability will be assessed at 18th months from the baseline. Dual primary outcomes are cognitive function in memory and executive domains using two sets of neuropsychological test batteries for cross validation (National Alzheimer's Coordinating Center Uniform Data Set V3 and NIH-Toolbox computerized tests). A secondary outcome is the change in objective assessment of instrumental activities of daily living (IADL) including medication adherence monitored by an electronic pillbox (MedTracker). Exploratory outcomes include changes in language characteristics and brain structure using MRI. Personality of both interviewers and participants will be assessed by NEO Five-Factor inventory to examine underlying reasons of potential variably in efficacy by individuals. Standardization of interviewers will be aided by assessing pre-post changes in participants' affect measured by PANAS (Positive and Negative Affect Scale) and number of turns between participants and interviewers which took place during the conversational interactions. These language analyses will be conducted using ASR (Automated Speech Recognition) derived speech and language data from audio files captured during conversations.

Conclusions: User-friendly Internet communication programs may improve the feasibility and cost-effective execution of social-interaction-based clinical trials. The protocol could serve as a model for future behavioral intervention trials.

Contributors

Elena Goodrich
Oregon Health & Science Universtiy

Jacob Lindsley
Oregon Health & Science Universtiy

Benjamin Hampstead
University of Michigan

Jeffrey Kaye
Oregon Health & Science University

CHALLENGES WITH CONDUCTING A TRIAL OF OPIOID ADMINISTRATION DURING ANESTHESIA IN PEDIATRIC PATIENTS UNDERGOING CARDIAC SURGERY

ANDREW WABERSKI

CHILDREN'S NATIONAL MEDICAL CENTER

Background: Postoperative high dose opioid strategies can result in respiratory depression, ileus, opioid tolerance, and fluctuations in analgesia. Methadone administered during anesthesia has been shown to reduce postoperative opioid requirement in adults, however, little is known about the effects in children undergoing high risk cardiac bypass surgery. Furthermore, the pediatric population has a different set of requirements pre- and postoperatively related to patient safety and positive outcomes and the interpretation of adequate analgesia and control of agitation in the postoperative period is difficult in critically ill pediatric patients. Therefore, conducting a randomized, double-blinded controlled anesthesia trial in this specialized population poses its own set of challenges.

Methods: The objective of the Methadone vs. Fentanyl Administration on Postoperative Pain Control in Pediatric Patients Undergoing Cardiac Surgery (MFAPP-PP) study is to evaluate if administration of methadone vs. fentanyl in pediatric patients (2-8 years old, weight >6 kg) undergoing cardiac surgery improves postoperative opioid requirements for pain control during the first 24 hours, and to evaluate any potential side effects. Continuous pain assessments are measured postoperatively and daily and total opioid consumption are recorded. Respiratory support, inotropic support, and discharge times are included in the postoperative measurements.

Results: To date, 25 of the anticipated 152 patients have been enrolled in the study. Numerous families have denied research consent due to the overwhelming nature and unfamiliarity with pediatric cardiac anesthesia. Among those who enroll, several factors routine to pediatric anesthesia care are challenging to ascertain and collect consistently across patients during a research protocol. Examples include anxiety and agitation which are common postoperative sequelae for pediatric patients however difficult to distinguish from inadequate analgesia. Furthermore, clinical and parental assessments of pain, which routinely dictate the administration of opioids based on scaled scores, are difficult to ascertain. Additional challenges include tolerance to opioids administered in the newborn or infant period. We utilized a well documented postoperative balanced analgesic approach, including acetaminophen and NSAIDs, which may have additional varying degrees of effect and influence on patient's pain. Variability among anesthesia and intensive care providers results in additional challenges to proper standardization of the study protocol across patients, this includes postoperative extubation criteria, inotropic support, and discharge criteria.

Conclusion: The analgesic picture for pediatric patients undergoing congenital heart surgery is often difficult to diagnose and treat. We rely heavily on signs rather than symptoms for interpretation of pain control in children less than eight years old. Providing a well balanced anesthetic technique to include a long acting opioid may benefit analgesia and agitation for pediatric patients following cardiac bypass surgery. It is the goal of this trial to optimize opioid analgesic strategies for pediatric cardiac bypass surgery patients.

Contributors

Barbara Braffett
George Washington University

EXPLORING TREATMENT PREFERENCE IN TRIALS WHICH COMPARE VERY DIFFERENT INTERVENTIONS: A PRE-TRIAL QUALITATIVE STUDY

LORETTA DAVIES

UNIVERSITY OF OXFORD

Background:

Strong treatment preferences of both clinicians and patients have been shown to have a detrimental impact on recruitment, adherence to treatment allocation and ultimately the success of a randomised controlled trial (RCT). Such preferences are often particularly evident when the interventions being evaluated are markedly different, such as the comparison of surgery with a non-surgical approach (e.g. physiotherapy).

If we could identify whether treatment preferences exist before a trial starts recruitment, this could offer opportunities to address potential issues earlier and optimise trial conduct. We chose the pre-trial stage of an anterior cruciate ligament (ACL) deficiency management trial, as it exemplified a surgical and non-surgical comparison with potential for strong treatment preferences, to explore this further.

Methods:

Semi-structured qualitative interviews were undertaken with a purposive sample of 15 patients recently diagnosed with ACL deficiency. We explored the extent of patients' preferences for the treatment options available (surgical reconstruction and physiotherapy) and identified what factors influenced their development. Interviews were analysed using Interpretative Phenomenological Analysis. Supplementary analysis of the content of information available online about the treatment of ACL deficiency was also conducted to provide a more complete understanding of how the treatments are presented.

Findings:

Strong preferences for treatment were evident. Overall patients viewed surgery more positively than a non-surgical approach with physiotherapy. Surgery was seen to provide mechanical stability, reassurance and confidence to return to activities. Whereas, not having surgery created feelings of vulnerability and was seen to potentially place considerable restriction on ability to return to activities. The uncertainty as to whether physiotherapy would provide adequate joint stability and possible subsequent requirement for surgery was a concern for patients because of implications on time to return to sport. Patients' perceptions of treatment and expectations of outcome were influenced by several sources: experiences of others e.g. friends, professional athletes; how treatments were portrayed by healthcare professionals and information on the internet. Analysis of online information available to patients demonstrated an unbalanced and outdated presentation of treatment options following this injury.

Conclusion:

We identified that strong treatment preferences exist. This highlights the need to implement strategies to address this at a feasibility stage, before potentially contributing to recruitment difficulties. Sources that influence patients' preferences were also identified, which improved our understanding of the potential for unbalanced presentation of these treatments. Trialists may wish to consider exploring the extent of preference for treatments before they start recruitment, particularly in trials where the interventions being evaluated are markedly different, to enable issues specific to a particular trial to be identified and addressed.

Contributors

Francine Toye
Oxford University Hospitals NHS Trust

Jonathan Cook
University of Oxford

Andrew Price
University of Oxford

David Beard
University of Oxford

IDEAS IN RANDOMIZATION METHODS FOR COMPLEX CLUSTER-RANDOMIZED TRIALS

JODY CIOLINO

NORTHWESTERN UNIVERSITY

Introduction: In cluster-randomized trials, constraining randomization based on baseline cluster-level variables is a technique often employed to control covariate imbalance across arms. Methods of constrained randomization in cluster-randomized designs tend to focus on categorical baseline variables with simpler, two-armed, equal allocation designs. The Patient Centered Outcome Research Institute (PCORI)-funded Mothers and Babies study, comparing the effectiveness of clinicians and paraprofessionals to reduce disparities in perinatal depression (PI: Darius Tandon, PhD; Northwestern University), serves as an example of a complex cluster-randomized study requiring unique methods of constrained randomization. We sought to randomize 42 sites with unequal allocation across three arms (1:3:3; control: intervention 1: intervention 2) while ensuring balance across the three arms with respect to three continuous site-level variables.

Methods: Logistical and practical constraints required randomization in three "waves". To control imbalance in important baseline site-level variables across arms, we employed a modified constrained randomization scheme that (1) controls imbalance across waves of randomization in a unique manner, and (2) can account for site dropout between waves. The algorithm applies ideas of the modified constrained randomization (Nietert, et al. 2009) and minimal sufficient balance (Zhao, et al. 2012), and it adds unique components to account for waves of randomization.

There are three variables for which we chose to control imbalance at the site level: percent non-White clients as reported by the site, site yearly client volume, and population density of the site area. The general procedure involved:

1.   Enumerating a large subset of possible allocation schemes.

2.   Evaluating imbalance for each variable for each enumerated sequence.

3.   If the imbalance is acceptable according to some pre-specified criterion, then we save this scheme in a smaller subset of potential allocations for implementation

4.   Of those in the smaller subset, randomly select one sequence for use in the current study

We chose the p-value corresponding to the Kruskal-Wallis test as our criterion for "balance" in step #3 above. If the p-value for each of the three variables is larger than 0.30 for a given simulated allocation scheme, that particular allocation is deemed "acceptable".

Results: To account for dropout, we randomized a total of 45 sites, with 38 ultimately remaining in the study (6:16:16) after randomization. We were able to achieve adequate overall balance according to our criterion for each variable of interest. The Kruskal-Wallis test results for non-White percentage, yearly volume, and population density were p-values equal to 0.629, 0.603, and 0.858, respectively.

Conclusion: With this novel, yet intuitive approach to constrained randomization for complex cluster-randomized study design, we achieved comparable intervention arms with respect to these important baseline variables at the cluster level. Analytic methods and inferences will require additional considerations for covariate adjustment given the design; however, this approach will allow for comparison of intervention across similar sites, thereby allowing for relative efficiency in analyses and overall face validity. This approach may be further adapted to other multiple arm studies with a variety of covariate types: continuous, ordinal, and categorical.

Contributors

Jessica Jensen
Northwestern University

Alicia Diebold
Northwestern University

Kimberly Koloms
Northwestern University

Darius Tandon
Northwestern University

ADAPTABLE - A PATIENT-CENTERED, PRAGMATIC RANDOMIZED CLINICAL TRIAL

AJAR KOCHAR

DUKE CLINICAL RESEARCH INSTITUTE

Background: ADAPTABLE (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long-Term Effectiveness) is a PCORI-funded, patient-centered, pragmatic clinical trial. ADAPTABLE is actively enrolling 15,000 patients with known coronary artery disease to compare the effectiveness and safety of two daily aspirin doses (81mg vs. 325mg) for cardiovascular outcomes (all-cause mortality, myocardial infarction, stroke) and major bleeding. ADAPTABLE has several innovative clinical trial features involving: the information technology (IT) infrastructure, the patient recruitment and retention strategy, and the endpoint ascertainment approach.

Trial Innovations:

Information Technology: ADAPTABLE is the first pragmatic clinical trial using the PCORnet infrastructure, a network of networks consisting of clinical data, patient powered, and health plan research networks (CDRNs, PPRNs, and HPRNs). These networks transformed their local electronic health data into a common data model (CDM), which is an analysis-ready standardized dataset for clinical research. Patient eligibility for ADAPTABLE was determined using a computable phenotype (Figure 1), an electronic identification process using the study inclusion/exclusion criteria. The base computable phenotype was created by the study coordinating center, implemented at each site, and validated using manual abstraction by site investigators.

Patient Recruitment and Retention: Recruitment strategies for ADAPTABLE can be divided into: (1) low-touch (letters, emails, EHR-based messaging) and (2) high-touch methods (in clinic, phone). Each participating site has tracked approach methods and number of contacts to assess recruitment efficiency (# of randomized patients / # of approached patients). To date, a strategy of multi-modal, multi-touch approaches using a staggered schedule has proved most efficient in maximizing enrollment.

Eligible patients are provided individual access codes and directed to the web-based patient portal. The portal provides details regarding participation, assesses eligibility including contraindicated medications, and allows patients to digitally consent and randomize online (Figure 2).

ADAPTABLE aims to maximize patient recruitment and retention through strategic clinical and patient engagement. The "Adaptors" are a group of patient partners representing each CDRN (Figure 3); they serve as key stakeholders for all aspects of the clinical trial including trial and patient-portal design and involvement in the steering and executive committees. The Adaptors have helped develop recruitment materials and the quarterly participant newsletter. The participant newsletter engages enrolled participants by providing trial updates and fostering a community. Many sites have also adopted a clinician newsletter to enhance clinician engagement and disseminate trial updates.

Outcomes: ADAPTABLE is using a multi-pronged approach for endpoint ascertainment. CDMs are routinely queried to identify potential endpoints. Linkage to health insurance claims data (Medicare and private insurers) allows for ascertainment of endpoints that do not occur within the enrolling health system. Participants complete follow-up visits in the web-portal every 3 or 6 months where they report hospitalizations and potential endpoints. The call center at the Duke Clinical Research Institute facilitates completion of participant visits for individuals who either do not have access to the web-portal (non-internet) or forget to complete the visit. A validation plan has been approved to determine the positive and negative predictive values of non-fatal endpoints that will be formally adjudicated by a clinical events classification group.

Contributors

Matthew Roe
Duke Clinical Research Institute

Holly Robertson
Duke Clinical Research Institute

Russell Rothman
Vanderbilt University Medical Center

Robert Harrington
Stanford University School of Medicine

Bill Larsen
OneFlorida Clinical Research Consortium

Daniel Munoz

Vanderbilt University Medical Center

David Crenshaw

Vanderbilt University Medical Center

Kevin Haynes

HealthCore

Iben McCormick-Ricket

Louisiana Public Health Institute

Adrian Hernandez

Duke Clinical Research Institute

Schuyler Jones

Duke Clinical Research Institute

IMPLEMENTING A SCALABLE RISK-BASED MONITORING STRATEGY

ERIK DOFFAGNE

CLUEPOINTS

Each organization faces different challenges when it comes to implementing RBM (Risk-Based Monitoring) and Data Quality Oversight. There is no 'one-size-fits-all' approach that overcomes all the barriers, as each organization has specific influencing factors such as therapeutic area, existing processes and available technologies. This presentation will explore how organizations are scaling their RBM implementations to ramp up utlization and return on their investment.

The aim of the session is to share the lessons learned in context of RBM implementation. We'll review the people, the technology, and the processes, to present an overview of what can be expected as part of the implementation process.

Learning Objectives

Participants will learn:

- Best practices for RBM study planning and execution and key pitfalls to avoid

- How to draw a clear distinction between study risk assessment, operational KRIs, and quality oversight method

- How to integrate RBM technology within new operational processes

<u>Contributors</u>

IMPROVING IMPLEMENTATION IN WALES THROUGH 'THE FORUM': AN UPDATE

PAUL BROCKLEHURST

BANGOR UNIVERSITY

Background: To improve the value and sustainability of health services, researchers, practitioners and policy makers must understand how health service interventions work in their context. These are usually inherently complex and commonly comprise of a number of components that can act both independently and inter-dependently. In Wales, a number of projects are underway to bring policy-makers and health service researchers together to facilitate the design of trials, their fidelity and the uptake of trial findings. The overall aim is to align government, organisations and individuals to create a research and evaluation informed knowledge management community at the level of the whole country.

Responsibility for all aspects of health and health services are fully devolved to the Welsh Government. The underlying approach in Wales is known as Prudent Healthcare, a universal philosophy that has its heart principles of 1) co-production, 2) proportionality of care to need, 3) avoidance of harm whether caused by commission or omission and 4) reduced unwarranted variation. Value based healthcare is an important executive element of a prudent approach, enabling measurement of progress.

Aim: The aim of 'the Forum' is to create an ongoing common space between research and practice: to improve the use of research evidence in NHS practice whilst equally enabling NHS practice and policy-makers to influence the research agenda. It brings researchers and policy-makers together in Wales to enable an understanding of the following:

• The clinical and research opportunities that arise in a value and values led fully devolved national health service;

• Perspectives of the different stakeholders in health services research and NHS practice;

• Contribution that implementation research can make to span boundaries;

• Importance and role of co-production in making research more relevant and applicable to healthcare services and NHS practice;

• Challenges of changing healthcare services and NHS practice in Wales;

• How research and trials can lead the agenda in a 'prudent' healthcare system.

Methods: To facilitate this, an initial workshop was run in June 2017, followed by a further national meeting in October 2017. The next stage is a national conference in Spring 2018 entitled "Bringing health services delivery and research together in Wales". At this meeting opportunities and priorities for health services research will be identified, along with identifying the necessary processes to promote better boundary spanning. Consensus methods using electronic voting methods will be used to develop final consensus. Stakeholders will be asked to score each priority: 1 to 3 signifying a priority of limited importance, 4 to 6 important but not critical, and 7 to 9 critical.

Consensus will be reached when 70% or more of the respondents score the priority between 7 to 9 and fewer than 15% scoring it as 1 to 3. Writing groups will then be set-up to address these priorities and influence future health services research calls from the Welsh Government.

Discussion: Spanning across research, health services and policy, 'the Forum' has the potential to reduce the 'implementation gap' and improve the value and sustainability of health services in Wales using research evidence.

Contributors

Chris Jones
Welsh Government

IMPROVING RCTS IN INVASIVE PROCEDURES; A PROPOSED NEW DEFINITION OF INVASIVE PROCEDURES

SIAN COUSINS

BRISTOL CENTRE FOR SURGICAL RESEARCH, UNIVERSITY OF BRISTOL

Introduction

The past decade has seen a huge increase in activity and a marked shift in culture in surgical research and related clinical areas such as interventional cardiology and radiology. It has led to significant growth in funded studies, support from the Royal College of Surgeons of England, methodological advances and productive multi-centre collaborations. One of the remaining challenges is to establish a clear definition of the entire spectrum of invasive procedures, including surgery. This would allow, i) selection of relevant methods during trial design to overcome challenges in evaluating invasive procedures (e.g. recruitment, intervention complexity and operator expertise, outcome selection, measurement and reporting), ii) streamlined evidence synthesis when developing a research question, iii) accurate categorisation of research in this field to track activity and inform investment.

Currently there is no universally accepted clear definition of invasive procedures. Trials are therefore often grouped based on other factors such as their clinical area (e.g. oncology, cardiology) rather than the characteristics of the procedure itself, which is what defines the challenges of undertaking research in this field.

Methods

Systematic searches were conducted in Medline (OvidSP), Embase and CENTRAL databases for RCTs of invasive procedures. Sub-sets of articles were read, and a preliminary definition developed. All articles were re-read, and the preliminary definition applied to each one and then sequentially to the remaining papers. Where necessary, the definition was amended. All articles were read again to confirm that the final definition encapsulated the entire spectrum of invasive procedures.

Results

The proposed definition is based on 1460 articles retrieved by the search and iterations of the definition. It has three key components, i) access to the body, ii) need for instrumentation, iii) operator skill.

 "An invasive procedure is one where purposeful/deliberate access to the body is gained via an incision, percutaneous puncture, or instrumentation via a natural orifice. It begins when entry to the body is gained and ends when the instrument is removed, and/or the skin is closed. Invasive procedures are performed by healthcare professionals using instruments, which include, but are not limited to, endoscopes, catheters, scalpels, scissors, devices and tubes.

Where invasive procedures involve the administration of a medicinal product (e.g. drugs, gases, stem cells), these could be categorised as an 'invasive procedure' when operator skill is required for its administration, ie:

i.   an internal action is performed to administer the product,

ii.   the product is administered to a targeted anatomical area."

Discussion

This definition encompasses procedures that share commonalities. Crucially, the definition excludes pharmaceuticals, except for those where administration occurs within the invasive procedure and is dependent on operator skill, e.g. products delivered internally via the 'stop-flow' technique; this is important as the methodological challenges (e.g. blinding of trial persons) of these procedures are not aligned with those where a drug is administered, e.g. intravenously during a procedure.

The sheer number, costs and fundamental role of invasive procedures in healthcare means the benefits that application of this definition would provide are crucial, particularly at a time when research funding is precious.

Contributors

Natalie Blencowe
Bristol Centre for Surgical Research, University of Bristo

Jane Blazeby
Bristol Centre for Surgical Research, University of Bristol

INTEGRATING NEW SITE COORDINATORS IN AN ONGOING LONG-TERM MULTI-SITE OBSERVATIONAL TRIAL

BRIAN BURKE

GEORGE WASHINGTON UNIVERSITY

The TODAY2 study is a multi-site observational follow-up of the TODAY randomized clinical trial, which was designed to treat and assess the progression of type 2 diabetes in youth and began in 2004. The study represents the largest cohort of youth with type 2 diabetes ever assembled recruited at 15 clinical centers in the United States. The Coordinating Center (CoC), located at the George Washington University Biostatistics Center, is responsible for training, certifying, and integrating new site coordinators to the TODAY2 study. All new site coordinators are first required to attend an hour-long training webinar led by the CoC research assistants. The webinar includes information on the history and the results of the original TODAY study, the structure and procedures of TODAY2, and the internal organization of the study group. New coordinators are also briefed on the differences between research data-collection procedures and clinical practice. Following the webinar, coordinators are granted access to the study website and are directed to review and familiarize themselves with the study protocol, the manuals of procedures (MOPs), and the data collection forms. New coordinators are also encouraged to read through the TODAY2 Manual for New Clinic Staff which provides a quick reference to frequently asked questions. A certification quiz is administered to evaluate study knowledge and proficiency in the study protocol. Once completed satisfactorily, new staff are granted access to the study's proprietary data-entry website to begin training and certification exercises focused on entering data accurately and editing forms according to study guidelines. After successfully completing the data management certification criteria, coordinators gain access to the study's live participant database, enabling them to view and work with real participant data. Coordinators may also be certified for other study measurements such as collecting participant blood pressure as dictated by their site staffing needs. As a final step in the training process, an experienced coordinator and a CoC research assistant visit the new coordinator's site for a one-on-one in-person training to reinforce good practices and to build the lines of communication between the new coordinator, other site personnel, the CoC, and the TODAY2 study group. We will describe these training and certification processes in more detail as well as share recommendations for integrating new staff into an established study group. Effective on-boarding of new site coordinators is an important factor in ensuring research data is collected with the highest degree of quality, consistency, and integrity, and contributes to the success of the research study as a whole.

Contributors

Jennifer George
George Washington University

Mihili Gunaratne
George Washington University

Melinda Tung

George Washington University

MONITORING AND ADJUDICATION OF COMORBIDITIES IN A MULTI-CENTER OBSERVATIONAL STUDY OF YOUTH WITH TYPE 2 DIABETES

BRIAN BURKE

GEORGE WASHINGTON UNIVERSITY

The TODAY2 study is a longitudinal follow-up study to track the progression of type 2 diabetes and related comorbidities and complications in the TODAY cohort. Because comorbidities are a critical study outcome, a rigid adjudication process is used to track and document them using an online data management system. A potential comorbidity is first identified by the site physician and recorded on worksheet Medical History and Conditions (MEDHX). The site coordinator then completes the Medical Records Tracking Form (MEDTRK) to track the medical record release form and the acquisition of medical records for the comorbidity. The coordinating center starts an adjudication form, which documents the final decision on the potential comorbidity. In the comorbidity tracker, site coordinators, the coordinating center, and the Comorbidities Assessment Committee (CAC) are able to communicate using a blog with attention statuses to ask questions or request more information about the potential comorbidity. Once the site coordinator uploads the medical records to MEDTRK, the site requests that the coordinating center review the documentation from the site. After this review, the coordinating center requests an additional review from the CAC committee. The CAC committee consists of expert physicians who assess potential comorbidities according to the study's predetermined diagnostic criteria. To close the adjudication process, CAC confirms or does not confirm the diagnosis and records the date of disposition. To date, there are 457 potential comorbidities that have been adjudicated through the CAC tracker. In year one, as the tracker was first implemented, 50 potential comorbidities were adjudicated; in year two, 218 potential comorbidities; and in the first nine months of year three, 189 potential comorbidities. The meticulous process using forms MEDHX, MEDTRK, and the adjudication form in the database system, along with the interprofessional collaborations, ensures that the TODAY2 study will provide valid, interpretable analytic results.

Contributors

Jennifer George
George Washington University

Meghan Turney
George Washington University

MONITORING OF MEDICAL EVENTS AND ADJUDICATION THROUGH CUSTOMIZABLE DIGITAL WORKFLOWS IN THE DIABETES PREVENTION PROGRAM OUTCOMES STUDY (DPPOS)

ANNA BOWERS

GEORGE WASHINGTON UNIVERSITY

Adjudication of major clinical events remains an integral aspect of data collection, ensuring the validity, integrity, and timeliness of study results. As paper-based adjudications have proven to be unwieldy, digital and online-based systems allow for both increased efficiency and productivity. Thus, creating a workflow that allows for systematic and expeditious adjudications is essential. For this presentation, we will discuss the evolution in developing custom adjudication workflows through a checklist within an electronic data capture system. The automated checklist is implemented as a form that integrates data and manages progression through the workflow. Features of the system include automated email notifications chronicling advancement in the workflow, alerts for new or updated file attachments, and notice of posted comments; permissions for adjudicators to manage the workflow themselves through attention status assignment; and automatically-generated and continuously-updated custom reports with relevant event data for quick reference during adjudications. These features provide a convenient and efficient way for adjudicators to review relevant medical records, summary reports, and the ability to directly enter their review and adjudication. Further, the workflow system allows quick and easy monitoring of the status for each reported event. The workflow system is highly customizable and has been implemented for a wide range of adjudication and review processes. In the Diabetes Prevention Program Outcomes Study, the virtual workflow has facilitated the adjudication of over 120 reported cardiovascular events and/or deaths and the adjudication of over 100 reported cancer events within the past 6 months. In addition to describing the various workflows established, this presentation will also discuss how the same program functionality can be modified for other procedural workflows and the benefits of using a central data system to track these processes.

Contributors

Hanna Sherif
George Washington University

Ashley Hogan
George Washington University

Alla Sapozhnikova
George Washington University

Marinella Temprosa
George Washington University

NEW GUIDANCE ON SPECIFYING THE TARGET DIFFERENCE ("EFFECT SIZE") FOR A RCT

JONATHAN COOK

UNIVERSITY OF OXFORD

Central to the design of a RCT is a calculation of the number of participants required (the sample size) which provides reassurance that the trial will be informative. Conventionally, this is usually based upon a difference in the primary outcome (target difference) between groups that is desired to be detectable; the corresponding number of participants needed to be recruited is then calculated. From a scientific, practical and ethical perspective, selecting an appropriate target difference is very important. However, it has been neglected until relatively recently. A variety of approaches have been proposed and addressed by a large recent review and limited guidance in this area has been produced. However, there is a need for better guidance to aid researchers and funders. The DELTA2 study was commissioned by two UK funders (Medical Research Council and National Institute for Health Research) to improve guidance in this area. It involved update review of the literature and extensive engagement with stakeholders to inform the scope and contents of the new guidance.

This session presentation will provide a brief introduction into the topic area and the DELTA2 project. This will be followed by the new guidance on specifying the target difference in a randomised trial sample size calculation and accompanied by guidance upon reporting of target difference and sample size calculations.

Contributors

Steven Julious
University of Sheffield

Will Sones
University of Oxford

 DELTA2 group

NHANES AS A COMPARATOR COHORT

CLAIRE BOYLE

JAEB CENTER FOR HEALTH RESEARCH

Introduction/objective

The National Health and Nutrition Examination Survey (NHANES) is a series of studies designed to assess the health and nutritional status of United States citizens. NHANES provides a wealth of data on the prevalence of major diseases and risk factors for diseases, which are available online and through publications. Many researchers have utilized NHANES data as a comparator cohort, though questions arise as to whether NHANES was designed for this purpose and whether formal statistical comparisons with external data are valid. This discussion aims to explore NHANES as a comparator cohort.

Background

The NHANES uses a complex multistage, stratified, clustered sampling design. Data is not obtained by a simple random sample and certain characteristics are deliberately oversampled. The publicly-released data includes survey weights to adjust for over-sampling, non-response, and non-coverage. These weights adjust the survey data to better represent the national population and are required in all analyses. Failing to account for them may result in biased estimates and significance levels. Although NHANES provides detailed analytic guidelines, they omit the possibility of comparing NHANES with external data. No guidelines exist on how to conduct such comparisons.

Methods

A review of abstracts from published articles comparing external data to NHANES was completed using the PubMed database using the search keywords 'NHANES cohort comparison'.

Results

The PubMed search returned 62 abstracts comparing external data with NHANES. Twenty-nine limited comparisons to descriptive statistics, while 33 conducted formal statistical comparisons.

Twenty-seven abstracts (12 statistical comparisons) used group matching to ensure a balance in demographic factors between the cohorts. Most only matched on age and sex, though there is a possibility that the samples are mismatched for other measured or unmeasured, and possibly confounding factors. Selecting a matched NHANES sample does not necessarily eradicate the issue of its complex survey design. Estimates produced from this matched sample are not generalizable to the US population unless NHANES survey weights are used. Of these abstracts, only one accounted for sample weights and limited comparisons to descriptive statistics only.

Among the abstracts that conducted formal statistical comparisons, ten adjusted for confounding factors in analyses. However, none utilized NHANES weights. Twenty-five (11 statistical comparisons) provided no method to adjust for the imbalance of factors between cohorts. Of these,

two properly accounted for sample weights when analyzing NHANES data, both of which only produced descriptive statistics.

Conclusions

Formal statistical comparisons with NHANES are questionable. The NHANES documentation never seems to consider the use of NHANES as a comparator cohort despite researchers considering it as such. The majority of past comparisons, and notably all that conducted formal statistical comparisons, failed to account for the complex NHANES sampling design by properly weighting the data. This indicates the need for increased awareness for proper analysis of NHANES data, and further discussion and literature on correct conduct of statistical comparisons with NHANES.

Contributors

Craig Kollman
Jaeb Center for Health Research

Nicole Foster
Jaeb Center for Health Research

ON THE ACCURACY OF ONCOLOGY EXPERTS' PREDICTIONS ABOUT SAFETY IN RANDOMIZED CANCER TRIALS

DANIEL BENJAMIN

MCGILL UNIVERSITY

INTRODUCTION: Predictions about the safety of interventions undergoing clinical testing, and understanding uncertainty surrounding those predictions, are vital for ethical review, priority setting, and protecting patients in randomized trials. However, little is known about the skill of experts in predicting a treatment's risk profile in research settings. OBJECTIVE: To assess the accuracy of medical experts' predictions about the frequency of drug-related serious adverse events in trials within their medical specialty. METHODS: We identified 18 on-going late-phase, randomized cancer trials in three highly prevalent cancer indications (genitourinary cancers, lung, and colorectal). Experts in each malignancy were identified based on a record of recent publication in the indication, or affiliation with top cancer institutions. We designed objectively verifiable prediction questions by asking for subjective probability predictions about the likelihood the proportion of patients in each trial experiencing drug-related serious adverse events would exceed a certain safety threshold. Safety thresholds were determined by referencing early phase trials and conferring with experts specializing in each malignancy. Predictions were scored for accuracy using Brier Accuracy Scores (BAS; a measure of how far a person's predictions deviate from actual outcomes where 0 means always correct and 1 means always incorrect), calibration (the alignment between predicted probability and observed frequency of a positive trial), and discrimination (the ability to discern positive from non-positive trials). Expert skill was compared to random guessing and a non-discriminating strategy of guessing 50% each time. RESULTS: A total of 137 experts participated. Experts had a mean age of 48 years (SD=8.3), a mean h-index of 24.2 (SD=16.5), and reported moderate expertise about trials in the survey (4.7, SD=1.7, on a 7-point scale). Predictions did not show a clear pattern of optimism or pessimism about trials exceeding the safety thresholds. Distributions of predictions were close to uniform across the full probability scale (mean prediction = 46%, SD=.30, skew=.14), and 62% of trials in our sample exceeded the threshold. Two trials were terminated early - due to futility or recruitment - leaving them unscorable. The mean BAS was 0.29, 95% C.I.[0.26, 0.31]. Overall, experts did better than random guessing (compared to 1000 randomly generated datasets, mean BAS=.33), but fell short of a noncommittal strategy of always guessing 50% (BAS=.25). Experts showed modest calibration (Cal=0.08, where 0 is perfect), but showed little ability to discriminate (area under the ROC curve = 0.58 where .50 is expected from random guessing), and were about equally successful for events that exceeded and did not exceed the safety threshold. A minority of experts show high forecast skill; 58% had a positive correlation between their forecasts and the results, and 37% of experts outperformed the noncommittal benchmark. CONCLUSION: Experts showed minimal skill in predicting safety outcomes in this setting, despite the availability of some safety information from early trials. This could be interpreted to mean either that our safety thresholds represent maximal uncertainty or that experts are insensitive to safety thresholds. In either case, expert judgment adds little predictive value to review systems beyond relying on early phase safety outcomes.

Contributors

Jonathan Kimmelman
McGill University

David Mandel
York University

Ian Tannock
Princess Margaret Cancer Centre

Natasha Leighl
Princess Margaret Cancer Centre

Monika Krzyzanowska
Princess Margaret Cancer Centre

Tristan Barnes

Princess Margaret Cancer Centre

OPTIMISING THE DESIGN OF PLACEBO-CONTROLLED RCTS OF INVASIVE PROCEDURES: AN ANALYSIS OF PUBLISHED RCTS

SIAN COUSINS

BRISTOL CENTRE FOR SURGICAL RESEARCH

Introduction

Randomized controlled trials (RCTs) are considered the gold standard for informing evidence based healthcare. High quality RCTs, however, have been difficult to conduct in trials of invasive procedures; one reason for this is linked to the challenges associated with minimising bias. Whilst invasive placebo procedures are desirable there are unique practical considerations, compared to pharmaceutical placebos, and there is debate regarding their ethical acceptability to patients and healthcare staff. Pilot work assessing the practical feasibility and acceptability of RCTs, including testing of procedures and estimation of recruitment/retention, is considered crucial in the development and evaluation of complex interventions, such as invasive procedures.

Aim: To study in detail and summarise reported development of placebo procedures to inform placebo-controlled RCTs of invasive procedures.

Methods

Randomised trials, including pilot RCTs, comparing invasive procedures with placebo were identified from systematic searches of Medline, Embase, CENTRAL and clinicaltrials.gov databases (Wartolowska, BMJ Open, 2016). Trials were published between database inception and 14th November 2014. Additional trials and pilot RCTs were identified by hand searching references and expert knowledge. Trial protocols, summaries from trial registries and texts referenced in the report outlining methods or pilot work were also retrieved. Data on type of invasive procedure, how the placebo was developed and tested, and lessons learnt from the placebo design were extracted.

Results

Seventy-three placebo-controlled trials were included. Procedures was mostly endoscopic (n=34, 47%), with 28 (38%) standard surgical procedures with an incision and 11 (15%) via a percutaneous puncture. Ten pilot studies were retrieved, however the focus of 8 of them was assessment of the standard invasive procedure and not the design or delivery of the placebo procedure. Two pilot RCTs assessed feasibility of recruitment, however only one commented on the feasibility of delivering the placebo procedure specifically. This report also outlined in-depth development work, which included interviews with key stakeholders exploring placebo development and the acceptability of the placebo-controlled RCT. Authors in 15 (21%) of the studies reflected on the quality of the placebo procedure in their trial. This was limited to discussion of the differences/similarities between the placebo and invasive procedures and how this may have impacted the blinding of patients and investigators, with none commenting on the need for pilot work to optimise the development of the placebo procedures.

Conclusion

Given the practical challenges and ethical considerations in designing and delivering placebo invasive procedures, the scarcity of pilot work revealed by this review is surprising. Pilot work is an ideal opportunity to assess and improve the placebo, and test whether it can be detected as such. Increased investment in pre-trial development and pilot work is recommended to optimise the design and delivery of placebo invasive procedures and improve the quality of RCTs in this area.

<u>Contributors</u>

Alexander Gormley
Bristol Dental School

Natalie Blencowe
Bristol Centre for Surgical Research

Jane Blazeby
Bristol Centre for Surgical Research

RESPONDER ANALYSIS WITH MISSING DATA: COMPARISON OF TWO IMPUTATION TECHNIQUES IN THE CONTEXT OF RANDOMIZED CONTROLLED TRIALS

MELANIE BELL

UNIVERSITY OF ARIZONA

Background: The desire to frame randomized controlled trial results in terms of clinical relevance has led to the so-called responder analysis, where the proportion of patients in each arm who have "responded" by some pre-specified threshold are compared. This approach has been criticized by some methodologists due to the potential for misclassification bias and reduced power. Another issue is missing data. A common approach for handling patients who drop out is to count them as non-responders, which may result in bias. Likelihood based mixed models yield unbiased estimates when data are missing completely at random or at random.

Objective: We aimed to investigate bias and power of the responder analysis on incomplete data when imputing missing follow-up responses as non-responders compared to using mixed model best linear unbiased prediction (BLUPs) for imputation.

Methods: We simulated data from a two arm randomized controlled trial with outcome values at baseline and three subsequent time points. We varied the trajectory over time, the amount of missing data and the missingness mechanism. Missing data were imputed using missing = non-response and using the BLUPs, and the difference in proportion of responders was compared using a chi-square test. BLUPs were calculated by combining the fixed effect estimates with subject specific random effect estimates.

Results: We found that, in general, BLUP imputation outperforms the missing = non-response imputation in terms of bias and power, although both approaches suffered from some bias. When data were missing not at random neither approach worked well.

Conclusion: If trialists want to use a responder analysis, a good choice for handling incomplete continuous data is to model the outcome with a mixed model and use the BLUPs to compute the proportion of responders.

Contributors

Lysbeth Floden
University of Arizona

RUN-IN PERIODS IN RANDOMIZED CONTROLLED TRIALS OF CHRONIC DISEASES: A META-EPIDEMIOLOGIC STUDY

DAVID COLLISTER

MCMASTER UNIVERSITY

Background: A run-in period may improve trial efficiency by removing participants at high risk of dropping out or being non-adherent to an intervention prior to randomization. However, how run-in periods are used and how they are designed is not well characterized.

Objectives: To systematically review and summarize the characteristics of trials that used run-in periods in randomized controlled trials (RCTs) of chronic diseases.

Methods: RCTs between 2011 and 2016 published in the four highest impact general medical journals were identified using the RCT MeSH term. RCTs were eligible for this review if they included participants 16 years of age or older with a chronic disease and the interventions of interest included an oral, subcutaneous or inhaled pharmacologic therapy administered at least daily for at least 6 months. Two reviewers abstracted trial characteristics using a standardized form in parallel.

Results: 217 RCTs were included of which 37 (17.1%) included a run-in period of which 3 included more than one run-in period. The median duration of a run-in period was 28 days (interquartile range 14-57.75) and 23 (62.2%) used active therapy during the run-in period. The most common reasons for a run-in period were to exclude non-adherent participants (n=17), or participants intolerant to the study treatment (n=13), or to ensure that eligibility criteria were met (n=15). Of the 24 (64.9%) studies that reported run-in completion rates, the mean run-in completion rate was 77.1% (standard deviation 16.2%).

Conclusions: Run-in periods are infrequently used and their impact on excluding participants unlikely to comply with or tolerate trial therapies is poorly reported, even in trials of therapies for chronic diseases where they may greatly increase trial efficiency.

Contributors

Jennifer Rodrigues
McMaster Universiy

PJ Devereaux
McMaster University

Gordon Guyatt
McMaster University

Michael Walsh
McMaster University

SITE ACTIVATION PROCESSES PREDICTIVE OF ENROLLMENT SUCCESS IN A PRAGMATIC PRECISION MEDICINE STUDY

KAITLYN ANTONELLI

AMERICAN SOCIETY OF CLINICAL ONCOLOGY

The American Society of Clinical Oncology's (ASCO) Targeted Agent and Profiling Utilization Registry Study (TAPUR) is a non-randomized, pragmatic precision medicine basket trial. TAPUR offers eligible patients with advanced cancer access to U.S. Food and Drug Administration (FDA)-approved, targeted drugs in non-indicated cancers to describe safety and efficacy outcomes. TAPUR launched in March 2016 at 4 cancer centers, comprising 37 clinical site locations and is now available at 26 centers, comprising 87 locations.

Sponsors often consider timeliness and completion of key steps in the site activation process as important predictors of time to first participant enrolled and overall site enrollment. In our experience to date, early identification and retention of a primary study coordinator in the site activation process is a strong predictor of timely completion of site activation activities and may be the proximate cause.

Turnover or transition of the study primary coordinator role during the activation process may cause disruptions in key activation steps such as study trainings and coordination of local activities (e.g. pharmacy processes, regulatory documentation, etc.) resulting in delays in activation. We examined time to site activation and time to first participant enrolled among 26 TAPUR study sites. Sites with consistent primary study coordinators through the activation process were nearly 1.5 times faster to activate after obtaining IRB approval, with an average of 65 days between obtaining IRB approval and site activation, compared to approximately 95 days for sites that did not identify or retain a primary study coordinator.

Sites with identified and consistent primary coordinators throughout the activation process were 3 times faster to enroll their first participant taking an average of 33 days post-activation, compared to 110 days for sites that did not identify or retain consistent primary coordinators. These sites also had approximately 6 times higher total accrual to the study in their first year of participation, enrolling an average of 35 participants in their first year, compared to an average of 6 participants for sites that did not have a consistent primary coordinator. Furthermore, sites with identified and consistent primary coordinators experienced a higher screen to enrollment rate with 62% of screens enrolling into the study, compared to 49% for sites that did not have a consistent primary coordinator.

Overall, turnover in the primary coordinator role during the activation process resulted in a cascading effect resulting in delays in activation and first participant enrolled, lower enrollment in the first year, and a lower screen to enrollment rate. Sites with turnover may be less likely to have defined participant identification processes, efficient mechanisms for relaying key information to staff, and a single point of contact responsible for not only the coordination of all local site activities (e.g. for legal purposes, regulatory submissions, pharmacy and drug distribution, and participant enrollment), but also for interacting with the sponsor.

The presentation will describe and discuss factors identified in the TAPUR site activation process that may be predictors of clinical site performance, focusing on the impact of primary coordinator turnover or transition as the key factor.

Contributors

Pam Mangat
American Society of Clinical Oncology

Suanna Bruinooge
American Society of Clinical Oncology

Richard Schilsky
American Society of Clinical Oncology

SITE INVESTIGATOR VS. SPONSOR SAE CAUSALITY: ARE THEY DIFFERENT?

HYUNSOOK CHIN

RHO, INC

A Serious Adverse Event (SAE) is any undesirable experience associated with a medical product in a subject where the subject outcome is any of the following: death; life-threatening event; hospitalization (initial or prolonged); disability or permanent damage; congenital anomaly/birth defect; required intervention to prevent impairment or damage; or other important medical event. When a SAE occurs the site reports the event to the sponsor and the site investigator determines the causality assessment (e.g. rating scale to a question such as "Was there a reasonable possibility that the drug caused the adverse event"). Once the sponsor receives the SAE report they also assign a causality assessment. In 2012, the Food and Drug Administration (FDA) updated their 'Safety Reporting Requirements for INDs and BA/BE Studies' to use the sponsor causality assessment to determine if a report needs to be reported to the FDA expeditiously. This is because the FDA believes the sponsor can better assess causality as they have access to SAE reports from multiple sites and studies along with a familiarity with the drug's mechanism of action. The causality assessment between the site investigator and sponsor can differ because the assessments are dependent solely on their own judgments.

To better understand the extent of the differences between site investigator and sponsor causality assessments we extracted data from the Argus safety database from multiple studies and sponsors. We included variables for preferred term, system organ class, site investigator causality, and sponsor causality. We will present a summary of our findings to better understand the differences.

We originally undertook this analysis for two reasons. First, we wanted to understand the differences as some of our studies have stopping rules written into the protocol that rely on causality (e.g. any occurrence of a life-threatening or fatal SAE that is related to the investigational agent). Originally site investigator causality was used when assessing the stopping rules because their assessment is done at the time of SAE reporting or soon afterwards. However, we wanted to quantify the difference with the sponsor reporting to understand if the assessment of stopping rules should rely on the sponsor assessment or some adjudication of the two when the differ.

Secondly, when reporting SAEs to the data and safety monitoring board (DSMB) the statistician needs to determine which causality assessment should be used.

The information presented can be used by study teams to inform their use of SAE causality in reporting (e.g. DSMB reports) and their use in stopping rules.

Contributors

Heather Kopetskie
Rho, Inc

STREAMLINING THE PROTOCOL OVERSIGHT PROGRAM FOR MAXIMUM EFFICIENCY IN THE DIABETES PREVENTION PROGRAM OUTCOMES STUDY (DPPOS)

TSEDENIA BEZABEH

GEORGE WASHINGTON UNIVERSITY

The conduct of clinical trials requires considerable surveillance to ensure data quality, protocol adherence, compliance to treatment interventions, and retention of participants. The Diabetes Prevention Program Outcomes Study (DPPOS) cohort has now been followed for over 18 years and consists of over 2500 participants at 25 clinics. A trial of this size and length requires substantial oversight to ensure data quality, compliance to study protocol, and retention of participants. The performance of DPPOS clinical centers is overseen by a Protocol Oversight Program (POP) committee consisting of clinical site Principal Investigators, Program Coordinators and Coordinating Center (CoC) staff. The committee monitors the performance of sites by examining data entered by the clinics and summarized by the CoC. To assist both POP and clinics in monitoring performance, clinic report cards ("POP cards") are issued twice each year, while time trend and comparison data are regularly updated and available on the secure DPPOS study website. POP cards include clinic outcomes completion, clinic operations issues (e.g. outstanding medical records), compliance with interventions (e.g. study medication adherence), retention issues (e.g. inactive, missed visits) and staff certification. The data for the items identified on the POP card are automatically extracted from the study reports and populated on the clinic's POP card. Additional detailed information of the POP card data is provided to clinics to help reconcile with their data, along with a key to interpret each of the values. To ease the data monitoring process, POP cards are automated HTML reports linked with a discussion board thread, which allows for a streamlined review process. Using this streamlined reporting process, preparation time and review time are minimized. The online review and message board allows real time discussion between key stakeholders to identify and troubleshoot issues, thus ensuring that clinics are operating at optimum standards.

Contributors

Hanna Sherif
George Washington University

Adrienne Gottlieb
George Washington University

Sharon Edelstein
George Washington University

Marinella Temprosa
George Washington University

TREATMENT PREFERENCE IN PEDIATRIC RANDOMISED CLINICAL TRIALS: SYSTEMATIC REVIEW AND QUALITATIVE SYNTHESIS

LUCY BEASANT

UNIVERSITY OF BRISTOL

Background: Recruitment to randomised clinical trials (RCTs) can be challenging. Patient preferences for treatment can hinder recruitment, prevent trial completion, and lead to post randomisation dropout. An understanding of the factors which influence preference for treatment is important if effective strategies to improve and facilitate recruitment and informed consent are to be developed. Existing systematic reviews have largely focused on preferences within adult RCTs and little is known about preference in pediatric trials.

Objective: To identify and synthesize qualitative and quantitative data in relation to treatment preference in pediatric RCTs.

Methods: Databases searched: CINAHL, EMBASE, MEDLINE, Cochrane Library and Cochrane Central Register of Controlled Trials (1950 – 2014). Studies inviting young people aged 0–17 years to participate in an RCT where a treatment preference was reported were included. Data extraction and quality checking were completed by two reviewers independently, discrepancies were resolved with a third reviewer. Synthesis of data drew on meta-ethnograpy. A reciprocal translational analysis (RTA) was undertaken to translate 'concepts' from individual papers into one another to develop evolved overarching 'constructs'.

Results: Forty-eight papers were included from 16,938 identified - 38 had descriptive data relating to treatment preference (reported via consort diagrams) and 10 had qualitative data (parent or participant quotes) for inclusion in the meta-synthesis. Descriptive data: 26 papers reported the number of families who declined the trial because of a specified treatment preference (range 2-74%), four reported the number of families who withdrew post randomisation because of a treatment preference (range 2-10%). Eleven publications reported RCTs with preference arms from the outset; two trials introduced preference arms after commencement because of high refusal rates. Declining accrual rates and a loss of clinical equipoise led to the closure of one RCT, and two required extensions because of slow recruitment. Meta-ethnography and synthesis: Qualitative data from 10 papers resulted in two overarching constructs: 1. Making sense of an RCT and parents' questions 2. Motivations and barriers to taking part in an RCT. These two constructs had six inter-linked sub-themes: i) Understanding of trial processes, ii) Understanding of treatment arms, iii) Perceived benefit, iv) Perceived risk, v) Access to treatment, and vi) Management of condition/practical implications. Parents' understanding of trial processes (e.g., randomisation and equipoise) impacted their perceptions of the benefits and risks associated with treatment arms, access to treatment, and long-term management of the condition in question. All of these factors contributed to the expression of treatment preference, and in some cases resulted in families declining RCT participation for their child.

Conclusions: Treatment preference can be a barrier to recruitment to pediatric RCTs. In some cases, this resulted in the need to change the design of the trial (introduction of preference arms), extend recruitment, or close the trial prematurely. Parents' understanding of trial processes are important

factors which underpin the context in which treatment preferences are constructed and emerge. Further investigation is needed to understand the impact treatment preference has on retention and outcomes in pediatric RCTs.

Contributors

Nicola Mills
Bristol Medical School, Population Health Sciences, University of Bristol

Roxanne Parslow
Bristol Medical School, Population Health Sciences, University of Bristol

Amberly Brigden
Bristol Medical School, Population Health Sciences, University of Bristol

Harry Apperley
Bristol Medical School, Population Health Sciences, University of Bristol

Tiffany Keep
Bristol Medical School, Population Health Sciences, University of Bristol

Antonia Northam

Bristol Medical School, Population Health Sciences, University of Bristol

Charlotte Wray

Bristol Medical School, Population Health Sciences, University of Bristol

Ryan Langdon

Bristol Medical School, Population Health Sciences, University of Bristol

Esther Crawley

Bristol Medical School, Population Health Sciences, University of Bristol

Hayley King

Bristol Medical School, Population Health Sciences, University of Bristol

Bridget Young

Bristol Medical School, Population Health Sciences, University of Bristol

VARIABILITY ADJUSTED TIME TO EVENT (VATE) METHOD FOR ANALYSIS OF RANDOMIZED CLINICAL TRIALS IN EPILEPSY AND OTHER EPISODIC DISEASES

DANIEL GOLDENHOLZ

BETH ISRAEL DEACONESS MEDICAL CENTER

OBJECTIVE: To develop a time-to-event method for epilepsy randomized controlled trials (RCTs) analysis that accounting for variability in seizure frequency during the baseline period.

BACKGROUND: Traditional epilepsy trial analysis is inefficient and costly. A proposed time-to-event method (French et al 2015) could decrease cost and boost efficiency, however it does not account for variability in seizure frequency. It has recently been shown that such variability is sufficient to reproduce typical placebo responder rates (Goldenholz et al 2017a).

METHODS: Artificially simulated clinical trials (generated with negative binomial distribution) were analyzed using time-to-event (TE) and variability adjusted time-to-event (VATE). VATE expands each trial into 100 trials sampled from the putative distribution of the baseline, and runs a mixed effects survival analysis on the result with trial as random effect, using the robust score statistic as the output. 1000 trials were analyzed for each set of parameters. The following parameter sweeps were conducted: number of patients (100,200 and 300), "size" a variability parameter (0.01,0.1,1,10), and drug strength (0% and 30%). Each trial randomly assigned drug to half the patients and placebo to the other. Drug and placebo are simulated using previously published methods (Goldenholz et al 2017b). Power was defined as proportion of the 1000 trials (among those with drug=30%) that had $p<0.05$ (i.e. trials that reject the null), and type 1 error rate was defined as the analogous proportion when drug=0%.

RESULTS: 24,000 trials were simulated using a supercomputer. For low variability (i.e. "size">0.1), TE has higher power to distinguish drug from placebo compared with VATE. For higher variability, VATE has higher power. Both techniques have comparable type 1 error rates. Figure 1 shows power and type 1 error for both methods.

CONCLUSIONS: When RCTs include patients with sufficient variability in seizure frequency, the VATE method is statistically superior than TE, and may allow for trials to have less patients and therefore lower cost while maintaining statistical power. Similar conclusions are likely relevant in other episodic diseases, such as cardiac arrhythmias, heart failure decompensation, diabetic hypoglycemic events, syncope, multiple sclerosis, migraines, lower back pain, etc.

1. French JA, Gil-Nagel A, Malerba S, Kramer L, Kumar D, Bagiella E. Time to prerandomization monthly seizure count in perampanel trials: A novel epilepsy endpoint. Neurology. 2015 May 19;84(20):2014-20. PMID: 25878175.

2. Goldenholz DM, Strashny A, Cook M, Moss R, Theodore WH. A multi-dataset time-reversal approach to clinical trial placebo response and the relationship to natural variability in epilepsy. Seizure. 2017 Oct 23;53:31-36. PMID: 29102709.

3. Goldenholz DM, Tharayil J, Moss R, Myers E, Theodore WH. Monte Carlo simulations of randomized clinical trials in epilepsy. Ann Clin Transl Neurol. 2017 May 24;4(8):544-552. PMID: 28812044.

<u>Contributors</u>

'IT'S TRYING TO MANAGE THE WORK': A QUALITATIVE EVALUATION OF RECRUITMENT PROCESSES WITHIN A UK MULTICENTRE TRIAL

KATIE GILLIES

UNIVERSITY OF ABERDEEN

As with many trials, recruitment within the TISU trial proved more difficult than had been anticipated at the outset and some sites did not meet their target expectations. The variability across sites relating to aspects of recruitment suggested nuanced differences in the processes relating to recruitment and we set out to explore trial site staff's perceptions regarding barriers and facilitators to local recruitment. For the purposes of the TISU trial, the aim was to identify trial-specific modifiable factors that could enhance the facilitators and remove the barriers to recruitment.

In this presentation we will present the findings of an in-depth qualitative investigation of trial teams perspectives on the barriers and facilitators to trials. Insights gained draw attention to the initial and ongoing burden of trial work that is involved throughout the duration of a clinical trial. In terms of building and sustaining a research culture, trial staff described the ongoing work of engagement that was required to ensure that clinical staff were both educated and motivated to help with the process of identifying and screening potential participants.

Having adequate and sufficient organisational and staffing resources was highlighted as being a necessary prerequisite to successful recruitment both in terms of accessing potentially eligible patients and being able to maximise recruitment after patient identification. The nature of the research study design can also potentially generate challenging communicative work for recruiting staff which can prove particularly problematic and adversely affect recruitment.

This work adds to existing research highlighting the importance of the hidden and complex work that is involved in clinical trial recruitment. Those designing and supporting the operationalisation of clinical trials must recognise and support the mitigation of this 'work'. Those designing interventions to improve recruitment could usefully widen their scope to consider not only patient-level factors but factors that mitigate the work required of staff. While much of the work is likely to be contextually sensitive at the level of local sites and for individual trials, some aspects are ubiquitous issues for delivery of trials more generally.

Contributors

A NOVEL RANDOMIZED PHASE 2 TRIAL DESIGN THAT CONSTRAINS SAMPLE SIZE BY REQUIRING A SUFFICIENT EXPERIMENTAL RESPONSE RATE

BRIAN EGLESTON

FOX CHASE CANCER CENTER, TEMPLE UNIVERSITY HEALTH

One difficulty with using one-sample Phase 2 designs, such as the Simon 2-stage design, is that a historical response rate might be difficult to specify. As a result, clinicians have been advocating in the literature that randomized Phase 2 designs be used instead. A further difficulty in using randomized designs in the Phase 2 setting is that the sample size required is often large. In 2017, Litwin et al. proposed a novel two-sample design that is a hybrid of the Simon one sample design and the traditional randomized design. In order to declare the trial a success, the experimental response rate has to be large enough (i.e. the one-sample component) and the treatment effect has to be large enough (i.e. the randomized component). Consider a two arm trial. Let E represent a random variable of the number of responses in the experimental arm and C represent the number of responses in the control arm. The criteria for declaring the study a success could be that $E>e$ and $E-C>d$. In a single stage design, the probability of declaring the study a success (i.e. power) would hence be $P(E>e)P(E-C>d|E>e)$ under a presumed hypothesis. One can optimize the design and minimize the sample size by judiciously setting both $P(E>e)$ and $P(E-C>d|E>e)$. That is, choosing both a desirable new treatment response rate and a promising preliminary treatment effect can provide a sensible study design with a reasonable sample size. The approach has been extended to a two-stage setting with early stopping.

Contributors

Samuel Litwin
Fox Chase Cancer Center, Temple University Health

Stan Basickes
Fox Chase Cancer Center, Temple University Health

Eric Ross
Fox Chase Cancer Center, Temple University Health

EXPLORING RETENTION IN CLINICAL TRIALS: A META-ETHNOGRAPHIC SYNTHESIS OF STUDIES REPORTING PARTICIPANT REASONS FOR DROP OUT

KATIE GILLIES

UNIVERSITY OF ABERDEEN

Background

Randomised controlled trials are integral for evidenced based clinical decision making. Issues around retention (i.e. ensuring that trial participants remain in the trial to provide primary outcome data), especially those issues reported by trial participants, have not received the same scrutiny in the literature as trial recruitment. This is a mistake. Poor retention is just as important for trial validity and is quite capable of fatally undermining a trial. Our aim was to undertake a meta-ethnographic synthesis of findings from primary qualitative studies that have explored factors influencing trial participant drop-out.

Methods and Findings

A systematic search of Embase, Ovid MEDLINE, PsycINFO, Cochrane CENTRAL, SSCI, CINAHL and ASSIA covering papers published from 1946 to August 2016 was conducted. Meta-ethnography was utilised to synthesise findings from eligible papers that contained qualitative data from trial non-retainers.

We identified 8 qualitative studies reporting data from 9 trials. The studies were undertaken between 2008 and 2015. Each study included between 3 and 40 people who had dropped out from a trial, with findings from a total of 137 people reported across the papers. Emergent from our synthesis was the significance of trial non-retainers' perceptions around the 'fit' of key aspects of the trial with their personal beliefs, preferences, capabilities or life circumstances. These related to: their own health state; preferences for receiving trial 'care'; individual capabilities; beliefs about or experiences of trial medication; and considerations about whether trial participation could be accommodated (or not) into their broader lives.

All of the factors raise important issues around the extent to which initial decisions to participate in the trial were fully informed. Our synthesis found only 8 eligible papers reporting findings across 9 trials, 4 of which had a mental health context and all of which were conducted in high-income countries. Studies in other disease areas are warranted to understand the extent to which all or some of these factors are shared across trials and which are more disease or trial-specific

Conclusions

Trialists need to reduce trial burden those taking part, both in terms of the intervention itself and also the ways that follow-up data are collected. Providing more detail on the nature of the trial interventions and what can be expected by 'participation' at the consenting stage may prove helpful in order to manage expectations. Early and meaningful patient/public involvement could be particularly important for accommodating future trial participants' preferences and capabilities. Taken together these findings could be considered during the design and delivery of follow-up procedures in trials and contribute to improving retention before it becomes a problem.

Contributors

Rumana Newlands
University of Aberdeen

Zoe Skea
University of Aberdeen

HOW DO TRIALISTS GO ABOUT DESIGNING THEIR RECRUITMENT STRATEGIES? A QUALITATIVE INTERVIEW STUDY WITH TRIAL STAKEHOLDERS

HEIDI GARDNER

UNIVERSITY OF ABERDEEN

Objectives

Trial recruitment is a challenge and many trials struggle to meet it. Currently, the literature around recruitment planning is limited; research has largely focussed on how the recruitment process functions once the process has started. Our aim was to understand if people involved with the process of participant recruitment to trials have explicit strategies for recruitment, if so, how they go about this process, and if not, what are the barriers to performing effective planning.

Design

One-to-one qualitative semi-structured interviews with a range of stakeholders. A framework analysis approach was used to analyse data, anticipated and emergent themes were identified, defined, and linked through constant comparison of data both within and across stakeholder groups.

Participants

23 trialists; 11 identifying primarily as 'Designers'; those responsible for the design of recruitment methods, and 12 identifying primarily as 'Recruiters'; those who actively recruit participants to trials. Participants spanned various roles including; Trial Manager, Research Nurse, Principal Investigator, Researcher, and Family Doctor. Participants' experience with trial intervention types, therapeutic indications, and level of recruitment expertise were diverse.

Setting

UK National Health Service sites that are actively involved in trials, academic institutions (UK, the Netherlands, Canada) and Clinical Research Organisations partnering with pharmaceutical companies (UK, South Africa, Italy).

Results

To varying degrees, respondents had strategies for recruitment, and two main themes emerged in relation to planning and development of these strategies: 1) content and timing of grant applications and trial protocols and 2) research governance practices and how they impact on recruitment in terms of increased work and timing delays.

Participants emphasised the need for effective planning, with many commenting that time-pressures as a result of tight deadlines for funding calls at the beginning of a trial often result in rushed planning and a so-called 'amendment cascade' once recruitment has started.

In relation to research governance, participants understood the need for rigorous processes, but having risk-appropriate procedures with a proportionate level of work, particularly in terms of

administrative burden on staff, was highlighted as an area that requires improvement. Many participants explained that interpretation of governance frameworks varied by institution. These participants went on to express that getting to know the nuances of the systems was an advantage that would allow them to build flexibility into trial protocols.

Further themes relating to the types of environments needed to facilitate successful recruitment emerged throughout the interviews; communication and relationship-building, both between individuals in a team, and across teams involved in multi-centre trials; support with recruitment, and the workload pressure experienced by both Recruiters and Designers.

Conclusion

Our respondents considered short grant preparation times and disproportionate approvals processes as major challenges to developing effective recruitment strategies. Poor initial planning is a mistake that trial teams then live with throughout the trial. In particular, poor planning underestimates the additional work the trial requires of site and trial staff. More efficient and effective recruitment requires strategies to increase the time available for proper trial planning.

This work is part of the Trial Forge initiative to improve trial efficiency.

Contributors

Shaun Treweek
University of Aberdeen

Katie Gillies
University of Aberdeen

IDEAL: FRAMEWORK AND RECOMMENDATIONS FOR EVALUATION OF SURGICAL INNOVATIONS

ASHA KHACHANE

MAIMONIDES MEDICAL CENTER

Background: Evaluation of new surgical innovations is complex. Factors unique to new procedures, such as iterations in procedural technique, operator learning curves, and surgeon and patient preferences, must all be considered. Historically there have not been well-established guidelines for developing and researching a new surgical method, and there is no minimum level of evidence required before a new surgical procedure is deemed safe for use. This is in stark contrast to the pharmaceutical industry, which has clearly outlined stages for testing and conducting research for all new drugs. Without a strong foundation of evidence, surgical procedures are often adopted by practitioners based on training experience, personal preference, and retrospective studies. The IDEAL (Idea, Development, Exploration, Assessment, and Long-term follow-up) Framework and Recommendations were developed to address these shortcomings and improve the quality of surgical research. The Framework describes the evolution of surgical techniques in five stages with clearly identifiable characteristics. The Recommendations propose study formats appropriate to each stage. Taken together these provide an integrated stepwise pathway for evaluation of surgical techniques. The early stages of IDEAL (Idea/Development) focus on prospective design, transparency in reporting with provision of adequate technical detail, and description of the learning and change that occurs as a technique is refined. The Exploration stage is a collaborative effort to eliminate the barriers to a randomized control trial, including analysis of learning curves and definition of indications, variations, quality of surgical delivery and outcomes. This acts as a precursor to the Assessment stage, which is, where possible, a randomized control trial. The Long-term follow-up stage emphasizes registry-based studies of late and rare outcomes. [1]

Approach: We will describe and explain the rationale for the IDEAL Framework, and present a recent update of the IDEAL Recommendations, showing how they can be integrated into research and practice. We will report on current projects that demonstrate the uptake of IDEAL in the literature since its inception and show its usage in developing high-quality surgical innovation studies. We will illustrate how use of the IDEAL Recommendations can improve the quality of evidence for surgical techniques in early studies and facilitate progress to successful RCTs.

References:

1. McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC et al., No surgical innovation without evaluation: the IDEAL recommendations. Lancet. 2009 Sep 26; 374(9695):1105-12

Contributors

Allison Hirst
IDEAL Collaboration, Nuffield Department of Surgical Sciences, University of Oxford

Peter McCulloch

IDEAL Collaboration, Nuffield Department of Surgical Sciences, University of Oxford

INNOVATIVE USE OF STACKED AREA PLOTS IN DMC REPORTING

ALLISON FUREY

UNIVERSITY OF WISCONSIN-MADISON

For many clinical trials, Data Monitoring Committees (DMCs) are charged with monitoring the safety of trial participants. As trials evolve, study arms may exhibit differences over time in subject disposition, dose level, adverse events (AEs), laboratory abnormalities, or other measures of interest.

The Statistical Data Analysis Center (SDAC) at the University of Wisconsin–Madison specializes in producing interim reports and analyses for DMCs. Our reports are graphically based, allowing DMC members to easily identify differences between treatment groups and/or changes over time, and to review a large amount of information in a short amount of time.

Among the graphical tools available, the stacked area plot can be particularly illuminating. A stacked bar plot may be used to display the distribution of ordered categorical data at a given point in time; when the number of time points is large, the individual bars become small, and the effect is a display of "stacked areas." Stacked area plots illustrate changes in patient state over time, where each patient's state can be known (at least in principle) throughout the time period of interest. Examples of "states" may include patient status, dose level, severity of an AE of interest, or level of a lab analyte. We illustrate the utility of stacked area plots using examples such as the following.

•   In a single figure, a stacked area plot can elucidate patterns in the timing of treatment withdrawal, study discontinuation and death. Displays using calendar time produce a cumulative accrual plot stacked by subject disposition, whereas displays using time from treatment start produce cumulative follow-up plots.

•   Stacked area plots lend themselves to trial designs incorporating titration phases and/or multiple dosing options throughout one or more phases of a trial. Visual displays of the number or proportion of patients at each dose level over time can reveal patterns in tolerability and adherence as a trial unfolds.

•   In study designs including an active treatment run-in, randomized withdrawal phase, or post-treatment follow-up (in addition to a double-blind phase), examining specific AEs over time can uncover carry-over effects of treatment. Stacked area plots can be used to examine AEs that continue beyond a phase change, elucidating the approximate duration and severity of the continued effect.

•   In trials of agents designed to affect lab parameters (e.g., hemoglobin in a trial of anemia, INR in a trial of atrial fibrillation), stacked area plots can display the proportion of subjects in the target range over time, with narrow and wide targets illustrated in a single plot. Summary statistics may include percent time in the target range.

While stacked area plots are primarily descriptive in nature, they can reveal interesting patterns in data over time that would not be captured using more focused analytic approaches. DMC members we have worked with have found them extremely useful in their deliberations.

<u>Contributors</u>

Robin Bechhofer
University of Wisconsin-Madison

Melissa Schultz
University of Wisconsin-Madison

INTERACTIVE REPORTING AND VISUALIZATION OF STANDARDIZED CLINICAL DATA

MICHAEL FRASKETI

THE EMMES CORPORATION

Introduction: We have built an interactive reporting application that allows the user the ability to access data from multiple protocols and perform exploratory analysis. The Emmes Clinical Trials Reporting Application is implemented with R and Shiny and accesses a data warehouse built on a standardized data model.

Background: Traditional reporting of clinical research data relies on static tables, figures, and listings in HTML, or Word/PDF documents. The user has no opportunity to interact with the data or create customized data views. Interactive reports would expand user options when investigating data questions, and customize data listings, graphs, and reports.

Methods: The Emmes Clinical Trials Reporting Application is built upon the open-source software platform of R using the Shiny package. The data warehouse stores data using the Clinical Data Interchange Standards Consortium (CDISC) Standardized Data Tabulation Model (SDTM). A standardized data model makes the application database driven.

The application can generate customized tables, figures, and listings and export user-defined data sets for further analysis. The user can design queries to answer research questions, with the ability to drill down to subject-level data. The application is tablet and mobile friendly, making it ideal for the investigator who wants to look into data on the go.

Domain-specific reports can be produced to describe common data points of clinical studies including demographics, adverse events, and research laboratory data, and templates can be created to rapidly clone related reports. The flexible layout and design can be tailored to user requirements. Specialized applications can be created to focus on a single domain of data, or comprehensive reports can provide access to all protocol data.

Conclusions: An interactive reporting application based on R and Shiny can complement static reports for reporting clinical research data. A data warehouse with a standardized data model such as SDTM helps make the application data driven so it can be easily implemented and replicated.

Contributors

Teresa Wilson
The Emmes Corporation

Jeanine May
The Emmes Corporation

Abbie Bellamy
The Emmes Corporation

Olga Vasilenko
NIH/NIAID/VRC

Dan Sinnett
The Emmes Corporation

Katie Menard

The Emmes Corporation

Amy Bray

The Emmes Corporation

Noble Shore

The Emmes Corporation

INTERIM EVALUATION OF EFFICACY OR FUTILITY IN GROUP-SEQUENTIAL TRIALS WITH TWO TIME-TO-EVENT OUTCOMES

TOSHIMITSU HAMASAKI

NATIONAL CEREBRAL AND CARDIOVASCULAR CENTER

We discuss logrank test-based methods for efficacy or futility evaluation in group-sequential clinical trials that compare two interventions with respect to two time-to-event outcomes. Evaluation is conducted under three situations: (a) both events are non-composite and non-fatal, (b) both events are non-composite, but one event is fatal, and (c) one event is composite, but other is fatal and non-composite. Based on group-sequential boundaries, we consider several decision-making frameworks for evaluating efficacy or futility. We consider two inferential goals, evaluating if a test intervention is superior to a control intervention on: (i) both outcomes (Co-primary endpoints: CPE), and (ii) at least one outcome (multiple primary endpoints: MPE). For the CPE goal, we incorporate the correlations among the outcomes into the calculations for non-binding futility boundaries and sample sizes (or event numbers) as a function of other design parameters, including mean differences, the number of analyses, and efficacy boundaries. We investigate the operating characteristics of the decision-making frameworks in terms of efficacy/futility boundaries, power, the Type I error rate, sample sizes, event numbers, while varying the number of analyses, the correlations among the outcomes, and hazard ratios. We provide examples to illustrate the methods and discuss practical considerations when designing group-sequential designs in clinical trials with two time-to-event outcomes.

Contributors

Koko Asakura
National Cerebral and Cardiovascular Center

Scott Evans
Harvard T.H.Chan School of Public Health

Tomoyuki Sugimoto
Kagoshima University

Chin-Fu Hsiao
National Health Research Institutes

Haruko Yamamoto
National Cerebral and Cardiovascular Center

INTERIM MONITORING USING ADAPTIVE, TIME-VARYING NON-INFERIORITY MARGINS

BRETT HANSCOM

FRED HUTCH CANCER RESEARCH CENTER

Background: In prior work we have shown that when there is evidence of a strong association between study-population characteristics and the effectiveness of the active-control therapy in a non-inferiority trial, those characteristics may be used in a pre-specified way to adapt the non-inferiority (NI) margin. Using fixed margins under non-constancy leads to poor trial operating characteristics, while adapting the NI margin can provide substantial correction of Type-I error and power in the primary analysis at the end of the trial. Often, however, interim monitoring is used mid-trial to evaluate early signs of significant effects. Group-sequential stopping bounds depend directly on the NI margin, and hence it is critical that early stopping rules consider early evidence of non-constancy. In addition, the effectiveness of an active-control therapy may not be stable over time, in which case the appropriate NI margin for interim monitoring may depend on elapsed study time at each interim analysis.

Methods: We use meta-analysis regression to model the association between population characteristics and the effectiveness of a standard therapy. We then combine pre-trial meta-regression results with study population characteristics observed during two interim analyses to adjust the NI margin and update the corresponding O'Brien-Fleming stopping bounds. We consider two scenarios: one where active-control effectiveness is stable over time, and another where active-control effectiveness wanes over time. In each scenario, we evaluate trial operating characteristics via simulation.

Results: We consider a hypothetical NI trial of an experimental HIV Pre-exposure Prophylaxis (PrEP) drug versus a standard PrEP drug (active control) designed to provide one-sided alpha equal to 2.5% and 90% power with a target maximum of 100 HIV events. The assumed level of adherence to the standard therapy is 75%, which leads to a meta-analysis estimated relative risk (active control versus placebo) of 2.1 and an NI margin of 1.45 to assure non-inferiority. If the observed level of adherence is on target throughout the trial, neither the NI margin nor the group-sequential stopping bounds need to change. However, when adherence to the standard therapy is lower than planned and drops over time, the planned boundaries fail to control type-I error at the pre-specified level. If the true average adherence starts at 65%, drops to 55% at the second interim analysis, and further drops to 45% at the end of the trial, the probability of a false-positive result jumps to 17% and power inflates to 99%. By adjusting the stopping rules at each time point (See Figure 1), type-I error and power are maintained at 2.5% and 90% respectively.

Conclusion: If prior placebo-controlled trials provide evidence of an association between population characteristics and the effectiveness of an active-control therapy, group sequential stopping bounds may be adjusted based on observed population features, effectively maintaining pre-specified levels of Type-I error and power.

Figure 1. Panel (a) shows the planned group sequential stopping boundaries for 75% adherence, which remains constant throughout to trial. Panel (b) shows time-varying stopping bounds for declining adherence, where adherence ranges from 65% to 45% over the course of the trial.

Contributors

James Hughes
University of Washington

Brian Williamson
University of Washington

Deborah Donnell
Fred Hutch Cancer Research Center

MEASURING THE IMPACT OF PATIENT PREFERENCE IN CLINICAL TRIALS

DENISE ESSERMAN

YALE UNIVERSITY

Patient preference can have a substantial impact on a study's outcome, particularly for cases where it is not feasible to blind the participant to treatment assignment. Further, patient preference can have an impact on clinical practice, and is at the center of patient-centered outcomes research. Ignoring its impact on clinical outcomes and its role in clinical decision making would be unrealistic. We will present work we have done extending the two-stage clinical trial design first proposed by Rucker, including extensions to binary and count outcomes, as well as the inclusion of stratification. The two-stage design allows for the estimation of patient preference effects, in addition to the standard treatment effects we can estimate from the traditional randomized clinical trial design. We will discuss how to implement this design, potential alternatives to this design, and ways that we can expand methodology to better measure the influence of patient preference on clinical outcomes.

Contributors

Briana Cameron
23andMe

Yu Shi
Yale University

Michael Kane
Yale University

Peter Peduzzi
Yale University

MEDICAL IMAGE INTEGRATION WITH CLINICAL RESEARCH SYSTEMS

MICHAEL FRASKETI

THE EMMES CORPORATION

Background: The scope of clinical research management and analysis typically includes standardized demographic, visit-based, and event related data. Increasingly medical imaging is being used to supplement clinical data in establishing primary endpoints for clinical trials. The result is a large and complex dataset with unique management needs requiring specialized knowledge. Logistical considerations must be made regarding image collection, de-identification, storage, and retrieval that have significant time and cost implications. Beyond the task of building an electronic image archive and linking with study data, project requirements may extend to other resource-intensive tasks such as analyses of the image data and access control to qualified researchers. Many of these challenges have now been overcome at Emmes using a variety of tools to integrate medical images into our clinical research platform to support a variety of clinical and analytical aims.

Methods: Open-source imaging tools coupled with web-based systems and analytical frameworks have been leveraged to enhance our electronic clinical research system. Additionally, a publicly accessible website provides researchers the ability to query and request image archives. Custom reports are utilized to provide transparency to clients regarding image resources available for analysis. Study-specific workflows have been deployed to enable efficient and economic solutions to support complex research objectives. Collaboration between the project management and software development teams contributes to the success of these solutions and ensures client needs are met by the image processing technologies noted above.

Results: Emmes has successfully implemented clinical trials and observational studies with integrated medical image repositories, including CT, MRI, and ultrasound images. These studies represent a diverse range of therapeutic areas with unique data management needs.

Conclusions: Integrating medical image and clinical research data to support a diverse scope of research objectives is a challenging but not impossible task. Emmes' efforts have led to extensive research utilizing these images in developing national growth standards, disease identification, disease progression and outcomes, and machine and deep learning.

Contributors

Seth Sherman
The Emmes Corporation

Lucas Barwick
The Emmes Corporation

Rob Gore-Langton
The Emmes Corporation

Anne Lindblad
The Emmes Corporation

Traci Clemons
The Emmes Corporation

Diane Brandt

The Emmes Corporation

MONITORING AND REPORTING BLOOD PRESSURE DATA QUALITY AND DATA COMPLETENESS AND BLOOD PRESSURE TREATMENT ARM SEPARATION FOR THE BLOOD PRESSURE IN DIALYSIS (BID) PILOT STUDY

JENNIFER GASSMAN

CLEVELAND CLINIC

The BID Pilot Study is a recently-completed NIDDK/Dialysis Clinic Inc. funded multi-center clinical trial (in press, JASN). 126 hypertensive patients were randomized to a predialysis Blood Pressure (BP) goal of either 155 to 165 mmHg (Usual BP control, n=64) or 110 to 140 mmHg (Intensive BP control, n=62) for 12 months. The BID protocol specified

- Standardized Dialysis Unit (SDU) BP measured at each dialysis session, i.e., 3 times/week

- Inter-dialytic (44-hour) Ambulatory Blood Pressure Monitoring (ABPM) quarterly

- Home BP measurement (three consecutive readings, morning and evening) weekly

The DCC programmed patient-specific "Ready to Randomize" reports showing eligibility criteria including whether a Baseline patient had 1) ABPM with a sufficient number of day and night readings and an adequate Home BP submission, 2) a sufficient number of SDU BP measures, and 3) a running mean SDU BP supporting hypertension/not requiring anti-hypertensive back-titration. Study Coordinators ran these reports from an on-line menu.

Over the course of the four-year study, the DCC developed increasingly detailed data completeness, data quality, and BP goal achievement reports which were emailed study-wide each week. DCC reports fed back (site-by-site and overall) recent and cumulative data monitoring

- site documentation of each scheduled dialysis as SDU BP measured, dialysis with no SDU measured, or (rarely) dialysis missed

- participant current achievement of Usual or Intensive BP goal as determined by 2-week running mean SDU BP

- participant completion of home BP weekly (per protocol) or at least once per month

- participant completion of ABPM quarterly

Tables of completeness and quality of BP data were discussed on weekly Study Coordinator calls, bi-monthly Steering Committee calls, and participating site clinical team meetings. Knowledge of missing or inadequate data allowed site teams to encourage patient Home BP and ABPM collection and prioritize SDU BP measurement and site data entry. Corrections and improvements were updated in the report as they happened, allowing study coordinators to watch their sites' progress with respect to missing data, inadequate data, or participants with out-of-range blood pressure.

Detailed reporting on patient achievement of Usual or Intensive 2-week running mean SDU BP, anti-hypertensives, treatment time, and dry weight enabled a team approach to achievement of

goal, with physicians brainstorming suggestions for bringing patients to goal on each Steering Committee call.

At study end, the majority of patients had measured SDU BP at more than 8 times a month, measured home BP more than once a month, and completed an adequate fourth quarterly ABPM. A 10 mmHg separation was achieved and sustained in running mean SDU BP and in home morning BP.

"What predialysis systolic BP goal is best?" is still an open question. BID Pilot Study completeness and quality of data, safety, and achieved separation between BP treatment arms support the feasibility of a full scale trial.

BID Participating Sites (PIs) included Medical University of South Carolina (David Ploth); University of Pittsburgh Medical Center (Manisha Jhamb); CWRU/University Hospitals Medical Center (Lavinia Negrea); DCI/University of New Mexico (Philip Zager); Tufts University Medical Center (Dana Miskulin).

Contributors

Dana Miskulin
Tufts Medical Center

Cynthia Kendrick
Cleveland Clinic

Karen Brittain
Cleveland Clinic

Susan Sherer
Cleveland Clinic

Kimberly Wiggins
Cleveland Clinic

Jack Mackrell

Cleveland Clinic

Brett Larive

Cleveland Clinic

Serena Cumber

DCI/University of New Mexico

Susan Paine

DCI/University of New Mexico

Philip Zager

DCI/University of New Mexico

MONITORING RARE EVENTS DURING AN ONGOING CLINICAL TRIAL

HALEY HEDLIN

QUANTITATIVE SCIENCES UNIT, STANFORD SCHOOL OF MEDICINE

We conduct a simulation study to evaluate approaches to monitoring rare events during an ongoing clinical trial. We consider the setting where a data and safety monitoring board (DSMB) or another entity is monitoring a rare event during the course of a two-arm study and desires guidelines around whether the event is occurring more in one arm than the other.

This work builds on a history of interim testing and event monitoring in clinical trials and considers the performance of previously developed methods when applied to a rare event. We do not consider stopping rules, but instead present guidelines to guide the monitoring in combination with other accumulating information during the trial. Our work is motivated by a real trial where a rare safety event was of interest. Most existing monitoring methods assume sufficiently large sample sizes or number of observed events. We aim to provide useful guidelines to be used in monitoring rare events while information is accumulating during a trial.

We evaluate the operating characteristics of two complementary approaches, one based on the sequential probability ratio test and the other on exact confidence intervals. The operating characteristics are evaluated via simulation in trials of sample sizes ranging from 1,000 to 10,000 and over a range of true relative risks. Importantly, we focus on settings where the initial sample sizes are relatively small and the proportion experiencing events are very small. Characteristics considered include proportion of trials where a threshold set by the Sequential Probability Ratio Test (SPRT) were crossed and the median and interquartile range of the time when the threshold was crossed. These thresholds are considered in combination with exact confidence intervals.

We present guidelines with desirable operating characteristics to guide rare event monitoring during an ongoing clinical trial.

Contributors

Victoria Ding
Quantitative Sciences Unit, Stanford School of Medicine

OPTIMAL COMPOSITE ENDPOINTS FOR CLINICAL TRIALS OF CHRONIC PROGRESSIVE DISEASE

STEVEN EDLAND

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Background. Composite scales have been proposed as outcome measures for clinical trials of chronic progressive disease. For example, the Prodromal Alzheimer's Cognitive Composite (PACC) is the sum of z-score normed component measures assessing episodic memory, timed executive function, and global cognition. We have recently described an alternative method of calculating composite total scores using the weighted sum of the component measures that maximizes signal-to-noise of the resulting composite score and optimizes performance as an endpoint for clinical trials [Ard, et al. Pharmaceutical Statistics. doi: 10.1002/pst.1701].

Methods. We use data from completed cohort studies to demonstrate the performance of optimal composite endpoints vis-a-vis established outcome measures. As a pragmatic metric of endpoint performance, we use standard power calculation formulas to calculate the relative sample size required for trials powered to detect a 25% slowing of decline on the respective outcome measures. Two examples are considered. The performance of the optimized PACC vis-a-vis the z-score normed PACC is demonstrated using data from cognitively normal subjects enrolled in the Alzheimer's Disease Cooperative Study (ADCS) Prevention Instrument cohort study. A second example uses volumetric MRI measures of brain atrophy in primary progressive aphasia to demonstrate the performance of a composite constructed from specific regions of interest relevant to this disease.

Results. In the context of an Alzheimer's disease primary prevention trial informed by the ADCS Prevention Instrument cohort, a three year clinical trial using the optimal PACC would require less than a quarter of the subjects required by the z-score normed PACC (600 subjects per arm versus 2500 subjects per arm assuming annual observations, equal allocation to arms, type I error rate 0.05, and 80% power; Edland et al. Alzheimer's & Dementia, doi: 10.1016/j.trci.2016.12.004). In the context of a treatment trial targeting brain volumetric decline in primary progression aphasia, an optimal composite composed of subregions of the left perisylvian temporal cortex would require 40% less subjects than a trial using total perisylvian temporal cortex volume (15 subjects per arm versus 25 subjects per arm assuming a two year trial comparing change first to last, equal allocation to arms, type I error rate 0.05, and 90% power; Edland et al. Alzheimer's & Dementia, doi: 10.1016/j.trci.2016.05.002).

Conclusions. The potential improvement in trial power and efficiency by optimal composite endpoints is substantial. This is especially relevant to primary prevention trials which may otherwise be cost-prohibitive, and to trials of rare conditions where the available subject pool may be otherwise be limiting. In general, more efficient outcome measures mean that greater statistical power can be achieved given fixed sample size.

<u>Contributors</u>

PRESENTING USER-TESTED EVIDENCE BASED TOOLS TO SUPPORT DEVELOPMENT OF TRIAL RECRUITMENT STRATEGIES

HEIDI GARDNER

UNIVERSITY OF ABERDEEN

Objectives

Evidence-based decision-making relies on access to reliable research evidence. The Cochrane recruitment review is the most relevant piece of research that trialists can use when designing recruitment strategies, but information is buried within the lengthy document. Stakeholders are often unaware of, or do not have time to read the entire manuscript; simply publishing a review is not enough to change research design or behaviour. We aimed to develop a way to effectively communicate summarised recruitment research evidence direct to trial stakeholders (e.g. trial managers, methodologists, research nurses) that was both understandable, and feasible to implement at scale and through a variety of communication channels.

Study Design

We conducted semi-structured scoping interviews with 23 researchers involved in aspects of trial recruitment, on the desired structure, content and delivery of research evidence about trial recruitment. Feedback suggested a web-based resource with uncluttered design.

We then worked with professional graphic designers on the first iteration of a prototype web-based resource, and then conducted iterative user-tests with 16 participants using a think-aloud protocol method. These data allowed us to sort problem into three categories: showstopper - leads to critical errors in understanding; medium - creates frustration; low - minor/cosmetic problems). We then integrated improvements to further iterations of the resource. In total, we conducted 4 rounds of user-testing before producing our final resource.

Setting

Participants from UK National Health Service sites that are involved in trials, academic institutions (UK, the Netherlands, Canada) and Clinical Research Organisations partnering with pharmaceutical companies (UK, South Africa, Italy).

Results

Stakeholders involved in both scoping and user-testing were initially surprised by the resource's simplicity after being exposed to research evidence as traditionally presented in journal articles. Interestingly, user-testing further simplified the resource, as stakeholders pushed their need for uncluttered presentation. User-testing revealed unexpected comprehension problems due to seemingly small design decisions; use of colour and size of particular pieces of text. Patterns and shapes were also recognised that were not intentionally integrated within the resource.

In order for stakeholders to assess the relevance of the research to their own specific trial environments, it became clear that context (e.g. population, location, setting) of the trials presented

was something that needed to be conveyed more clearly. Participants reported that they required an easy-to-understand resource that did not require them to spend significant amounts of time processing (unlike, e.g. a conventional systematic review), so language was edited to reflect a more conversational style.

Conclusion

User-testing led to a stripped-down resource, presenting minimal information in order to produce maximal impact for stakeholders looking for evidence to inform the way they design and conduct the process of recruitment. After user-testing, the resource was available online with an open 'comments' section to facilitate ongoing improvements. We will demonstrate the final tool at the conference, and designs from round 1 and 2 of user-testing can be accessed here: http://bit.ly/2hLtnp1.

This work is part of the Trial Forge initiative to improve trial efficiency.

<u>Contributors</u>

Katie Gillies
University of Aberdeen

Shaun Treweek
University of Aberdeen

SAFETY AND TOLERABILITY OF TRANSCRANIAL DIRECT CURRENT STIMULATION TO STROKE PATIENTS - A PHASE I CURRENT ESCALATION STUDY

WAYNE FENG

MEDICAL UNIVERSITY OF SOUTH CAROLINA

Background and objective: A prior meta-analysis revealed that higher doses of transcranial direct current

stimulation (tDCS) have a better post-stroke upper-extremity motor recovery. While this finding suggests

that currents greater than the typically used 2 mA may be more efficacious, the safety and tolerability of

higher currents have not been assessed in stroke patients. We aim to assess the safety and tolerability of

single session of up to 4 mA in stroke patients.

Methods: We adapted a traditional 3 þ 3 study design with a current escalation schedule of 1»2»2.5»3»3.5»4 mA for this tDCS safety study. We administered one 30-min session of bihemispheric montage tDCS and simultaneous customary occupational therapy to patients with first-ever ischemic stroke. We assessed safety with pre-defined stopping rules and investigated tolerability through a questionnaire. Additionally, we monitored body resistance and skin temperature in real-time at the electrode contact site.

Results: Eighteen patients completed the study. The current was escalated to 4 mA without meeting the

pre-defined stopping rules or causing any major safety concern. 50% of patients experienced transient

skin redness without injury. No rise in temperature (range 26 C to 35 C) was noted and skin barrier

function remained intact (i.e. body resistance >1 kU).

Conclusion: Our phase I safety study supports that single session of bihemispheric tDCS with current up

to 4 mA is safe and tolerable in stroke patients. A phase II study to further test the safety and preliminary

efficacy with multi-session tDCS at 4 mA (as compared with lower current and sham stimulation) is a

logical next step. (ClinicalTrials.gov Identifier: NCT02763826.)

Contributors

Rong Chen
Medical University of South Carolina

Rachael Deardorff
Medical University of South Carolina

Blair Dellenbach
Medical University of South Carolina

Steven Kautz
Medical University of South Carolina

Mark George
Medical University of South Carolina

Pratik Chhatbar

Medical University of South Carolina

SELECTING TECHNOLOGY DERIVED NOVEL ENDPOINTS FOR DEVELOPMENT AND INCLUSION IN CLINICAL TRIALS

JENNIFER GOLDSACK

CLINICAL TRIALS TRANSFORMATION INITIATIVE

Background: Currently, when data are gathered from participants in clinical trials, these assessments typically take place in healthcare settings rather than in the context of patients' daily lives. In addition, many assessments of physical function rely on relatively subjective outcome measures that are reported by participants or their healthcare providers. In addition to being limited in terms of their objectivity, these measures also may capture only brief "snapshots" of the participant's functionality and/or disease burden at a given point in time.

Mobile technology offers new ways to capture objective measurements as clinical trial participants go about their daily lives by utilizing novel endpoints, defined as 1) new endpoints that have not previously been possible to assess, or 2) existing endpoints that can be measured in new and possibly better ways. These novel endpoints have the potential to provide high-quality data pertaining to outcomes that are meaningful to patients while theoretically enabling larger trials with reduced barriers to participation, thus making possible more sensitive, generalizable, and patient-centric assessments.

Methods: The Clinical Trial Transformation Initiative (CTTI) convened a multi-stakeholder team comprised of sponsors, academics, regulators, and measurement science, technical and patient experts to define and test a framework to optimize the selection of technology derived novel endpoints for development and inclusion in clinical trials.

Results: When selecting outcome assessments for development, the approach should be patient-centered with the patient voice included as standard in the work of clinician experts in the therapeutic area. Selection should address an unmet need for assessments that directly measure or indirectly reflect an aspect of the disease or illness that, if relieved, improved, or prevented would be meaningful to patients.

Developing novel endpoints is a time-consuming and resource-intensive process. For this reason, early successes should be made public as use cases to inform future efforts. As such, sponsors, consortia and grant-making organizations should take a systematic approach to identifying key novel endpoints to be developed for use in clinical trials. CTTI has developed an Interactive Selection Tool consisting of fourteen weighted domains to support decisions between viable technology-derived novel endpoints for development. This tool could also be used by other groups, such as mobile technology companies, who wish to assess the potential role of a sensor or device under development.

Conclusion: There are significant benefits associated with the selection, development and inclusion of technology-derived novel endpoints in clinical studies. In order to maximize these benefits, selection of appropriate outcome assessments is critical. CTTI has developed recommendations and resources to support this selection.

Contributors

Lauren Bataille
The Michael J. Fox Foundation

Rob DiCicco
GlaxoSmithKline

Cheryl Grandinetti
US Food and Drug Administration

Will Herrington
University of Oxford

Martin Landray
University of Oxford

Kaveeta Vasisht

US Food and Drug Administration

Ashish Narayan

Mount Sinai Medical School

Elektra Papadopoulous

US Food and Drug Administration

Nirav Sheth

Micromedicine, Inc.

Ken Skodacek

US Food and Drug Administration

Komathi Stem

monARC Bionetworks, Inc.

SHOULD WE PLAN FOR THE UNPLANNED? MANAGING UNDESIRABLE BUT UNAVOIDABLE PROTOCOL AMENDMENTS IN CLINICAL TRIALS.

SARAH FINLAYSON

UNIVERSITY OF SYDNEY

Background: Protocol amendments in clinical research are common, and are associated with significant time and economic costs. Approximately 55% of amendments are unavoidable, for reasons including new safety data and changes in standard of care1. While some data exist on approval timelines for pharma-sponsored studies, limited data exist for investigator-initiated studies. We present the approval timelines for the most recent of 3 protocol amendments in the investigator-initiated Australian Placental Transfusion Study.

Methods: The amendment was finalized on July 1st, 2016. Dates of release to site and lead ethics submission were collected from Sponsor records, while approval date was taken from the final approval letter. All other dates were by site report. All 25 APTS sites were included. Sites under a lead ethics committee were grouped, unless sites required subsequent institutional approval – in which case, they were grouped only when measuring ethics times.

Results: Median time to approval was 89 days, however the range was very large (13-256d). Approval times for Australian sites were considerably shorter than for international sites (45d vs. 119d). Site time accounted for the greatest proportion of approval time, followed by ethics, other institutional approval, and Sponsor time. There was a strong relationship between approval time and the level of central resourcing required, as measured by site prompts ($R2=0.93$). Results were also collected for internal approval timelines, different approval models and methods of site contact.

Discussion: Approval timelines were lengthy, and delays were associated with increased central resourcing. Co-location of the coordinating center may explain the shorter approval times for Australian sites. Given the impact of site time, early additional coordinating resources (possibly targeted at international sites lacking a local coordinating centre) may provide an opportunity to reduce approval timelines.

Protocol amendments delay studies, and result in a loss of enthusiasm at site. To minimize amendments and their impacts, investigators should ensure appropriate planning and feasibility assessments are completed prior to study commencement.

References:

1.   Getz K. Acknowledging cycle time impact from protocol amendments. Applied Clinical Trials 2016; 25(4): 16-17

Contributors

Alpana Ghadge
University of Sydney

Wendy Hague
University of Sydney

STRATEGIES TO IMPROVE PARTICIPANT RETENTION TO RANDOMIZED CONTROLLED TRIALS: WHAT DO NON-RANDOMIZED EVALUATIONS OF RETENTION STRATEGIES HAVE TO TELL US?

ADEL GOMAA IBRAHIM EL FEKY

UNIVERSITY OF ABERDEEN

Background

Despite significant investment in trial infrastructure, a remarkably large number of trials continue to face challenges with retention. UK studies suggest that over half of trials lose more than 10% of participants.There is systematic review evidence summarizing randomized evaluations of retention interventions but the modest quantity and quality of primary research limits its use for decision-making. This review collates evidence from non-randomized evaluations to judge whether it could provide complementary information to guide trialists developing retention strategies.

Methods

Non-randomized studies (observational studies) including a comparison of two or more strategies to increase participant retention in randomized trials, or comparing one or more strategies with no strategy were eligible. Our primary outcome was the proportion of participants providing primary outcome data. Risk of bias was assessed using the Cochrane ROBINS-I ("Risk Of Bias In Non-randomized Studies - of Interventions") tool. The Grading of Recommendations Assessment Development and Evaluation (GRADE) criteria will be used to assess the overall certainty in the evidence for all included studies. Where studies can be pooled, GRADE has provision for upgrading bodies of evidence from observational studies depending on a number of factors, including consistency.

Results

The search was conducted for the period 2007 through October 2017 in 5 electronic databases, including MEDLINE and EMBASE. The search retrieved 7609 records. Once de-duplicated 5371 records remained. Full-text reports for 92 studies were screened. Abstract screening and full text check were conducted by two independent reviewers with disagreements resolved by discussion. A total of 15 retention studies were included, with studies from across a wide range of disease areas, countries and healthcare settings.

Most of the included studies were at low risk of bias for all ROBINS-I risk of bias domains (meaning the study is comparable to a well-designed randomised study) except for confounding where most of them were at moderate risk of bias, meaning that these studies were robust for a non-randomised study with respect to this bias domain but cannot be compared to a well-conducted randomised study. Only 3 studies were found to be at serious risk of bias on the confounding domain.

The retention strategies evaluated fell into seven broad categories:

- Participant incentives (e.g. use of a monetary incentive) (2 studies),

- Communication strategies (e.g. telephone call or text messaging for follow-up) (7 studies),

- Different mode of questionnaire completion (e.g. online questionnaire for follow-up) (1 study),

- Different questionnaire format (e.g. short version of online questionnaire) (1 study),

- Home follow-up (1 study),

- Design strategies (e.g. use of a run-in period) (1 study),

- Multifaceted strategies (e.g. intensive tracing efforts to locate study participants) (2 studies).

This preliminary categorisation will be finalised in early 2018 prior to discussion of whether any studies can be pooled and whether GRADE assessments can be upgraded in light of pooling.

Conclusion

Initial indications are that non-randomized evaluations of trial retention interventions may offer useful evidence for trial teams planning their retention strategies although this will be confirmed or refuted in early 2018 through formal GRADE assessment. If confirmed, well-done, non-randomized evaluations may provide sufficient certainty for trial decision-making for at least some classes of retention interventions.

<u>Contributors</u>

Shaun Treweek
Health Services Research Unit, University of Aberdeen

Katie Gillies
Health Services Research Unit, University of Aberdeen

Heidi Gardner
Health Services Research Unit, University of Aberdeen

Cynthia Fraser
Health Services Research Unit, University of Aberdeen

THE DATA EXTRACT VERIFY FILE – AN OPEN SOURCE AND CROSS PLATFORM XML FILE TO VERIFY DATA EXTRACTS

JONATHAN GIBB

ROBERTSON CENTRE FOR BIOSTATISTICS

Objective

Data transfers are a common and vital part of modern clinical trials. There are many potential transfer types during the different stages of a clinical trial. For all such transfers it is vital that the validity of the exported data can be verified i.e. it is both complete and accurate.

Background

Hash functions that are available in most modern programming, database and statistical tools provide the fundamental mechanism required to verify exported data. However conventions are required to ensure that hashes are constructed in the same manner across different systems.

Additionally simply hashing each record does not provide full verification of the data. Each record imported, hashed and compared would be confirmed accurate but this does not confirm that all of the data has been imported and the user could be unaware that the dataset is incomplete and overall the import process was inaccurate.

Methods

A simple XML document format was devised that should accompany any data extract.

The document provides a very small amount of metadata about the extract, combined with a common set of conventions that were devised to allow any system importing the extract to recreate suitable hashes across different record types and different system types.

It is important to note that the XML document is independent from the underlying format of the data extract; therefore the XML document may describe a CSV, XML or proprietary binary formatted extract.

Results

Existing extracts within the CTU were enhanced with data extract verify files and verifications were performed in Microsoft SQL Server, Microsoft C# (.Net) and SAS.

Conclusions

The data extract verify file provides a means to ensure data transfers are complete and accurate between a wide number of technological platforms and systems and should help data extracts be readable into the future, regardless of their underlying format.

Contributors

Sharon Kean
Robertson Centre for Biostatistics

THE UGLY TRUTHS OF TODAY'S DMCS – AND WHAT THE SDAC CAN DO TO HELP

DAVID KERR

AXIO RESEARCH

Background:

Data Monitoring Committees (DMCs) are an important component of the clinical trial process. They are charged with reviewing interim data and making recommendations to protect the safety of trial participants and ensure the scientific integrity of the study. DMCs in practice are not performing as effectively as they should be. A real-world assessment will be provided of today's DMC environment as well as suggestions for improvements that can be led by the SDAC (Statistical Data Analysis Center).

Methods:

The author is a leader of a Contract Research Organization (CRO) that specializes in providing independent DMC services for government and industry sponsored trials. He has worked with over a hundred unique DMCs and facilitated approximately 500 DMC meetings. He will discuss the ugly truths about today's DMCs, but provide insight into how an SDAC can be an agent for improving the DMC process.

Results:

Topics that will be covered include:

- Ugly truths of today's DMCs – DMC, Sponsor, and SDAC all can be guilty of sub-optimal performance

- Key outputs that should be provided to DMCs

- Successful approaches for the SDAC to assist the DMC through the outputs

- Training the SDAC team

- Prepping the SDAC before the DMC meeting

- Intangibles for facilitating a quality DMC meeting

- Useful SOPs for an SDAC

- SDAC training potential or current DMC members

- Examples of tricky SDAC/DMC interactions

- Some final truths

Conclusions:

DMCs will perform a more thorough and efficient review with assistance from a skilled SDAC. An improved DMC process will better protect the safety of the trial participants and the validity of study results.

Contributors

THE WIN RATIO APPROACH IN KIDNEY TRANSPLANT TRIALS

NICHOLAS FERGUSSON

UNIVERSITY OF BRITISH COLUMBIA

Background:

Trials in kidney transplantation have often suffered from low short-term event rates and low efficiency. In order to increase trial event rate and statistical power, an endpoint of kidney function decline is often incorporated with death and transplant failure into a single, time-to-event composite outcome measure. Although these composites may increase the event rate and trial efficiency they contain components of unequal significance—a patient death is not equal to a decline in kidney function. A composite with endpoints of unequal significance and frequency can lead to misleading trial conclusions. The win ratio approach attempts to mitigate these concerns by ranking and sequentially assessing data.

Objective: Determine the impact and potential utility of using the win ratio approach within the setting of a kidney transplant trial.

Methods: We present an application of the win ratio approach to a trial in kidney transplantation. The original primary outcome of the trial was a composite of time to either death, kidney transplant failure, or doubling of serum creatinine. Additionally, we used a post-hoc alternative composite outcome substituting doubling of serum creatinine for a >40% eGFR decline. We compare the win ratio approach to a conventional time-to-event hazard ratio analysis. A win ratio with a lower 95% confidence limit greater than 1 indicates a positive treatment effect with statistical significance.

Results: On the original composite of death, transplant failure, and doubling of serum creatinine, ramipril treatment was associated with a win ratio of 1.02 (95% CI, 0.54-1.83) versus a hazard ratio of 0.96 (95% CI, 0.55-1.65). In the time-to-event analysis, the endpoints of death and transplant failure accounted for 38% and 23% of the event data respectively, while a doubling of serum creatinine accounted for 38%. In the win ratio analysis, the endpoints of death and transplant failure accounted for 46% and 38% of the event data respectively, while a doubling of serum creatinine accounted for 15%

On a composite of death, transplant failure, and a 40% eGFR decline, ramipril treatment was associated with a win ratio of 1.13 (95% CI, 0.59-1.57) versus a hazard ratio of 0.89 (95% CI, 0.54-1.48). In the time-to-event analysis the endpoints of death and transplant failure accounted for 30% and 16% of the event data respectively, while a >40% eGFR decline accounted for 54%. Using the win ratio analysis, the endpoints of death and transplant failure accounted for 39% and 32% respectively, while a 40% eGFR decline accounted for 30% of the event data used.

Conclusion: Within a randomized trial in kidney transplantation the win ratio approach maximized data from more clinically significant endpoints. Consequently, this method may limit concerns surrounding composite outcomes that possess endpoints of unequal clinical significance. Going forward, the win ratio approach may be a highly useful analysis method for kidney transplant trials– particularly if lesser kidney function decline endpoints are integrated into trial composites.

<u>Contributors</u>

Greg Knoll
University of Ottawa

Tim Ramsay
University of Ottawa

Michaël Chassé
Université de Montréal

Shane English
University of Ottawa

TRANSLATING ASSESSMENTS IN CLINICAL TRIALS CONDUCTED IN THE US

EVE JELSTROM

THE EMMES CORPORATION

Background:

Roughly 20% (63 million) of US residents do not speak English. These non-English speakers are frequently excluded from participation in research studies in situations where it is logistically difficult to accommodate them. Study eligibility often requires the ability to read and write in English, which rules out many participants during screening. However, in some cases the exclusion of these potential participants may bias the research findings and omit individuals from clinical trials from which they could potentially benefit.

Objective:

To include non-English speakers by translating key study documents. Translation should allow for higher inclusion of non-English speakers and potentially increase the representativeness of the study sample and the generalizability of the study results to the increasingly diverse US population.

Methods:

The Emmes Corporation, in its role as coordinating centers for studies conducted by The National Drug Abuse Treatment Clinical Trials Network (CTN) of the National Institute on Drug Abuse (NIDA), explored new implementation practices during the recent planning for a study that had the potential for a high number of Spanish-only speaking study participants. Incorporating this population required identifying the assessments and study materials that would be translated, contracting with a commercial translation company with certified translators to prepare the required materials, and customizing the EDC system with a Spanish user interface. Additionally, Spanish speaking research staff were made available at each site to communicate with non-English speaking participants, and documents were created that outline their involvement.

Results:

The translation of the informed consent form(s) (ICFs) and participant self-report assessments is an important step to meet IRB requirements and assure the study participant fully understands the study design and activities they agree to if they participate. In addition to participant self-reported assessments, case report forms (CRFs) where study staff read questions and instructions to the participant should be translated to ensure there is no variability in the way these items are read to participants within a site and across sites, and to reduce the impact of translation errors. This study used an electronic patient reported outcomes (ePRO) module, which also required translation of all system text, and navigation and pop-up messages.

Conclusions:

This presentation will discuss the key considerations and knowledge gained in the implementation of best practices regarding translations for clinical trials including forward and back translation,

considerations for different dialects found across the US, and how best to hire and train the research staff to conduct the clinical trial with non-English mono-lingual study participants.

Even though the effort required to implement a Spanish component may be costly and time consuming, the generalizability and increase in available participants to enroll may be beneficial to a clinical trial.

<u>Contributors</u>

Julia Collins
The Emmes Corporation

Ashley Case
The Emmes Corporation

Lauren Yesko
The Emmes Corporation

Dikla Shmueli-Blumberg
The Emmes Corporation

UTILIZATION OF A PAIRED COMPARISON ANALYSIS FRAMEWORK TO INFORM AND PRIORITIZE DECISION-MAKING IN A HIGHLY MATRIXED CLINICAL RESEARCH PROGRAM

MARCUS JOHNSON

DURHAM VA HEALTH CARE SYSTEM

Background

Project prioritization is often an arduous task for any organization. This is particularly evident in large, complex, organizations with matrixed management structures, such as the VA Cooperative Studies Program (CSP). CSP is responsible for the planning and conduct of large multicenter clinical trials and epidemiological studies in the Department of Veterans Affairs (VA). CSP Health Systems Specialists (HSSs) have the primary responsibilities of facilitating alignment and coordination of program-level activities, and leading projects and initiatives to meet program goals. Project Prioritization is best informed through a methodology that is structured, well-organized, and inclusive of multiple stakeholder perspectives to increase the likelihood of an initiative's completion, sustainability, and efficiency. There is an abundance of literature on Paired Comparison (PC) analyses to inform decision-making but there is limited publicly available information on its use in research administration settings.

Objective

To utilize a PC analyses framework to select and prioritize projects assigned to the CSP HSS group.

Methods

Participants were ten HSSs representing the 10 VA CSP Centers: 3 epidemiology coordinating (EC) and 7 clinical trial (CC) coordinating centers. To practice the PC method and create a shared framework under which all participants would subsequently assess projects, members completed a Values Comparison (VC) exercise using the PC worksheet downloaded from www.mindtools.com. Group members were shown a visual of the worksheet, which was populated with five values, and given verbal instructions for completion. Members were given three minutes to complete the worksheet individually. Members provided the results of their worksheet and scores were tabulated and shared by the facilitator. Members were subsequently provided PC worksheets, populated with the group's projects, and asked to complete the worksheet using the previous instructions keeping the shared values in mind. 90% of the worksheets were returned to the facilitator for scoring.

Results

The PC exercise resulted in a clear ranking of the group's shared values, with Safety rising to No. 1. The Project PC exercise generated two distinct findings. When results were stratified across the "EC HSSs" and "CC HSSs", the EC HSSs placed a higher value on projects that would provide training for

their role, while others placed higher value on projects that attempted to address program-level issues. When all participant scorings were tabulated together, three projects aimed at addressing program-level issues clearly rose to the forefront, with the highest-ranking project being one that had been given high-priority by VA leadership.

Conclusion

This effort successfully utilized the PC analysis framework to prioritize a list of projects that the HSS group would undertake. These PC exercises allowed the HSS group to identify shared values as they related to program-level projects and to use those values to assess the urgency and feasibility of group-assigned projects prior to investing time, effort, and money into them. There are numerous challenges to effectively performing decision-making in the context of prioritizing organizational projects. Therefore, the strategies that have been outlined here may be transferable to other settings.

Contributors

Melissa Middleton
New Mexico VA Health Care System

Tara Burke
Perry Point VA Medical Center

Tammy Barnett
Edward Hines Jr. VA Hospital

"ESSENTIAL DATA FOR DATA MONITORING COMMITTEE (DMC) REVIEW – THE IDEAL OPEN AND CLOSED SESSION REPORTS"

SHARAVI PEERAMSETTI

AXIO RESEARCH

Background:

The DMC is appointed to review quality of trial conduct and/or trial operations as well as to provide ethical oversight to reach a benefit-risk conclusion for a clinical trial. They should be able to review data regarding: enrollment, demographics and baseline characteristics, disposition, violations (study, and protocol), adherence, retention, safety (adverse events and labs) and efficacy (outcome measures).

Methods:

The sponsor will present the current study status which includes the enrollment, violations, adherence and retention to the DMC in the open session. The reporting statistician will present quality to trial conduct, safety and efficacy data by treatment arm to the DMC in the closed session.

Results:

An example of each of these reports will be presented to show proper ways to summarize these data for the DMC.

- Disposition – consort diagram

- Enrollment – planned vs. actual figure

- Demographics and baseline characteristics – stratification factors and baseline covariates

- Study violations (CCC criteria – complete, correct and current)

- coming in for scheduled visits, but not complete information is captured

- coming in for scheduled visits, but not correct information is captured

- not coming in at regular scheduled visits, overdue visits

- Protocol violations – summary of the current violations in the study till date

- Adherence – drug exposure and study follow-up of subjects who are on and off treatment

- Retention – retain subjects all the way through the follow-up visits

- Safety

- summary of Laboratory abnormalities (shift tables or display box plots over time)

-        adverse events which should include summary of events grade 3 and higher, events of special interest (AESI)

-        graphics are helpful to display – volcano plots, dot plots with or without relative risk (95% CI) – explain when they are necessary and when not

• Efficacy

-        summarize primary and secondary outcome measures

-        swimmer plots that display response assessments

-        adjudicated event summaries (ex: cardiac events or tumor responses)

-        Kaplan-Meier curves

Conclusion:

These reports enable DMC to propose changes or potential improvements to the study design and conduct and, most importantly, to identify any potential safety concerns for the trial participants.

<u>Contributors</u>

"FORGET-ME-NOT" – A PROTOCOL FOR A PILOT CLUSTER RANDOMISED TRIAL TO PILOT A CORE INFORMATION SET INTERVENTION FOR INFORMED CONSENT TO INVASIVE PROCEDURES INCLUDING SURGERY

BARRY MAIN

UNIVERSITY OF BRISTOL

Background

All surgeons are familiar with consent. It is the ethical, professional and legal obligation to gain a patient's authorisation to operate. Recent legislation in the United Kingdom and international initiatives promoting shared decision-making emphasise the need for patient-centred, information provision that results in the patient choosing the treatment (or no treatment) that best meets their needs, beliefs, and circumstance. But this is difficult. For surgeons, these conversations require time, training, and skill. For patients, who often have little baseline knowledge of what is being proposed, formulating ideas about materiality and weighing-up treatment options is challenging. There is a need for better means of stimulating conversations that are meaningful and useful to individual patients about to undergo invasive procedures.

Core information sets (CISs) are a potential means of facilitating patient-centred information disclosure and discussion.[1,2,3] These are scientifically developed to contain information rated as most important for discussion by patients and surgeons, and are intended as a standardised baseline from which further discussion of importance to the individual patient can take place. While the methods for developing CISs have been peer reviewed and published, the model has not yet been tested for effectiveness in practice.

Aims

The overall aim of the "Forget-me-not" study is to develop a communication intervention to optimise informed consent discussions for patients about to undergo invasive procedures, and to investigate the feasibility of evaluating the intervention in a cluster randomised trial of cancer multidisciplinary teams (MDTs). The specific objectives are:

i)   To develop a core information intervention for informed consent for surgery to inform the design of the intervention under evaluation in a main trial.

ii)   To establish the proportion of potentially eligible MDTs who can be approached about the trial, who can be confirmed as eligible, who can be successfully recruited and randomised, and who are willing to undergo research assessments

iii)   To establish outcome measures that are recognised as comprehensive, valid, and reliable assessments of informed consent

Methods

There will be three work packages

i)   Development of consultation interventions incorporating the core information set(s)

A novel overview of systematic reviews of patient-centred communication interventions will inform the design of up to three potential interventions incorporating the CIS. These will be refined by discussion in focus groups involving patients and healthcare professionals. The optimal intervention design will be taken forward.

ii)   Feasibility cluster randomised controlled trial to establish if a main trial is possible

Six MDTs will be randomised to the CIS intervention or standard care. The primary outcome will be whether a full cluster randomised trial is feasible. The acceptability of the intervention and the trial methods will be considered. A process evaluation will assess the fidelity of intervention delivery.

iii)   Study findings and dissemination

Key stakeholders will be invited to a workshop at which the results of the feasibility study will be discussed. If a full trial is deemed feasible, the facilitated workshop will seek stakeholder input on the design and outcome measures for that trial.

Contributors

Suzanne Audrey
University of Bristol

Leila Rooshenas
University of Bristol

Chris Metcalfe
University of Bristol

Richard Huxtable
University of Bristol

Steven Thomas
University of Bristol

Jane Blazeby

University of Bristol

A UNIQUE COLLABORATION FOR A DATA COORDINATING CENTER

KELLY MILLER

UNIVERSITY OF WISCONSIN-MADISON

Data Coordinating Centers (DCCs) are primarily responsible for providing expertise in statistical methods and data management of clinical trials. At the University of Wisconsin-Madison (UW), the DCC is a component of the Clinical Trials Program in the Department of Biostatistics and Medical Informatics (BMI) in the UW School of Medicine and Public Health. The DCC supports and is engaged in collaborative clinical research projects which are NIH- or industry-funded data coordinating centers. Ideally staffing would lend itself for all aspects of data management to be done by one team at one institution/organization but given space and resource limitations this does not always occur. Sometimes partnerships outside one's own institution are necessary to accomplish the end goal. The UW has had a long-standing relationship with Frontier Science and Technology Research Foundation (FSTRF) through the Eastern Cooperative Oncology Group since the1970s. FSTRF headquartered in Boston, Massachusetts has been engaged in clinical trials collaboration and support for over 40 years, providing efficient and cost-effective data management as well as statistical expertise and software development and maintenance for the collection and analysis of data to research networks, pharmaceutical companies and investigators. FSTRF has extensive experience in a wide range of studies, involving all phases. Focusing on the strengths and expertise of both groups and the limitations of each group independently, a unique collaboration was formed that addresses all aspects of clinical trial data coordination activities. Together both teams formed one stronger team to the benefit of the researchers that engage with them to answer important clinical questions. The INfluenza Vaccine to Effectively Stop Cardio Thoracic Events and Decompensated heart failure (INVESTED) trial (ClinicalTrials.gov: NCT02787044) showcases this unique collaboration. This presentation will describe the arrangement as a data coordinating center (U01 HL130204) between UW and FSTRF and their respective roles for the INVESTED trial.

Contributors

Dorianne Kalish
University of Wisconsin-Madison

Kenneth Wood
Frontier Science and Technology Research Foundation

Sue Siminski
Frontier Science and Technology Research Foundation

KyungMann Kim
University of Wisconsin-Madison

APPLICATION OF MULTI-STATE MODELS IN CANCER CLINICAL TRIALS

JENNIFER LE-RADEMACHER

MAYO CLINIC

Time-to-event endpoints are common in cancer trials and are commonly analyzed with Kaplan-Meier curves, logrank tests and Cox models. However, in trials with complex disease process and/or treatment options, multistate models (MSM) add important insights. This talk will focus on simple Aalen-Johansen estimates - the multistate analog of the Kaplan-Meier - via the analysis of a leukemia trial. The canonical path for a subject in the trial is a conditioning regimen (A or B), which leads to a complete response (CR), followed by consolidation therapy, and eventually followed by relapse and death. While standard survival methods look only at A vs. B in terms of overall survival, MSM can track all the intermediate states in a manner that is simple to compute and interpret. In our leukemia trial, MSM provides significant insights that the survival advantage observed in the experimental treatment results from its ability to both induce a faster CR and prolong survival once a patient achieved CR. Our goal is to encourage the use of MSM in cancer trials as they complement standard survival methods and may facilitate a better understanding of cancer disease process and its treatments.

Contributors

Ryan Peterson
University of Iowa

Terry Therneau
Mayo Clinic

Sumithra Mandrekar
Mayo Clinic

BUILDING THE FOUNDATION FOR A STATISTICS AND DATA COORDINATING CENTER WITH LIMITED RESOURCES: THE HERO STUDY, A NATIONWIDE MULTICENTER RANDOMIZED HEPATITIS C TREATMENT TRIAL

CRISTINA MURRAY-KREZAN

UNIVERSITY OF NEW MEXICO HEALTH SCIENCES CENTER

A Statistics and Data Coordinating Center (SDCC) that efficiently runs multiple clinical trials takes years of development to become fully operational. The staffing, infrastructure, and resources required are formidable, especially for relatively isolated institutions located in poor, rural, under-resourced states such as New Mexico. In an effort to become competitive with established academic SDCCs, the University of New Mexico Health Sciences Center (UNMHSC) must be innovative, collaborative, and use existing resources to build capacity and infrastructure for coordination of multiple multisite trials. To lay the groundwork, we formed the UNMHSC SDCC for an 8-site, two-arm, randomized, controlled trial, initiated in spring 2016. The Hepatitis C Real Options (HERO) study is evaluating modified directly observed therapy (mDOT) versus patient navigator (PN) assistance to deliver treatment for chronic hepatitis C (HCV) to participants who inject drugs. Participants receive ledipasvir-sofosbuvir in one of two health settings: an opiate agonist treatment program or a community health center. We utilize existing resources from our Clinical and Translational Science Center (CTSC) including REDCap for our electronic data capture system and programming support to automate data downloads for report generation. Data security and storage are maintained by the UNMHS Library and Informatics Center. Staff include two data managers, two biostatisticians, and the clinical site project director, contributing a combined 2.15 FTEs for drafting and building electronic data collection forms (eDCFs) in REDCap, data quality, database maintenance, and weekly and other periodic reporting. Standard operating procedures govern these processes and serve to train new staff. Our strengths include free access to REDCap and low-cost programming support via the CTSC; a SDCC biostatistician with industry SDCC experience; and supportive site collaborators who assist with development and testing of eDCFs. Challenges include securing experienced programming and data management staff, and inefficient institutional security processes such as granting external sites access to UNMHSC REDCap. Our successes are highlighted by the creation of bilingual eDCFs (Spanish and English), efficiently training sites with heterogeneous clinical trials experience, and maintaining high fidelity of data capture across sites. An innovative aspect to our SDCC's capabilities is collecting adherence data that is critical to the study's primary and secondary endpoints. HCV treatment is dispensed in blister packs with an electronic tracking function that records the date and time a dose is removed. Participants randomized to mDOT receive smart phones loaded with the eMocha virtual video capture app that records a video of the participants taking their HCV treatment along with a timestamp. The data are uploaded to a centralized, secured, back-end server. Participants who receive HCV treatment at a treatment program or community health center have adherence data manually recorded by nurses who directly observe participants taking their medication. The SDCC is developing the capability to seamlessly transfer data collected from these diverse platforms into the REDCap data repository. Future steps to build capacity at the UNMHSC SDCC include hiring additional staff to expand report generation, improve REDCap capability, coordinate central IRB oversight, increase training for users, and perform site and medical monitoring.

Contributors

Katherine Wagner
University of New Mexico Health Sciences Center

Herbert Davis
University of New Mexico Health Sciences Center

Dawen Liu
University of New Mexico Health Sciences Center

Moonseong Heo
Albert Einstein College of Medicine

Alain Litwin
University of South Carolina School of Medicine

Kimberly Page

University of New Mexico Health Sciences Center

CHALLENGES AND SUCCESSES IN RECRUITMENT OF HIV INFECTED PERSONS TO A CANCER PREVENTION STUDY

SHELLY LENSING

UNIVERSITY OF ARKANSAS FOR MEDICAL SCIENCES

Background: As deaths from AIDS decrease, people with HIV are at increasing risk of anal cancer. The incidence of anal cancer is 15 to 90 times higher in HIV-infected populations than in the general population, which has rates of 1.5 and 1.9 per 100,000 for men and women. Men who have sex with men (MSM) are particularly susceptible with rates of 131 cases per 100,000 for men who have sex with men (MSM), followed by heterosexual men and women, who have rates of 46 and 30 per 100,000.

Given the success in cervical cancer prevention through routine screening and removal of precancerous cells, a similar strategy was proposed for anal cancer prevention. The ANCHOR study was designed to test the hypothesis that treatment of anal high-grade squamous intraepithelial lesions (HSIL) will lead to a 75% reduction of incident anal cancer by comparing two strategies (HSIL treatment vs. watchful waiting). The treatment arm included infrared coagulation, electrocautery, or laser procedures, or imiquimod or 5-fluorouracil topical treatments, selected at the discretion of the treating physician. To achieve the goal of randomizing 5058 men and women 35 years of age and older with HSIL, approximately 15000 participants are targeted for screening.

Results: Since March 2015, over 4000 participants have been screened from 19 clinics in 13 cities, comprising a diverse HIV population with 21% women and 76% minority race/ethnicity. A national campaign was implemented in conjunction with support for local recruitment strategies and monetary incentives for participants. Clinic capacity in terms of staffing and clinic availability has been the biggest challenge for recruitment. In response, the number of sites have been increased from the 10 sites originally proposed, and site funding has been restructured to reward high enrollers.

Prevention studies differ from therapeutic clinical trials since the medical care is not medically necessary; there were concerns that some participants would not enroll due to copays and deductibles. Approval from the Office of the Inspector General was obtained for the study to pay for expenses not paid by insurance in the third year of the study. Twenty-two percent of participants have private insurance, 30% Medicare, and 58% Medicare.

Most screenees have heard about the study from their doctor/provider (44%) and/or friends/family (33%). Analysis of factors related to referral sources found that African Americans were more likely to be referred from family or friends, as were participants with Medicaid or without private insurance; however, non-African Americans or participants without Medicaid or Medicare insurance were more likely to be referred by a MD or provider.

Conclusion: The ANCHOR study has been successful in reaching a diverse population. Sites have unique populations and have been tailoring their recruitment strategies. To date, providers and family/friends have been by far the greatest source of referrals, but the national campaign may have a greater role later as providers deplete their pool of regular patients. Interestingly, there have

been minority and insurance differences for referral sources, and increasing insurance costs for privately insured participants and site capacity have been challenges.

Contributors

Jeannette Lee
University of Arkansas for Medical Sciences

J. Michael Berry
University of California, San Francisco

Joel Palefsky
University of California, San Francisco

CHALLENGES OF CONDUCTING A CLINICAL TRIAL IN YOUTH CORRECTIONAL FACILITIES

JEANNETTE LEE

UNIVERSITY OF ARKANSAS FOR MEDICAL SCIENCES

Chlamydia (CT) is the most prevalent bacterial sexually transmitted infection and causes significant morbidity. CT treatment trials have been limited by inability to control for re-exposure to untreated partners, making it difficult to distinguish treatment failure from reinfection. We conducted a randomized clinical trial of azithromycin (AZT) versus doxycycline (DOXY) for CT treatment in persons (aged 12 -21 years of age) in gender-segregated youth correctional facilities (YCFs), a setting with a high CT prevalence and minimal opportunities for re-exposure post-treatment. The study objective was to demonstrate the noninferiority of AZT to DOXY with respect to treatment failure rate at Day 28. Because the duration of treatment differed between doxycycline and azithromycin, the study was not masked. The primary analysis population for this study was the per-protocol population which was defined as all randomized participants who completed therapy, whose failure status could be established at the day 28 visit, and who took a single dose of azithromycin or 10 of 14 prescribed doxycycline doses.

Conducting a trial in YCFs poses significant challenges including: 1) protocol review and approval by multiple IRBs, the federal Office for Human Research Protections, the local Superior Court Juvenile Division, and other stakeholder agencies including the local child welfare and juvenile corrections departments, 2) the inability to use incentives, and 3) uncertainty of a participant's release date, and 4) frequent transfers between detention facilities. These challenges can lengthen and complicate the planning prior to starting the study, prolong the time necessary to reach target enrollment, and create hurdles in adhering to a rigid protocol. In our study, a high proportion of enrolled participants were released early due to policy changes within the local juvenile delinquency system; this meant that these participants were unavailable to assess treatment outcome and were not included in the per-protocol population.

Strategies were implemented to address these challenges. For example, although the trial was originally directed at females, it was expanded to include males both as an effort to increase enrollment and because of emerging evidence suggesting azithromycin was less effective in males. The study ultimately enrolled 567 persons to obtain 310 individuals in the per-protocol population, which was the analysis population for this study. Of the 257 study participants who did not meet the criteria for inclusion into the per-protocol population, 185 (72%) were discharged early from the facility.

The rationale for conducting this study in youth correctional facilities was to minimize the risk of new chlamydial infections that would be difficult to distinguish from persistent infections. In this environment, only 55% of the enrolled participants successfully completed the trial.

Contributors

Shelly Lensing

University or Arkansas for Medical Sciences

Raymond Perry
Los Angeles Department of Health Services

William Geisler
University of Alabama at Birmingham

CHALLENGES IN CONDUCTING A MULTI-CENTER TRIAL IN US MILITARY HOSPITALS

ROBERTA SCHERER

JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH

Background: Tinnitus is the perception of sound in the absence of external sound and may sometimes be caused by loud noises. For some individuals, tinnitus is intolerable and impacts quality of life. Tinnitus retraining therapy (TRT) uses tinnitus-specific counseling (TC) and sound therapy (ST) to facilitate habituation to the tinnitus, improving quality of life.

Methods: The Tinnitus Retraining Therapy Trial (TRTT), a randomized, double-blind, placebo-controlled, multi-center trial for individuals with intolerable tinnitus, evaluated the efficacy of TRT and its two components. Treatment groups included TRT (TC and conventional sound generators (SGs)), partial TRT (TC and placebo SGs) and standard of care. All groups were encouraged to use environmental sound enrichment.   The primary outcome was change in score on the Tinnitus Questionnaire assessed longitudinally between treatment and follow-up (3, 6, 12 and 18 months). The TRTT was conducted at US military hospitals, based on the assumption that these sites would provide an enriched population due to noise exposure and as federal sites, were exempt from indirect costs associated with NIH funding and clinicians could donate time and effort to the trial.

Results: We experienced unexpected difficulties in conducting the TRTT in US military hospitals. Site selection took considerable up-front time and effort by the trial's principal investigator. Of 19 potential sites visited, 10 expressed interest, 8 completed certification, 6 enrolled study participants, and 5 completed the trial. Once selected, each site needed a Cooperative Research and Development Agreements to receive funds; obtaining all required signatures often took longer than expected (range = 6 to 26 months). Delays in institutional review approval were due to separate independent boards with most meeting infrequently. We learned that support from the entire chain of command (site investigators, commanders, consultants, and if possible, Surgeon General) is essential because continuity of support at a site was not assured when changes in authority occurred. Staff replacement due to retirement or transfer occurred frequently and often resulted in an extended period of inactivity due to military hiring restrictions and prolonged security checks. During the trial, 54 persons were certified at 6 participating sites, including 19 coordinators, 29 audiologists, and 16 others. Rules regarding recruitment materials were neither uniform across all sites nor consistent over time at single sites. Difficulties in scheduling study visits resulted from staff or participant transfers, temporary duty assignments, and military priorities overriding scheduled visits. Federal government barriers manifested as Base Realignment and Closure at trial start and sequestration during the trial; the latter resulted in restricted workloads with priority for clinical responsibilities. Security issues included firewalls preventing internet access to the TRTT web-based database and restrictions on civilian staff prior to clearance for computer access. On the positive side however, the TRTT study population was diverse (30% women; 27% minorities, matching the US armed forces demographics). We had few protocol violations, and audiologists were enthusiastic about learning a new approach for tinnitus treatment.

Conclusion: Although it was possible to conduct a multi-center trial in US military hospitals, planning and perseverance was required.

<u>Contributors</u>

Lee Sensinger
Johns Hopkins Bloomberg School of Public Health

Craig Formby
University of Alabama

 TRTT Research Group

DESIGN OF AN ELECTRONIC CIRB SUBMISSION MODULE INTEGRATED WITHIN A CLINICAL TRIAL MANAGEMENT SYSTEM

KEITH PAULS

MEDICAL UNIVERSITY OF SOUTH CAROLINA

The National Institutes of Health recently issued a policy effective January 25, 2018 on the use of a single Institutional Review Board for multicenter clinical trials. While this policy offers many opportunities for efficiencies, there remain inherent redundancies associated with site submission of regulatory documents and team member information to both the Clinical Coordinating Center and the central Institutional Review Board (cIRB). Redundant submission of this information is an inefficient use of resources and makes the data vulnerable to discrepancies, typos and other data entry errors. To avoid these pitfalls, a comprehensive cIRB submission module was developed and integrated into a web-based EDC-CTMS system for the SIREN Network.

The SIREN CTMS already contained an electronic regulatory document management system, electronic delegation of authority log, and various other site and project data points which are also required by the cIRB, while other information needed by the cIRB had to be added to the CTMS interfaces for collection. It's important to note that certain data points required for cIRB review are the same for a clinical site regardless of the protocol. Likewise, certain information will be the same across all clinical sites participating in a specific trial. By centrally capturing this data, it can be stored in one place and exported as needed. Using this model, the data for cIRB review is aggregated in the CTMS, reviewed for correctness, converted into an electronically transferable package, and sent by secure FTP to the cIRB where it is imported into their review system. By working with the cIRB and utilizing standard technology, a fast and efficient process has been developed to streamline cIRB submissions by mitigating redundancies, data entry errors and bottlenecks that slow down and hamper the review process.

Contributors

Wenle Zhao
Medical University of South Carolina

Valerie Durkalski
Medical University of South Carolina

Robert Silbergleit
University of Michigan

Deneil Harney
University of Michigan

Chris Arnaud
Medical University of South Carolina

DO SITE VISITS IMPROVE PARTICIPANT RE-ENGAGEMENT IN LONGITUDINAL CLINICAL RESEARCH?

MATHEW MARGOLIS

WASHINGTON UNIVERSITY SCHOOL OF MEDICINE

Introduction: Despite the near universal reliance on site visits in multi-center clinical research, there has been little documentation on their value for participant re-engagement. This study reports the impact of site visits on improving participant re-engagement and ascertainment in a cohort of older adults approximately twenty years after study onset.

Methods: A retrospective review of all site visits during OHTS 3 was conducted. Clinics were site visited if less than 50% of the surviving participants had been seen. Data regarding the number participant visits and ascertainment (visits, declines, no further contact requests, and notice of death) were collected for the five months pre- and five months post-site visit. This ten month window was selected to avoid influence of a yearly coordinators meeting, which addresses re-engagement. Data regarding the cost of site visits were collected from coordinating center records.

Results: A total of 11 site visits were conducted. None were excluded from analysis. In the five pre-site visit months the average number of patient visits was 3.6 (2.6 to 4.7) compared to 3.1 (1.5 to 5.8) visits five months post-site visit. In the five pre-site visit months the average number of ascertainments per month was 7.2 (5.0 to 8.7) compared to 5.7 (1.7 to 10.3) in the five months post-site visit. In the second month after site visits (weeks 4-8), both patient visits and ascertainments were at their highest levels in our ten-month window (5.8 and 10.3, respectively). When adjusted proportionally to account for the decreased participant pool over time, patient visits were elevated the second month after site visits, and ascertainments were elevated two to four months (weeks 4-12) post-site visit. Adjusted values for monthly visits and ascertainments were higher in the five months following site visits than the five preceding months (12.8 vs 4.2 and 9.0 vs 4.0, respectively). The average cost per visit was $2105.65 ($960 to $2587.76).

Conclusions: This study demonstrates a transient absolute increase in participant re-engagement and ascertainment following site visits. Correction for diminishing pool of participants via proportionate adjustment showed a modest sustained increase in re-engagement and ascertainment throughout the five-month post-site visit months. This report demonstrates the ability of site visits to potentiate other re-engagement efforts in long term clinical research. Coordinating centers must weigh the financial costs of site visits against potential gains in patient re-engagement and ascertainment status.

Contributors

Bradley Wilson
Washington University School of Medicine

Patricia Morris
Washington University School of Medicine

Mae Gordon
Washington University School of Medicine

Michael Kass
Washington University School of Medicine

GRAPHICAL SUMMARIES OF TRIAL CONDUCT IN DMC REPORTS

MELISSA SCHULTZ

UNIVERSITY OF WISCONSIN

For many clinical trials, Data Monitoring Committees (DMCs) are charged with monitoring, not only the safety and efficacy of an intervention, but also the conduct of the trial itself. A study with low accrual, high dropout, or an unacceptable lag in data collection, adverse event coding or endpoint adjudication may not have information of sufficient quality for monitoring, and may ultimately prove unable to answer the clinical questions of interest.

The Statistical Data Analysis Center (SDAC) at the University of Wisconsin--Madison specializes in producing interim reports and analyses for DMCs. Our reports are graphically based, allowing DMC members to easily identify differences between treatment groups, and/or changes over time, and to review a large amount of information in a short amount of time.

In order to interpret the safety and efficacy data presented in a DMC report, the DMC must be able to place the information in context and evaluate whether the trial is being conducted in a way that allows them to fulfill their responsibilities.

The following questions are relevant to evaluating recruitment and retention:

* Is enrollment meeting projections? Are a few clinical centers dominating enrollment?

* Which countries are contributing the most subjects? Is treatment assignment roughly balanced within each country? What are the enrollment patterns over time in different geographic regions?

* Are treatment groups balanced within stratification levels?

* How many participants are currently in each phase of the study (e.g., double-blind, open-label, safety follow-up)?

* Are subjects adhering to the treatment plan? What are the patterns of treatment termination, by treatment arm? Are there differences in incidence, timing or reasons?

* Are study discontinuations clearly distinguished from treatment terminations?

* What is the distribution of follow-up time in each treatment group?

While cumulative accrual is often presented graphically, many other aspects of recruitment and retention are typically presented in tables, if at all. We have found that graphical approaches make it much easier for DMC members to get a feel for the data.

Data quality and quantity are also key components of trial conduct. It is important for members of the DMC to understand how much follow-up information is available, how current the data are, whether there are sites with a large number of eligibility violations or other important protocol deviations, and whether adverse event coding and endpoint adjudication are acceptably up to date.

Our DMC reports typically include several figures summarizing both data availability and adjudication of clinical events, allowing the DMC to easily comprehend, not only what they are seeing, but also what information is missing and potentially impeding their ability to adequately monitor the trial.

We will present examples of figures which address many of the above questions and which have proven invaluable to the DMCs we have served.

Contributors

Robin Bechhofer
University of Wisconsin

IS INTRAOCULAR PRESSURE, A SURROGATE OUTCOME, IMPORTANT TO PATIENTS WITH GLAUCOMA?

JIMMY LE

JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH

INTRODUCTION/OBJECTIVE

Choosing the wrong outcomes can make a clinical trial irrelevant. We sought to identify outcomes that are important to patients with glaucoma and prioritize outcomes for use in clinical trials in glaucoma.

BACKGROUND

The validity of a surrogate outcome depends on the level of certainty that it predicts an effect on the true clinical outcome of interest. In glaucoma research, intraocular pressure (IOP) is the most widely-used surrogate outcome for adequate disease control. But is IOP an outcome that is important to patients with glaucoma? We have previously shown that less than a third of outcomes in studies that investigate a new type of glaucoma device are patient-centric.

METHODS

Working in collaboration with investigators at the Food and Drug Administration, American Glaucoma Society, and Johns Hopkins University, we conducted a mixed method study to identify and prioritize outcomes that are most important to patients with glaucoma.

We first conducted semi-structured, one-on-one interviews with patients recently diagnosed with glaucoma at the Wilmer Eye Institute, Baltimore, MD. We used the Framework Approach to analyze the qualitative data and identify outcomes patients discussed. We then designed a cross-sectional survey using Best-Worst Scaling ("Case-1") to elicit patients' preferences for outcomes. We surveyed participants from Arkansas, California, Maryland, and Texas. We analyzed response patterns using conditional logistic regression to determine the relative importance of different outcomes.

RESULTS

We interviewed 25 patients between May and December 2016 and identified 13 outcomes that patients expressed as important. One hundred patients responded to our survey between September 7 and November 4, 2017, and we will continue recruiting until reaching 300, the target sample size. Of the first 100 participants, the majority were 65 years of age or older (55%) and currently taking IOP lowering eye drops (66%) to manage their glaucoma. We found that control of IOP is among the outcome domains that patients identified as most important. Patients also expressed that the ability to perform other vision-dependent activities (e.g., driving a car, navigating outside of the house) and maintain visual function and quality of vision (e.g., depth perception) are more important than burden of treatment (e.g., reduction of number of drops, cosmesis).

DISCUSSION

There is no question that glaucoma control is related to the level of pressure within the eye. While our survey suggests that patients prioritize IOP when making treatment decision, it also demonstrates that other outcomes, particularly those associated with vision-dependent activities such as driving and navigating, also matter. It has been postulated that the perceived significance of IOP "is not of concern to patients until they have been educated, or rather miseducated, to believe that there are linear relationships" between eye pressure and what the patient can do or how well they feel (Spaeth, IOVS, 2011). While using surrogate outcomes can help expedite the development of trials and approval of new interventions, we should pay more attention to what keeps patients functioning and feeling optimally. The purpose of glaucoma treatment is more than just controlling IOP; it is also about maintaining or enhancing the health and well-being of patients.

Contributors

Amanda Bicket
Wilmer Eye Institute, Johns Hopkins School of Medicine

Michelle Tarver
United States Food and Drug Administration

Malvina Eydelman
United States Food and Drug Administration

John Bridges
Johns Hopkins Bloomberg School of Public Health

Tianjing Li
Johns Hopkins Bloomberg School of Public Health

LEVERAGING THE ELECTRONIC HEALTH RECORD IN A GLOBAL OUTCOMES TRIAL: DESIGN AND RATIONALE OF THE HARMONY OUTCOMES EHR ANCILLARY

SUDHA RAMAN

DUKE UNIVERSITY SCHOOL OF MEDICINE

Background. Electronic health records (EHR) are a rich source of clinical data that are increasingly used in pragmatic health research initiatives, but the hypothesis that EHR data are fit for use in clinical research has not been rigorously evaluated.

Methods. The HARMONY EHR Ancillary Study (AS), will evaluate the fitness of EHR data for use in identifying a trial-eligible cohort, populating baseline characteristics and identifying trial endpoints as part of a large, multinational clinical trial. The EHR AS is being conducted alongside the HARMONY Outcomes Trial, an ongoing, double-blind, randomized controlled study to evaluate the effect of albiglutide on cardiovascular events in patients with Type 2 diabetes. Objective 1 of the AS will use site data from study coordinator surveys, workflow assessment, and qualitative analysis of a site-level enrollment measure to enhance understanding of how EHR data are used to facilitate clinical trial enrollment at sites with an EHR system. Objective 2 will assess the level of agreement between select data documented on the baseline eCRF (reference standard) and 1) data extracted from the EHR at select US sites participating in the HARMONY Outcomes Trial, and 2) data extracted from existing national-level electronic health data and administrative claims data within select countries having sites participating in the HARMONY Outcomes Trial. Objective 3 will apply case-finding algorithms to these same data sources to evaluate concordance between select clinical endpoints identified in EHR datasets and those documented in the main trial database.

Conclusions. The HARMONY Outcomes EHR AS will evaluate the potential for EHR data to be used for high-quality clinical trials. By using EHR data to identify a trial-eligible cohort, populate baseline characteristics and identify trial endpoints alongside a multinational clinical trial, this study will provide evidence to support the optimal use of EHR data in pragmatic clinical trials.

Contributors

Emily O'Brien
Duke University School of Medicine

Bradley Hammill
Duke University School of Medicine

Lisa Berdan
Duke Clinical Research Institute

Tyrus Rorick
Duke Clinical Research Institute

Stefano del Prato
University of Pisa School of Medicine

Salim Janmohamed

GlaxoSmithKline, Inc.

Adrian Hernandez

Duke University School of Medicine

Lesley Curtis

Duke University School of Medicine

MODELING CORRELATED OCULAR OUTCOMES WHEN BOTH EYES FROM SOME BUT NOT ALL PARTICIPANTS ARE ENROLLED

DANNI LIU

JAEB CENTER FOR HEALTH RESEARCH

Many ocular disorders affect both eyes, so it is not uncommon for clinical trials to randomize one or both eyes of study participants, referred to as unilateral and bilateral enrollment respectively. An example of this is the hybrid design, which permits both unilateral and bilateral enrollment into a study depending on the number of eyes meeting eligibility criteria at enrollment. Advantages of the hybrid design include faster recruitment of eyes—the unit for sample size—and lower costs, but the design introduces additional complexity to the protocol, site operations, and analysis requiring careful forethought about the tradeoffs involved.

As ocular outcomes from the eyes of bilateral participants are usually correlated, it is necessary to appropriately account for the inter-eye correlation when conducting statistical analyses. Two approaches have been widely used to address this correlation: 1) mixed effects models using maximum likelihood estimation (MLE) with a subject-specific random effect, and 2) marginal models using generalized estimating equations (GEE) with a working correlation structure for residuals.

Several studies have discussed the characteristics of the two models when analyzing data with correlated ocular outcomes. In most cases, the choice of approach to account for the inter-eye correlation has little impact on statistical inference, especially when the sample size is relatively large. However, there are unusual cases where this choice greatly impacts both statistical inference and interpretation of the study.

We will present two examples where the two statistical approaches yielded substantially different results when both eye- and participant-level covariates were included in the models for eye-level outcomes. The examples are taken from two studies, one conducted by the Diabetic Retinopathy Clinical Research Network (DRCR.net) and the other by the Pediatric Eye Disease Investigator Group (PEDIG). Analyses will be performed using the MIXED and GENMOD procedures in SAS® 9.4, which implement the MLE and GEE methods respectively.

We will illustrate how the two procedures handle between-subject correlation and within-subject correlation, and we will provide some recommendations for modeling correlated outcomes in similar situations.

Contributors

Trevano Dean
Jaeb Center for Health Research

Michele Melia
Jaeb Center for Health Research

Adam Glassman
Jaeb Center for Health Research

Ray Kraker
Jaeb Center for Health Research

(null)

STANDARDIZED REPORTING OF RANDOMIZED CLINICAL TRIALS FOR SYSTEMATIC REVIEWS: SYSTEMATIC REVIEW DATA REPOSITORY

RIAZ QURESHI

JOHNS HOPKINS UNIVERSITY, BLOOMBERG SCHOOL OF PUBLIC HEALTH

Background:

Despite the existence of reporting standards for clinical trial design and methods (e.g., CONSORT), reporting is incomplete, inconsistent, across multiple data sources, and threatens the ability of systematic reviewers to assess the risk of bias in the trials and synthesize data appropriately. The open data movement is driven by principles of increased transparency, improved access to and use of trial data, and minimizing waste. The same principles apply to systematic reviews ("reviews") of randomized clinical trials (RCTs). Cochrane Eyes and Vision, US Satellite (CEV@US) has been working to make available trial data it extracts for reviews, and data included in the related meta-analyses, through the open access Systematic Review Data Repository (SRDR): https://srdr.ahrq.gov.

Objectives:

To examine extracted trial design and methods data in CEV@US reviews in preparation for making Cochrane systematic review data open access.

Methods:

We examined a convenience sample of six CEV@US reviews conducted between 2012 and 2016; the six reviews included 29 RCTs. We collected data from the reviews which overall covered 27 extracted design and methods data items. We mapped the data items to the CONSORT 2010 checklist for reporting of clinical trials and one item (description of interim analyses or stopping guidelines) had to be added as it was not present in any of the reviews. We relied on the systematic reviewers' extractions for each study and thus do not know whether trial investigators failed to report the elements that were missing or the items were reported but not found by the reviewers. We assume that, per Cochrane standards, all public sources of trial data were checked by two extractors; thus, we assume that the information collected for each trial and presented in the review was the sum of information available for the trial across sources.

Results:

The characteristics of design and methods for six CEV@US reviews that were most commonly reported in the clinical trials (>90%) were: countries where participants were enrolled (100%), description of interventions (100%), description of outcomes (100%), number randomized (100%), losses to follow up (97%), description of masking (93%), and length of follow up (93%). None (0%) of the included trials reported a description of planned interim analyses or stopping rules. We found that trial registration in ClinicalTrials.gov was also poorly reported for each trial, with only 2/29 (7%) trial reports providing information. Table 1 presents the full list of design and methods characteristics and the proportions of included trials reporting each characteristic. We

used these data to develop a standardized SRDR data extraction form to facilitate consistent reporting in CEV@US systematic reviews.

Conclusions:

We found that reporting of RCT design and methods appears to have been inconsistent among the 29 RCTs included in our sample of six CEV@US reviews. We recommend that trial investigators publicly report trials using CONSORT and make all sources of trial information easy to find. Systematic reviewers should extract data consistently from trial reports, following PRISMA standards, into an open access database such as SRDR.

Contributors

Kristina Lindsley
Johns Hopkins University, Bloomberg School of Public Health, Cochrane Eyes and Vision

Bryant Smith
Brown University Center for Evidence Synthesis in Health, SRDR Administrator

Kay Dickersin
Johns Hopkins University, Bloomberg School of Public Health, Cochrane Eyes and Vision

THE HABIT FORMATION TRIAL TO EVALUATE BEHAVIORAL ECONOMIC INTERVENTIONS TO REDUCE THE RISK OF CARDIOVASCULAR DISEASE: LESSONS LEARNED

MARY PUTT

UNIVERSITY OF PENNSYLVANIA

Background:

Using principles of behavioral economics, the Habit Formation trial assessed interventions designed to improve sustained adherence to statin medication and reduce lipid levels in subjects at risk of cardiovascular disease. The trial considered several financial incentives to first achieve reductions in low density lipoproteins (LDL), and then maintain LDL levels after incentives end.

Study Design:

Habit Formation is a four-arm randomized controlled trial. All subjects received wireless-enabled pill bottles, to record adherence through bottle opening and provide daily adherence reminders. The Control arm received no further intervention. The financial incentives arms were active for six months, and included Simple Daily Sweepstakes based on daily medication adherence, a Habit Formation strategy requiring adherence prior to a daily timed reminder, and a Hybrid Sweepstakes/Deposit strategy with incentives divided between daily sweepstakes and a monthly deposit. Collection of all trial-related data used our web-based Way to Health platform. Six months after the incentives ended, we measured the primary outcome, change from baseline LDL, and determined the effectiveness of the incentives in yielding sustained reduction in LDL.

Recruitment:

We recruited subjects at risk of complications of cardiovascular disease with evidence of poor statin adherence and sub-optimal LDL levels. The target recruitment (800 subjects) was designed to provide over 90% power to detect a 10 mg/dl difference in LDL between interventions and control. Two different recruitment strategies yielded strikingly different success rates. First, using pharmacy records for employees from several corporations, we identified potential subjects with a statin Medication Possession Ratio less than 80%, offered these individuals a lab measurement and then, if eligible, enrollment. Our yield was only 0.7%. Subsequently, working with practices in the University of Pennsylvania Health System, we identified potential subjects from usual-care electronic medical records. These individuals were eligible if they self-reported incomplete adherence, and generally did not require a separate baseline lab measurement since lab results were available in the medical records. After the initial invitation by mail, we followed up by phone or text. Here, our yield improved by over 20 fold, to 16.7%.

Managing a Crisis:

We overcame significant challenges with our pill bottles. Our first device experienced catastrophic failure as the network removed their 3G cell towers and transitioned to 4G. Our second device had high failure rates and was not reliably supplied by the vendor, causing us to suspend our trial for several months. Our third device proved reliable and readily available. We learned to be vigilant in order to quickly identify and trouble-shoot device problems. We incorporated sensitivity analyses into our statistical analysis plan to assess any impact of device malfunction and/or the effects of different pill bottles.

Conclusions:

For the Habit Formation trial, recruitment was most effective when the burden of meeting eligibility criteria was minimized for the subject, and when multiple avenues of reaching potential subjects were available. Appropriate strategies for managing device failure were critical to trial success. Results of the Habit Formation trial will add to our understanding of optimally structuring incentives to motivate durable behavior change.

Contributors

Peter Reese
University of Pennsylvania

Iwan Barankay
University of Pennsylvania

Louise Russell
University of Pennsylvania

David Pagnotti
University of Pennsylvania

Ryan McGillovary
University of Pennsylvania

Dara Finnerty

University of Pennsylvania

Sakshum Chada

University of Pennsylvania

Karen Hoffer

University of Pennsylvania

Kevin Volpp

University of Pennsylvania

THE NATIONAL INSTITUTE OF NEUROLOGY DISORDERS AND STROKE CLINICAL TRIALS METHODOLOGY COURSE: OUTCOMES OF TRAINEES AFTER 3 YEARS

NICHOLAS SEEDORFF

UNIVERSITY OF IOWA

Introduction

The NINDS Clinical Trials Methodology Course (CTMC) is a joint venture involving investigators and program directors from the Neurological Emergencies Treatment Trials Network (NETT) and NeuroNEXT. The primary aim of the CTMC is to help train the next generation of clinical trialists in the clinical neurosciences. The overarching goal of our course was to provide the expertise in neurology clinical trials from the domains of clinical medicine and biostatistics to a broad group of investigators who were new to clinical trials. Here we characterize the outcomes of our participants, with the focus on if participants receive funding and enroll patients in projects stemming from the course.

Methods

We recruited and enrolled our first cohort in 2014. We initiated our website (http://neurotrials.training) to serve as a central repository of information about the course, and to provide learners with resources. We used a blended approach of webinars, small groups, and a residential course. The primary deliverable of the course participants was a functional clinical trial protocol. We longitudinally surveyed our participants for outcome information relating to the study. As a way to summarize results from the course, the proportions of participants to reach several key outcomes were calculated over time and plotted longitudinally.

Results

Over 3 years we enrolled 23 trainees who were designing trials in the clinical neurosciences. Several specialties and disciplines were represented, including neurology, emergency medicine, PM&R, and trauma surgery amongst others. By year three, over 40 percent of the trainees had enrolled their first patient (Figure 1), and nearly two thirds had submitted proposals.

Conclusions

Since many academic investigator initiated clinical trials fail to enroll any subjects, the preliminary results from the NINDS CTMC are encouraging. Clinical trials are challenging to initiate for early stage investigators. A limitation of this work is that we are not reporting on projects that may have been transitioned to non-interventional designs given the maturity of the science. Additional efforts will focus on further accelerating the growth of new investigators starting early phase clinical trials in the neurosciences.

<u>Contributors</u>

Joy Black

Robin Conwit
Roger Lewis

Chris Coffey

Laurie Gutmann
William Meurer

TREATMENT ADHERENCE IN THE TINNITUS RETRAINING THERAPY TRIAL

ROBERTA SCHERER

JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH

Background: Tinnitus, the perception of sound in the absence of external sound, is extremely distressing for some individuals. There are no medical interventions known to alleviate the tinnitus signal, so current treatment options are typically based on minimizing the reaction to and perception of the tinnitus to improve quality of life. The Tinnitus Retraining Therapy Trial (TRTT), conducted at US military hospitals, randomized individuals with severe tinnitus to tinnitus retraining therapy (TRT), Partial TRT or standard of care (SoC). TRT involves tinnitus-specific counseling (TC) and sound therapy. The TRT group used conventional sound generators and the Partial TRT group placebo sound generators to implement sound therapy. SoC involved counseling using a patient-centered care approach. To avoid a counselor effect, study audiologists administered both types of counseling in the TRTT, creating a challenge for managing treatment adherence.

Methods: We developed methods to maintain treatment adherence that involved training, certification, use of scripts and visual aids, and fidelity monitoring. To prepare for certification, protocol monitors identified critical topics for each type of counseling session, prepared scripts, and helped develop checklists covering those topics. Certification required attendance at full-day regional training sessions and review of prepared TC and SoC videos. Audiologists not able to attend a training meeting viewed a webinar or videos that emulated the training session. Study audiologists submitted a voice recording of one TC and one SoC counseling session, each with a non-study individual. Protocol monitors reviewed the recordings, and certified only those audiologists who adhered to the protocol. During the trial, audiologists used treatment-specific aids: scripts, a model of the ear, and handouts for both TC and SoC; and a flip-chart for TC with talking points identified for the counselor. Audiologists completed treatment-specific checklists during each counseling session, indicating topics covered or discussed. They submitted their recordings of the first two study counseling sessions of each type, which the protocol monitors reviewed by completing a separate checklist to verify topics covered. If no deficiencies were noted, then subsequent recordings were randomly selected for review. Whenever deficiencies were noted, the Protocol Monitor communicated with the audiologist in writing and by phone to discuss the deficiency and the next two counseling sessions were submitted for review. Continued non-adherence resulted in de-certification.

Results: 25 audiologists were certified for TC and/or SoC counseling and 24 completed at least one counseling session. Excellent adherence was obtained on 33 items for 74 TC counseling sessions reviewed by the Protocol Monitor (median, 100%, range: 70.3% to 100%), with no difference between adherence for TRT (median, 97.3%) and Partial TRT (median, 100%). Adherence on 44 items for 30 checklists for SoC counseling also showed reasonably good adherence (median, 88.5%, range: 42.3% to 100%).

Conclusion: By using multiple methods to address treatment adherence, the TRTT found audiologists demonstrated fidelity to distinct types of counseling sessions, a critical step in evaluating the efficacy of the study interventions.

Contributors

Sue Ann Erdman

Susan Gold
Craig Formby
University of Alabama

 TRTT Research Group

USE OF QUALITATIVE DATA ANALYSIS SOFTWARE TO FACILITATE A SYSTEMATIC LITERATURE REVIEW

BRIAN PERRY

DUKE UNIVERSITY

Introduction

A systematic literature review is a valuable tool for identifying important research gaps and/or best practices in clinical research. The process, however, can be cumbersome and overwhelming. To allow for replicability and for others to assess the review's rigor and potential relevance, the reviewers' methods must be clearly documented and auditable. The use of qualitative data analysis (QDA) software is routinely used to store, extract, and categorize text and visual data by qualitative researchers. Likewise, it can support processes involved in a systematic literature review. Here, we describe our use of QDA software (NVivo 11) to systematically screen, extract, and summarize data for a literature review on the use of mobile technologies to measure outcomes in clinical research, conducted by the Clinical Trials Transformation Initiative, and the benefits and challenges of using such software for this purpose.

Methods

We retrieved potentially relevant citations in PubMed using a discreet search framework. We screened each title for specific inclusion/exclusion criteria developed by a multi-disciplinary team and removed duplicate citations. Annotated citations (including abstracts) of remaining articles were uploaded to NVivo. In NVivo, nodes pertaining to each of the inclusion/exclusion criteria were applied to the citations/abstracts to screen for relevance. After screening, PDFs of relevant articles were uploaded to NVivo for data extraction. We created a codebook in NVivo to deductively capture and categorize elements in the literature related to our objectives, including the research design, therapeutic condition/s, mobile device(s) used, device placement, sampling rate, mobile outcomes, and research endpoints. Sub-nodes were used to explicate the type and variety of information in each element.

Results

There were benefits and challenges to using NVivo for this purpose. Various types of sources can be uploaded; however, the ability to extract data verbatim can vary depending on the source format. Additionally, queries can be used to identify potentially relevant information, but the capability is limited to text documents (TXT, text-based PDF, DOC, etc.). Classifications can help to organize sources (e.g., by date, author). Data visualization tools can be used to examine crosscutting themes and concepts (e.g., word clouds, word trees, coding matrices). Coding matrices can identify text coded at multiple nodes (e.g., "mobile outcomes" and "therapeutic condition/s"). When multiple reviewers are used, coding comparison queries can assess inter-reviewer reliability. Finally, all sources, text segmentation, and queries can be archived and are available for further analysis or audit. However, reviewers may have limited expertise with the software. Additionally, NVivo is not

commonly available across organizations, though a free 14-day trial version can be obtained with limited functionality. Finally, text segmentation and code application can vary by reviewer. It is important to establish a robust data analysis plan and to train each reviewer to segment text appropriately.

Conclusion

The use of QDA software can aid in storing, extracting and summarizing text. Challenges to using the software to facilitate a systematic literature review are minimal. Tools offered in NVivo may aid clinical researchers in reviewing and summarizing large, complex literature datasets following a systematic process.

<u>Contributors</u>

USING A PROFESSIONAL TRACING SERVICE TO LOCATE PARTICIPANTS 6 YEARS AFTER STUDY CLOSE-OUT

PATRICIA MORRIS

WASHINGTON UNIVERSITY

Purpose: The Ocular Hypertension Treatment Study (OHTS) was a glaucoma treatment trial that randomized 1,636 participants across 32 clinical centers between 1994-1996, with extended follow-up to 2009. Six years after study close-out, the NIH awarded funds for a 20-year follow-up to determine clinical and QOL status of participants. We describe how a professional tracing service helped to increase ascertainment rate by providing information on those lost to follow-up.

Methods: The OHTS Coordinating Center identified and sent last known contact information for participants lost to follow-up to Battelle Memorial Institute's Tracing Department. Last known contact information included name, maiden name (if applicable), last known address, date of birth, SSN, and spouse/next of kin. Battelle searched several proprietary databases for current contact information and vital status. Participant data were sent to Battelle for re-tracing up to three times. Battelle sent tracing results to the OHTS Coordinating Center, which relayed this information to clinical centers. A participant was defined as "ascertained" when they completed a clinic visit, QOL survey, or were confirmed deceased.

Results: To date 380 participant data lost to follow up have been sent to Battelle for tracing. Of these participants, contact information was found for 72% (275 of 380), and deaths confirmed for 20% (77 of 380). This information was generated over three successive rounds of tracing, which resulted in 81%, 16%, and 3% of Battelle's findings, respectively. Contact information for participants provided by Battelle resulted in 44 completed clinic visits, and 35 telephone QOL surveys. Overall ascertainment rate increased by 10% from 66.6% (1,089/1,636 of the randomized participants) to 76.1% (1,245/1,636). Still pending are 224 participants who have been traced by Battelle, but have either not yet completed visits, or been successfully contacted by clinics. The cost of tracing averaged $159.84 per participant ($60,739.20 in total).

Conclusions: Battelle Memorial tracing provided contact information and vital status for participants who would otherwise never have been ascertained. To date, the retrieval of contact/vital status data by Battelle increased the overall ascertainment rate from 66.6% to 76.1% of the 1,636 participants originally randomized more than 20 years ago. Researchers must weigh the benefits of tracing services against their costs.

Contributors

Deborah Dunn
Washington University

Bradley Wilson
Washington University

Leonard Haertter
Washington University

Ruth Gatiba
Battelle Memorial Institute

Lisa John
Battelle Memorial Institute

Mathew Margolis

Washington University

Michael Kass

Washington University

Mae Gordon

Washington University

USING PHASE 2 CLINICAL TRIAL DESIGN TO ENHANCE THE PHASE 3 PROCESS

RON MARKS

CLINIPACE WORLDWIDE

Introduction: Phase 2 clinical trial design is often underutilized to provide critical information that can improve the chances for a successful Phase 3 clinical trial program to lead to product registration. This talk will describe the opportunities for the Phase 2 program to provide needed information to benefit Phase 3.

Phase 3 Research Design Needs: Critical clinical trial needs for a Phase 3 clinical trial to lead to successful registration generally include ensuring final dose has been determined, final defined study population is set, identification of primary and secondary objectives, most relevant primary and secondary endpoints, most relevant comparator arm(s), determination of relative efficacy of IP to comparator (not enough to just document IP efficacy), determine whether result will be for superiority of IP or non-inferiority, and determination of sample size. There may in some studies be additional needs for a successful Phase 3 program, and it may be that not all this list is needed. However, this list should serve as a fairly complete set of needs for all Phase 3 clinical trial programs.

Phase 2 Clinical Trial Design Needs: The Phase 2 clinical trial program has the opportunity to provide all the needed info on the critical Phase 3 questions needed for a sponsor to maximize their confidence in the Phase 3 program to be successful in achieving product registration. For example, some Phase 2 studies start with a small study providing all subjects with the IP to establish that some level of efficacy is present to continue with Phase 2 research. Such a study should also include a comparator group to also provide critically important information on the relative efficacy of the IP as compared to the most likely comparator that might be used in the Phase 3 studies. If the comparator arm may not be obvious, the sponsor can include two comparator arms to also help determine which comparator is most relevant. The cost of such a study is outweighed by the important information gained. Note that the Phase 3 program should not begin unless the comparator group is clearly established based on data from the Phase 2 studies. To do otherwise means the information is likely insufficient on comparator arm. This enhances the chances of Phase 3 study failure.

The same can be said for all the issues raised in the Phase 3 needs described above. Hard data is needed in the Phase 2 program to provide information to make needed decisions for the Phase 3 program so you are not simply making decisions based on opinions.

Conclusion: This talk will identify the needed Phase 2 clinical trial design criteria that should be considered before designing the Phase 2 clinical trials that need to be run. Following these guidelines will provide the necessary information from Phase 2 studies to include in your Phase 3 registration studies, and the confidence that your Phase 3 studies are based on the best information available to provide more confidence in the conclusions to expect from your Phase 3 registration studies.

Contributors

Frank Wierckz
Clinipace Worldwide

Raquel Billiones
Clinipace Worldwide

USING QUALITATIVE DATA TO DEVELOP AND REFINE INTERVENTIONS

KELLEE MILLER

JAEB CENTER FOR HEALTH RESEARCH

Introduction/Objective:

Management of type 1 diabetes (T1D) in youth is challenging, with the greatest demands placed upon families of children with T1D <8 years of age. Despite modern technologies and therapies, care of these youngest children remains difficult due to their unpredictable behaviors and eating patterns, inability to articulate symptoms of out-of-range blood glucoses, and frequent intercurrent illnesses.

We investigated challenges to diabetes management in youth <8 years old to identify barriers to intensive insulin therapy and the use of advanced diabetes technologies using both web-based surveys and semi-structured interviews.

Methods:

Semi-structured qualitative interviews were conducted with parents (85% mothers) of 79 youth aged 1 to <8 y/o with T1D for ?6 months (mean age 5.2±1.5y, T1D duration 2.4±1.3y, 77% white, A1c 7.9±0.9%, 66% pump-treated). Interview transcripts were coded and evaluated using content analysis to derive central themes. Parents also completed surveys on healthcare needs.

Results:

The combination of survey responses from parents with qualitative data from parent interviews revealed modifiable and non-modifiable barriers to optimizing glycemic control and technology use and identified and confirmed core psychosocial constructs. The data from the qualitative interviews both supported existing hypothesis surrounding parental burden but also revealed less commonly known perceived barriers to technology use.

Conclusions:

Qualitative data is useful and informative for formulating and tailoring behavioral interventions. We were able to successfully design an intervention based on qualitative responses from parents of youth with T1D and incorporate into a randomized clinical trial. We have used these data to develop an intervention to determine if real-time continuous glucose monitoring (CGM), and the combination of real-time CGM with family behavioral interventions can improve glycemic control and parental quality of life in young children <8 years old with T1D.

Contributors

Linda DiMeglio
Indiana University

ACHIEVING PROJECTED DATABASE LOCK IN A MULTI-SITE STUDY - A REPORT FROM A NIDA CTN STUDY

KEATON SOMERVILLE

THE EMMES CORPORATION

Introduction: Closing a multi-site study requires extensive and careful preparation and accommodation to meet contractual project deadlines. Failure to meet this critical project milestone can become costly as it delays conducting data analysis, publishing the Final Study Report, study publications, and may result in resubmitting the protocol for IRB continuing review.

Objective: The Emmes Corporation's Data and Statistics Center (DSC) and Clinical Coordinating Center (CCC) support multi-site substance use treatment studies for the National Institute on Drug Abuse (NIDA) Clinical Trials Network (CTN). Through experiences gained across numerous clinical trials, the centers have developed strategies to achieve a smooth process during the final stages of a study. This involves routine collaborative efforts to assess, plan, and adapt study procedures to achieve the projected database lock date for the trials.

Methods: The DSC and the CCC staff begin study closeout endeavors with the identification and coordination of required database lock activities six months prior to the projected database lock date, including an assessment of potential obstacles that could impede the timely closing of the study. Two key database lock challenges involve identifying and then resolving data quality discrepancies. Utilizing the project-specific standardized internal tracking study close-out checklist, the team evaluates database lock tasks and makes appropriate modifications to fit the needs of the study, specifically, outlining and prioritizing data quality tasks (e.g., tracking the status of manual data changes and the review of protocol monitoring discrepancies). In addition to the internal tracking checklist, the CCC develops a protocol-specific closeout checklist for study sites to aid research staff in tracking protocol deviation and adverse event submissions, as well as, reviewing missing values, forms and systems checks. The DSC follows up with research staff to ensure they are updating data accordingly. As a method to review and resolve complex data quality items for some trials, the DSC and CCC may conduct internal weekly data quality meetings amongst the protocol specialist, data manager and biostatisticians, in addition to their weekly protocol team meeting. The DSC often provides additional support to research staff to accommodate sites with depleting staff resources near the end of study. Three months prior to the projected database lock date, the DSC may begin scheduling weekly tele-conferences with research staff to efficiently address and resolve outstanding site-specific data queries in real time.

Conclusion: Many tools are necessary to achieve a database lock study milestone, including internal and site checklists and scheduling additional internal and site-specific meetings as deemed necessary. Due to the collaborative work of the DSC and CCC to proactively plan and adapt to the needs of the study and research staff, the CTN studies are better equipped to meet their projected database lock dates.

<u>Contributors</u>

Dagmar Salazar
The Emmes Corporation

Dikla Shmueli-Blumberg
The Emmes Corporation

Jeanine May
The Emmes Corporation

Patrice Yohannes
The Emmes Corporation

Jacqueline King
The Emmes Corporation

Beth Jeffries

The Emmes Corporation

ADVERSE EVENTS IN 103 RANDOMIZED TRIALS ON 1ST LINE GLAUCOMA EYE DROPS

JIAJUN WEN

JOHNS HOPKINS BLOOMBERG SCHOOL OF PUBLIC HEALTH

Background

Medical treatment (i.e., topical eye drops) is considered as the first-line therapy in published guidelines for managing primary open-angle glaucoma (POAG). Building upon a recent systematic review and network meta-analysis (NMA) on the comparative effectiveness of 4 classes (14 types) of eye drops for POAG, we attempted to analyze the adverse events reported in the included randomized controlled trials (RCTs) to understand the comparative safety of the treatment options.

Methods

We extracted verbatim descriptions of adverse events from 104 published RCTs included in our NMA. Because inconsistent terminologies were used to describe the same adverse event, a glaucoma specialist recoded adverse events across all trials for analysis. For example, 'ocular itching', 'eye itching', 'itching eye', 'ocular pruritus', 'pruritus', and 'eye pruritus', all describing one symptom, were recoded to 'eye itching'. To display the results, we further grouped adverse events into sign/symptom categories. For example, symptoms such as 'eye irritation and discomfort', 'foreign body sensation', 'eye watering', and several others were grouped into the category 'symptom: ocular surface'. To enable analysis of the most important adverse events common across drugs, we plotted the frequency of reporting of each adverse event in a heat map (Figure).

Results

We extracted 419 distinct descriptions of adverse events and recoded them into 152 unique ones across 17 sign/symptom categories from 103 trials. The heat map shows that different classes of eye drops were associated with the reporting of different types of adverse events. "Symptom: Ocular Surface", "Symptom/ Sign: Eye Redness", "Sign: Eye exam finding - non-vision-threatening" were reported most often among all trials across all classes of drugs. We were not able to conduct any NMA using these safety data because of inconsistent, incomplete, and under-reporting of adverse events in trials.

Discussion

Evidence users, especially patients need information about both benefits and safety of interventions to make informed healthcare decisions. Describing the same adverse event inconsistently across trials makes it challenging for systematic reviewers to synthesize them across trials, limiting their values for decision making. Standard medical terminology such as MedDRA could be used to facilitate the documentation and analysis of safety data. In addition, based on our heat map, we propose that a core safety outcome set could be developed for future POAG drug trials in which the adverse events in the core set can be collected systematically to improve the collection and reporting of adverse events.

Contributors

Amanda Bicket
Johns Hopkins Wilmer Eye Institute

Kay Dickersin
Johns Hopkins Bloomberg School of Public Health, US Cochrane Center

Tianjing Li
Johns Hopkins Bloomberg School of Public Health, US Cochrane Center

ALL HANDS ON DECK FOR ELIGIBILITY CHECKS

STEPHANIE WELLFORD

RHO, INC.

In order to take part in a clinical trial, participants must meet specific eligibility criteria. Eligibility criteria protect participants' safety and ensure participants in the clinical trial represent the population identified for analysis. To maintain the scientific integrity of the trial, it is critical for clinical sites to confirm participant eligibility prior to enrollment and to ensure ongoing eligibility throughout the trial. While confirming participant eligibility is ultimately the clinical site investigator's responsibility, the statistical and clinical coordinating center (SACCC) can assist by providing tools that facilitate eligibility assessment and minimize the chance for enrollment errors. Our SACCC team at Rho is responsible for monitoring study databases and site documentation to confirm participants enrolled in the clinical trial meet eligibility criteria. The project managers, data managers, and statisticians work together to develop multiple tools and processes that help the clinical site teams with real-time eligibility confirmation.

The first eligibility checks are in place to confirm inclusion and exclusion criteria at enrollment and randomization. These criteria are individually captured on CRFs and are entered into the EDC system in real-time. A separate trial-specific eligibility CRF includes each inclusion and exclusion criterion. Clinical site teams enter this eligibility CRF into the EDC system, which is programmed with edit checks to confirm that the responses match those on the CRFs containing the clinical data. If a participant does not meet eligibility criteria or if discrepancies are present, the system will generate a real-time query notifying the site the participant is ineligible or the CRFs are discrepant. The site investigator reviews and signs the CRF as a final check to confirm eligibility. Only after the eligibility CRF is complete and eligibility is confirmed will the clinical site team be able to access CRFs needed for the remainder of the visit procedures.

As an additional eligibility confirmation, checklists, which include all expected CRFs, procedures, and biological specimens, are used at each visit as a tool to guide the order of visit activities. Those activities required to confirm initial and ongoing eligibility are listed first. These checklists minimize the risk of clinical sites performing study activities on participants who are later found to be ineligible. For study procedures that require specific criteria to be confirmed prior to performance, additional processes are developed. For example, the EDC is programmed such that the CRFs needed to complete a procedure and collect specimens are not available until eligibility is confirmed.

While creating and implementing these eligibility checks requires coordination across the entire study team, these tools assist clinical sites in making real-time eligibility determinations and allow the sponsor and CRO to monitor eligibility, ultimately meeting the overall goal of avoiding enrollment of ineligible participants.

Contributors

Amanda Farrell

Rho, Inc.

AN INITIATIVE USING INFORMATICS TO FACILITATE CLINICAL RESEARCH PLANNING AND RECRUITMENT IN THE VA HEALTH CARE SYSTEM

KANDI VELARDE

U.S. DEPARTMENT OF VETERANS AFFAIRS HEALTH CARE SYSTEM

Background: Randomized clinical trials are the gold standard for evaluating healthcare interventions and, more generally, by adding to the medical knowledge related to the treatment, diagnosis and prevention of diseases and conditions. Recent literature continues to identify health informatics methods that can help improve study efficiency throughout the life cycle of a clinical trial. Electronic medical record (EMR) data provides a mechanism to facilitate clinical trial research during the study planning and execution phases, and ultimately, can be utilized to enhance recruitment. The Veterans Health Administration (VHA) has a strong history of clinical and epidemiological research with over four decades of data collected from Veterans it has served nationwide. The VA Informatics and Computing Infrastructure (VINCI) provides VA research investigators with a nationwide view of high value VA patient data. The Department of Veterans Affairs (VA) Cooperative Studies Program (CSP) is a clinical research infrastructure embedded within the nation's largest integrated health care system. The VA Network of Dedicated Enrollment Sites (NODES) is a consortium of nine (VA's) sites intended to provide systematic site-level solutions to issues that arise during the conduct of VA CSP clinical research. This abstract describes the collaboration initiated by the SLC NODES that brings informatics and clinical trials together to enhance study planning and recruitment within the VHA.

Methods: The Salt Lake City (SLC) VA Medical Center physically houses both VINCI and a NODE site and the proximity of these two groups prompted a natural collaboration on both a local and national level. One of the functions of the SLC NODES is to enhance recruitment and promote the success of CSP projects. VINCI supports these efforts by providing VA researchers access to potential population pools. VINCI can provide 1) feasibility data during study planning, and 2) active patient lists during recruitment. The process for CSP study teams to utilize these services involves regulatory documentation, development of queries, multiple meetings, revisions to the initial data request, and ongoing communications with several key individuals including the requesting research team, study statisticians, and VINCI data managers.

Results: The early efforts of SLC NODES and VINCI aimed to provide patient lists exclusively to the SLC CSP study teams for the following purposes: 1) increasing recruitment for trials that are struggling to meet their respective enrollment goals, and 2) decreasing the time required by study coordinators to complete chart review activities. This effort was expanded to include multiple CSP sites and studies. To date, SLC NODES has facilitated the delivery of these VINCI services to nine active CSP studies.

Conclusion: The ability of clinical trial study teams to successfully plan and execute their respective trials is contingent upon their proficiency in obtaining data that will help them efficiently and effectively recruit and enroll eligible participants. This collaboration demonstrates that the utilization of a model that partners two distinct entities, with similar objectives, was effective in the provision of feasibility and patient lists to clinical trial study teams and facilitation of clinical trial research within a large, integrated healthcare system.

<u>Contributors</u>

Marcus Johnson
Durham VA Health Care System

Jennifer Romesser
VA Salt Lake Health Care System

Daniel Clegg
VA Salt Lake Health Care System

Jeffrey Scehnet
VA Salt Lake Health Care System

Olga Effimova
VA Salt Lake Health Care System

Grant Huang

Cooperative Studies Program Central Office, VA Office of Research and Development

CALCULATING ANTIRETROVIRAL DRUG RESISTANCE: AN INNOVATIVE OPEN-SOURCE TOOL

ADINA STOICA

FRONTIER SCIENCE FOUNDATION

Background. HIV-1 resistance interpretations are most commonly algorithmic, using rules based on mutations, as with the Stanford HIV Drug Resistance Database (HIVDB) or the France Recherche Nord & Sud Sida-HIV Hépatites (ANRS). Most computer implementations of algorithms have the rules programmed in, perhaps using a database to allow some flexibility in calculations. The Algorithm Specification Interface (ASI) approach, by contrast, aims to decouple scientific rules from computer code. ASI involves writing computer code that instantly compiles and executes any properly stated set of scientific rules. This approach, pioneered by the HIVDB group, has now been flexibly and robustly implemented by Frontier Science Foundation (a statistical and data management institute specializing in clinical trials and observational research) in collaboration with the HIVDB group. The result is a new body of open-source software, released on 01-Dec-2017, available to all resistance researchers.

Methods. The ASI approach has two components: (1) an English-like and flexible but structured grammar that allows scientific rules and elements used in analysis (genes, drugs, etc.) to be easily but fully described in an XML file; and (2) program code, written in Java and drawing on the open-source SableCC library (a grammar compiler used to create a skeleton of strictly-typed abstract syntax trees and tree walkers), that receives a list of mutations, parses the XML file, evaluates the nodes of the syntax tree by which to evaluate resistance for each specified drug, and then works through the trees, reaching resistance interpretations for each drug/sequence combination. The code was extensively tested, including by the HIVDB group.

Results. The software has been posted on GitHub, under the Apache 2.0 license, at https://github.com/FrontierScience/asi_interpreter. Before becoming open-source, it was used by HIVDB, by the New York Department of Health, and by Frontier Science Foundation for NIH-sponsored networks: AIDS Clinical Trials Group, International Maternal Pediatric Adolescent AIDS Clinical Trials, HIV Prevention Trials Network, and Pediatric HIV/AIDS Cohort Study. The code is a library, not a stand-alone program. A simple wrapper is provided, but many users will want to write code that accesses input data, invokes the library, and organizes the resistance results according to the user's needs; doing so is not hard but requires programming skill.

Conclusions. The ASI library offers many advantages: flexibility (an algorithm can be customized or revised quickly and without programming skill); transparency (all rules are stated in the XML file), reproducibility (results are the same no matter where ASI is run); simplicity (no database or data files beyond the algorithm XML are required); portability (implemented in Java); ease of experimentation (one can quickly try algorithmic variations and see the results); robustness (removing the risk of programming errors in version updates), and user control (any version of any algorithm can be run, whereas using a web system, even HIVDB, usually restricts one to using a predetermined version). Open-source release provides new possibilities for extension to organisms beyond HIV-1, to enhancements scientists may desire, and to broader use.

<u>Contributors</u>

Robert Shafer
Stanford University

Stephen Hart
Frontier Science Foundation

CHALLENGES AND LESSONS LEARNED FROM ENGAGING THE CLINICALTRIALS.GOV PROTOCOL REGISTRATION AND RESULTS SYSTEM (PRS): AN NRG ONCOLOGY EXPERIENCE

WENDY SEIFERHELD

NRG ONCOLOGY

The U.S. Public Law 110-85 (Food and Drug Administration Amendments Act of 2007), Title VIII, Section 801 mandated the expansion of the clinical trials data bank, ClinicalTrials.gov, to include information on the results of clinical trials. Specifically, responsible parties are expected to enter summary information on study participants and study outcomes for certain trials, within one year of the Primary Completion Date (date on which the data for the primary endpoint collection is completed), and creating penalties for noncompliance. NRG Oncology is a non-profit research organization running National Cancer Institute (NCI)-sponsored multi-center clinical trials, with the primary goal of improving the lives of cancer patients by conducting practice-changing clinical and translational research and reporting their results. This abstract focuses on some challenges and lessons learned from entering results of NRG trials into ClinicalTrials.gov PRS.

From 2012 to 2017, NRG SDMC has entered the results of more than 75 clinical trials into ClinicalTrials.gov PRS. These trials range from phase I (in combination I/II and I/III trials) to phase III. The vast majority of these trials were written/activated prior to the implementation of the required results reporting. Rather than having each trial statistician enter results into ClinicalTrials.gov PRS, NRG identified selected individuals to submit results, in order to allow for focused, consistent, and efficient efforts to meet required compliance.

Challenges along the way have involved determining best methods for collating and entering data into the PRS system (technically and logistically), integrating ClinicalTrials.gov-related tasks into established processes seamlessly to minimize additional work, and ensuring that analysis and reporting plans accommodate ClinicalTrials.gov requirements. Other issues have been intrinsic in the reporting requirement itself, due to across the board, fixed entry requirements, for which there are not always straightforward /satisfactory solutions. This is particularly challenging in trials that did not complete planned accrual. Addressing these challenges has been problematic and time consuming.

Specific examples of and approaches to challenges encountered by NRG will be presented. A primary goal of this presentation is to stimulate discussion. Shared challenges exist among the large number of organizations conducting clinical trials and an exchange of ideas and approaches may be of value.

Contributors

Kathryn Winter
NRG Oncology

CLASSIFYING NO IMPROVEMENT IN VISUAL ACUITY IN CHILDREN WITH AMBLYOPIA USING SIMULATIONS

RUI WU

JAEB CENTER FOR HEALTH RESEARCH

Background: In clinical trials involving novel treatments of amblyopia in children, a pre-randomization run-in phase is often used to ensure that the amblyopic-eye visual acuity (VA) has stopped improving with current treatment. In a previous amblyopia trial, children randomly assigned to continue spectacle wear alone improved by a mean of 0.5 line during follow up despite having VA classified as not improving after completing a spectacle run-in phase. This suggests that our current classification criteria may not be ideal for judging no improvement in VA. We used simulations to evaluate the performance of different rules for prospective classification of no improvement in VA and make recommendations for future clinical trials.

Methods: Monte Carlo simulation was used to generate a true amblyopic-eye VA (20/40 to 20/400) and 3 observed VA (initial test, retest, and 2nd retest) at each of 3 clinic visits for 10,000 hypothetical subjects. The observed VA was calculated by adding random measurement error (based on established test-retest error) to the true VA. Eighteen rules for classifying no improvement based on the observed VA were developed by: 1) adding up to two retests at each visit; 2) adding a third visit; and/or 3) using the best VA or mean VA of observed test and retest measurements to calculate changes. These rules were evaluated where true VA did not change over time or improved by a constant or decreasing amount between successive visits. For each rule, the sensitivity and specificity were calculated.

Results: The typical practice of comparing a single test of VA at 2 consecutive visits (for subjects with 1 line true improvement, same below if not otherwise specified) had a sensitivity of 70% and a specificity of 72%. Adding both a same-day retest (at each visit) and a subsequent visit to confirm "no improvement" and comparing best VA tested at each visit improved the performance (sensitivity: 80%, specificity: 80%), as did comparing the average of test and retest VA at each visit (sensitivity: 73%, specificity: 95%). Adding 2 same-day retests (at each visit) and a subsequent visit to confirm "no improvement" and comparing average of 3 tests at each visit also improved performance (sensitivity 86%, specificity: 93%). Generally, rules that required 3 successive visits to demonstrate "no improvement" had higher specificities (24% to 62% increase) but lower sensitivities (14% to 61% decrease) compared with rules that required only 2 visits.

Conclusions: Adding 1 or 2 same-day retest(s) at each visit and a 3rd visit to confirm "no improvement" improved accuracy in classification of "no improvement" in VA among children with amblyopia. If the intent is to identify as many subjects as possible who are likely not improving to be enrolled in a randomized trial, rules that require only 2 successive visits should be considered. If ensuring that a participant's VA has truly stopped improving before randomization is important, rules requiring 3 visits should be considered. Simulations were helpful for quantifying the expected effect of retest(s) and/or additional visits on sensitivity and specificity of eligibility criteria.

Contributors

Ruth Manny
University of Houston College of Optometry

Michele Melia
Jaeb Center for Health Research

Elizabeth Lazar
Jaeb Center for Health Research

Jonathan Holmes
Mayo Clinic

Eileen Birch
University of Texas Southwestern

Susan Cotter

Southern California College of Optometry Marshall B. Ketchum University

Raymond Kraker

Jaeb Center for Health Research

David Wallace

Indiana University

DESIGN OF NON-INFERIORITY RANDOMIZED TRIALS USING THE DIFFERENCE IN RESTRICTED MEAN SURVIVAL TIMES

ISABELLE WEIR

BOSTON UNIVERSITY SCHOOL OF PUBLIC HEALTH

Background: Non-inferiority trials with time-to-event outcomes are becoming increasingly common. Designing non-inferiority trials is challenging, in particular they require very large sample sizes. We hypothesized that the difference in Restricted Mean Survival Times (RMST), an alternative to the Hazard Ratio (HR), could lead to smaller required sample sizes.

Methods: We aimed to re-design a sample of non-inferiority trials and compare the required sample sizes based on the difference in RMST and the HR. We systematically searched MEDLINE for non-inferiority trials published between 2013 and 2016 in 7 major journals. Based on the protocol and article of each trial, we determined the clinically relevant time horizon of interest. We reconstructed individual patient data for the primary outcome. We tested the proportional hazard assumption and also for non-inferiority using the difference in RMST and the HR. We calculated the margin for the difference in RMST that corresponded to the HR margin. Using both measures, we determined the required sample size with the type I error risk and power from the original trial design.

Results: We reconstructed data for 35 trials. We found evidence of non-proportional hazards in 5 (14%) trials. The HR and difference in RMST were consistent regarding non-inferiority testing, except in one trial where the difference in RMST led to evidence of non-inferiority while the HR did not. The median HR margin was 1.49 (range 1.15 to 2.85). The median of the corresponding margins for the difference in RMST was -21 days (Q1-Q3 -36 to -8) for a median time horizon of 2.0 years (range 0.5 to 9.0 years). The required sample size according to the difference in RMST was smaller in 71% of trials, with a median relative decrease of 8.5% (Q1-Q3 0.4%- 38.0%). Across the 35 trials, about 25,000 participants would be spared from enrollment using the difference in RMST as opposed to the HR as the measure of treatment effect.

Conclusions and Relevance: The HR margins may seem large and lenient but translate to relatively small differences in RMST. The assumption of proportional hazards is often violated but the use of the hazard ratio persists. The difference in RMST offers a meaningful and clinical interpretation and can result in considerable reductions in sample size when used for trial design. This alternative measure should be considered more widely in non-inferiority trials with time-to-event outcomes.

Contributors

Ludovic Trinquart
Boston University School of Public Health

DESIGN OF THE TEAMMATE TRIAL FOR CHILDREN WITH HEART TRANSPLANT AND DEVELOPMENT OF A NOVEL EFFICIENT ENDPOINT

LYNN SLEEPER

BOSTON CHILDREN'S HOSPITAL/HARVARD MEDICAL SCHOOL

BACKGROUND: Clinical research in pediatric heart transplant has historically been conducted to assess short-term outcomes at single centers. Efficient approaches are needed to test hypotheses related to long-term outcome.

Median survival after pediatric heart transplantation is only 15 years in the current era, due to the occurrence of late complications after heart transplant, most of which stem from the medications used to suppress the immune system in order to prevent graft rejection. While graft survival has improved significantly with the current standard of care, tacrolimus (TAC) and mycophenolate mofetil (MMF), most of the improvement has come from a reduction in early mortality. Late mortality is driven by 6 major adverse transplant events (MATE): acute cellular rejection (ACR), antibody-mediated rejection, coronary artery vasculopathy (CAV), post-transplant lymphoproliferative disorder, infection, and chronic kidney disease (CKD). In recent years, everolimus (EVL), a proliferation signal inhibitor, in combination with low-dose tacrolimus (LDTAC), has emerged as a potential treatment that may improve longer-term survival by reducing the risk of CAV, CKD and ACR, as well as cytomegalovirus infection.

THE CHALLENGE: Evaluation of an immunosuppressive regimen to reduce the rate of long-term complications of transplant is hindered by a large sample size requirement and lengthy follow-up in order to achieve sufficient power to detect treatment differences. The MATE Score was developed by a network of pediatric heart transplant investigators; it was validated against long-term registry outcomes and then approved by the FDA, paving the way for implementation of an efficient trial design.

DESIGN: The FDA-regulated, multicenter randomized TEAMMATE Trial network infrastructure was built in 2017 and launched enrollment in 2018. The trial is led by a Clinical Coordinating Center at Stanford University and a Data Coordinating Center at Boston Children's Hospital. Innovations include the MATE score and the use of a central IRB for over 20 sites to facilitate efficient trial start-up. The target of 210 patients will be randomized to either the standard TAC/MMF or EVL/LDTAC at 6 months post-transplant. The primary efficacy and safety endpoints are two forms of the MATE score. Follow-up is 2.5 years per subject. Central adjudication of clinical MATE events and a core laboratory for angiogram interpretation are in place to minimize variance.

CONCLUSIONS: Leveraging the cumulative burden of multiple complications into a single, continuous validated endpoint enables the TEAMMATE trial to test a hypothesis regarding long-term outcome with relatively short follow-up. This talk will highlight trial endpoint development, statistical power considerations, and the logistics of efficiently launching a complex multicenter trial with multiple stakeholders. We anticipate that the TEAMMATE trial will result in FDA approval of the first immunosuppression regimen for pediatric heart transplantation and that the

trial infrastructure will facilitate the development of a sustainable clinical research network for the pediatric heart transplant field.

Contributors

Kevin Daly
Boston Children's Hospital/Harvard Medical School

Joseph Rossano
Children's Hospital of Philadelphia

Manisha Desai
Stanford University School of Medicine

Scott Auerbach
Children's Hospital Colorado/University of Colorado Denver

Matthew Bock
Loma Linda University

Chesney Castleberry

Washington University School of Medicine

Matthew Fenton

Great Ormond Street Hospital

Seth Hollander

Lucile Packard Children's Hospital/Stanford University School of Medicine

Ashwin Lal

Primary Children's Hospital/University of Utah

Elfriede Pahl

Ann & Robert H. Lurie Children's Hospital of Chicago

Christopher Almond

Lucile Packard Children's Hospital/Stanford University School of Medicine

EVALUATION OF DIABETIC RETINOPATHY USING THE ETDRS SEVERITY SCALE – IS THERE A GOLD STANDARD?

RALPH TRANE

UNIVERSITY OF WISCONSIN--MADISON

Purpose: Automated image grading is gaining popularity and accuracy is currently being evaluated against evaluation data from experienced readers. Therefore, it is important that the reproducibility of the primary outcomes used as the golden standard is as high as possible. We compared reproducibility of three methods of evaluation workflow - dual read with adjudication, majority rule and single reader scenario.

Methods: To compare the reproducibility, color fundus photographs (CFP) from 129 eyes were evaluated by experienced certified readers. The severity of diabetic retinopathy (DR) was documented on a 12 step scale as the primary outcome. For each eye, the primary outcome was obtained in three different ways: (1) dual read with adjudication, where two independent readers evaluate images with adjudication by a third senior reader in cases of discrepancy (2) majority rule, where three or more readers document the primary outcome and the majority grade is identified as the final score (3) single reader system, where each image is evaluated by an individual reader and monitored using inter and intra-grader agreement. Each method was repeated in its entirety. For example two readers evaluated an image with an adjudicator's review and two additional readers evaluated the same image with an alternate adjudicator's review. Both percentage agreements, kappa statistics, and data loss due to ungradability within each method were compared.

Results: Agreement rates and weighted kappa are shown in the table below. Weighted kappa was > 0.8 for all three methods and highest with the majority rule system. Ungradable rate was highest with the majority rule system at 11% compared to 1.5% with the other two methods.

Conclusion: Our results suggest that all three methods of evaluating primary outcome for DR severity are reproducible. Dual read with adjudication provides high agreement rates with least compromise on data loss due to ungradable images and should be considered the gold standard method for evaluating DR.

Contributors

Amitha Domalpally
Fundus Photography Reading Center, University of Wisconsin--Madison

James Reimers
Fundus Photography Reading Center, University of Wisconsin--Madison

Barbara Blodi
Fundus Photography Reading Center, University of Wisconsin--Madison

FEASIBILITY OF A SEVEN-DAY THERAPEUTIC WASHOUT IN EARLY STAGE PARKINSON'S DISEASE

MALLORY HACKER

VANDERBILT UNIVERSITY MEDICAL CENTER

Objective: To report the feasibility of conducting a seven-day therapeutic washout in a clinical trial evaluating subjects with early stage Parkinson's disease (PD).

Background: Parkinson's disease is a relentlessly progressive neurodegenerative disease that affects more than one million Americans. Since there is currently no biomarker that objectively monitors disease progression, clinical trials to explore potentially disease-modifying therapies must utilize study designs suited to demonstrate a lasting effect on the disease course.

Methods: Vanderbilt University completed a prospective, randomized (1:1), single-blind pilot clinical trial assessing the safety and tolerability of deep brain stimulation (DBS) in early stage Parkinson's disease (IDE G050016, NCT 00282152, IRB 040797). To evaluate the progression of underlying motor symptoms without robust interference of symptomatic therapies, a seven-day washout of all Parkinson's disease medications and stimulation (if applicable) was implemented at baseline and every six months during the two-year trial (five total). After the trial concluded, 27 of 29 subjects completed a survey that collected information regarding their experiences in the study.

Results: Subjects completed all of the 147 week-long washout periods without complications or early termination, and none of the subjects experienced a dopaminergic withdrawal syndrome. After the trial concluded, subjects were asked to select all that applied for the following question: "Which parts of the study, if any, were most burdensome?" Responses included the financial commitment (8/27, 30%), the washout periods (7/27, 26%), and the time commitment (4/27, 15%), although 10 subjects (37%) indicated that they did not find any part of this study to be burdensome. One respondent stated, "A week out of my schedule is expensive," and another considering dropping out due to "problems with time away from work." Conversely, two subjects indicated that the washout periods were the most interesting part of the study, and a majority of subjects (14/27, 52%) indicated that the daily clinical evaluations during the washout were among the most enjoyable parts of the study.

Conclusions: This pilot trial demonstrated the feasibility of conducting week-long therapeutic washouts in patients with early stage Parkinson's disease. None of the washouts were terminated early due to safety or tolerability reasons. While subjects with early stage Parkinson's disease patients tolerated the therapeutic washouts, some subjects found them to be burdensome due to the required time commitment. Therefore, fewer washout assessments will be implemented in the future pivotal, phase III safety and efficacy clinical trial of DBS in early stage Parkinson's disease (IDE G050016).

Contributors

Anna Molinari
Vanderbilt University Medical Center

David Charles
Vanderbilt University Medical Center

IMBALANCED

THEVAA CHANDERENG

UNIVERSITY OF WISCONSIN MADISON

Randomization is a common technique used in clinical trials to eliminate potential bias and confounders in a patient population. Randomization eliminates a systematic difference between subjects in treatment groups inducing approximate balance with respect to covariates, both observed and unobserved. Equal allocation to treatments groups is the standard due to its high efficiency in many cases. Statistical efficiencies are directly related to statistical power. Thus, increasing statistical efficiencies improves the likelihood of correctly rejecting the null hypothesis when the alternative is true. However, in certain scenarios, unequal allocation can improve efficiency. In superiority trials with more than two groups, the optimal randomization is not always a balanced randomization. Non-inferiority trials are clinical trials designed to establish a new treatment that is not much worse than the current treatment or control. A margin ? is introduced to allow a small loss in efficacy in the new treatment compared to the control treatment. The margin of ? could be additive or multiplicative and it is important in obtaining the efficient randomization ratio. In non-inferiority trials, additive margin with equal variance is the only instance with balanced randomization. The Cardiovascular Patient Outcomes Research Team (CPORT), which is a multi-center trial conducted to compare fatality rate between presence of on-site cardiac surgery in performing angioplasty, serves as a motivation to this research. Patients were randomized 1:3 (with and without onsite cardiac surgery) due to several constraints, therefore reducing the statistical efficiency and statistical power for the same sample size (N) compared to the optimal randomization ratio. This paper describes each case and computes the optimal randomization ratio and efficiency gains compared to a balanced randomization.

Contributors

Rick Chappell
University of Wisconsin Madison

Lee McDaniel
Louisiana State University

IMPLEMENTING ECRFS IN AN EMERGENCY DEPARTMENT ENVIRONMENT: CHALLENGES, ENHANCEMENTS, AND LESSONS LEARNED

LAUREN YESKO

THE EMMES CORPORATION

Introduction

The Emmes Corporation, in its role as the Data and Statistics (DSC) center for studies conducted by The National Drug Abuse Treatment Clinical Trials Network (CTN) of the National Institute on Drug Abuse (NIDA), implemented data management systems and additional electronic administrative tools in a clinical trial in high volume Emergency Department (ED) settings. Due to the fast-paced nature of the ED environment, approaches used in a typical fully electronic clinical trial needed to be re-evaluated to manage site staff, participants, and data collection.

Background

Conducting a clinical trial in an ED setting requires additional considerations beyond a traditional clinical trial site to accurately and completely collect data and maintain high data quality while meeting the needs of the protocol. Creating an efficient data system can be obtained by adding electronic administrative enhancements outside a typical eCRF environment and performing targeted, clear data cleaning processes.

Methods and Results

In this presentation, we will review system enhancements that were developed to meet the administrative needs of study site staff, including screening logs, enrollment reports, tools for scheduling and identifying upcoming and overdue follow-up visits, appointment cards, reminder letters, unable-to-contact letters, fax cover sheets, thank you cards, and database tools for entering data related to area providers and provider surveys; we will also review tools for easily identifying and addressing data quality issues. Additionally, we will present items that led to successful support of the trial, as well as areas for potential modification for similar future studies. Finally, we will review the unique challenges and solutions in cleaning data in this atypical setting, including engaging and developing relationships with research staff within the ED.

Conclusions

To successfully support a large scale, multi-site, fully electronic trial in an ED setting required flexible and innovative solutions. While most tools created prior to implementation met the needs of the trial, during study conduct we identified alternate options that have potential for improved trial management. An additional lesson learned is the utility of being flexible and willing to accept varying levels of clean data.

Contributors

Casey Nelson

The Emmes Corporation

Dagmar Salazar
The Emmes Corporation

Mitra Lewis
The Emmes Corporation

Aimee Wahle
The Emmes Corporation

Mateusz Makowski
The Emmes Corporation

Felipe Munoz

The Emmes Corporation

Eve Jelstrom

The Emmes Corporation

INCORPORATING AGING MEASURES IN A LONG-TERM MULTICENTER OBSERVATIONAL STUDY: TRAINING AND CERTIFICATION PROCEDURES

VICTORIA TRAPANI

GEORGE WASHINGTON UNIVERSITY

Background: Participants with type 1 diabetes (T1D) enrolled in the Diabetes Control and Complications Trial (DCCT, 1982-1993) and subsequently in the Epidemiology of Diabetes Interventions and Complications (EDIC, 1994-present) study have been followed for over 33 years and extensively phenotyped using consistent and validated methods. The DCCT/EDIC study has demonstrated the long-term benefits of intensive vs. conventional diabetes therapy on reducing the risk of micro- and macrovascular complications and mortality in T1D. The next phase of the study (2017-2022) will focus on the interaction between and the effects of aging and long-duration diabetes on physical and cognitive function. Incorporating these new measures into the already multifaceted EDIC annual visit poses challenges related to participants and examiners. Well-designed training and certification processes are crucial for ensuring uniform administration and accuracy.

Methods: An expected 1,200 EDIC participants (mean age 58 years, mean T1D duration 35±5 years) at 27 EDIC clinical centers will undergo physical and cognitive function testing. Cognitive testing will be completed twice (~3 years apart) for each participant and includes measures of learning, memory, executive function, information-processing speed, and mood state administered via iPad, interviewer-administered paper-and-pencil tests, and self-administered questionnaires. The physical function testing will be completed once with each participant and includes a series of objective performance tests and self-administered questionnaires. During a 6-month period, various written, video, and e-learning methods were used to supplement two in-person training sessions to certify 37 study personnel, most of whom were registered nurses, to ensure proficiency in all aspects of test administration. An EDIC-specific study operations manual and a series of online videos augmented in-person demonstrations. Trainees were required to practice the physical function tests on at least one volunteer before administering the entire protocol to an expert certifier in-person. A written quiz was given in which a score of 90% or higher was required. For the cognitive protocol, trainees reviewed an NIH Toolbox e-learning module and videos that demonstrated the interviewer-administered tests. Practice testing of the full set of cognition measures was required on three volunteers. Certification consisted of in-person administration of the entire protocol to 1 of 4 expert examiners. Electronic submission of the completed test packet was required to ensure accurate data transfer procedures.

Results: The comprehensive training and certification process required on average 12 hours of training for cognition and 3 hours for physical function, but ultimately resulted in the certification of 89% and 100% of study personnel, respectively. Certification for the cognitive protocol posed more of a challenge due to trainees' lack of prior cognitive testing experience. Additional challenges included the need for sufficient self-study time; identification of practice subjects and performance of practice tests; adaptation to iPad technology; and identification of experts to oversee the certification process.

Conclusions: A 6-month training and certification period allowed the majority of personnel to demonstrate proficiency in carrying out each component of the complicated and time-intensive physical and cognitive function measures. The methods described promote uniform test administration to ensure precision, quality, and accuracy of the assessments across clinical centers.

Contributors

Barbara Braffett
George Washington University

Gayle Lorenzi
University of California San Diego

Catherine Martin
University of Michigan

Jye-Yu Backlund
George Washington University

Christopher Ryan
University of Pittsburgh

Rose Gubitosi-Klug

Case Western Reserve University

Susan Hitt

University of Missouri

Alan Jacobson

Winthrop University Hospital

ONE DECADE OF IMPACT. ONE VISION AHEAD. OPTIMIZING THE POWER OF PUBLIC-PRIVATE PARTNERSHIP: LESSONS LEARNED FROM THE CLINICAL TRIALS TRANSFORMATION INITIATIVE

PAMELA TENAERTS

CLINICAL TRIALS TRANSFORMATION INITIATIVE

CTTI has over 10 years of experience in building a collaborative community. Experience and lessons learned related to the following will be shared:

1)   How to select projects for success

2)   Team dynamics and execution

3)   General themes emerging from 10 years of recommendations

4)   Measuring the impact of a collaborative organization on the clinical trials enterprise

CTTI is a public-private partnership co-founded in 2007 by Duke University and the FDA to develop and drive adoption of practices that will increase the quality and efficiency of clinical trials. CTTI's experience in bringing more than 80 diverse organizations from across the clinical trials enterprise to the table to address specific issues affecting quality and efficiency of clinical trials will be described.

CTTI projects engage and value all stakeholders equally; project teams are comprised of volunteer stakeholders representing critical perspectives on the topic and who are potential change agents. Having conducted over 25 projects in topics ranging from patient engagement to use of mobile devices, CTTI has identified what works (and what doesn't) in creating a collaborative community. CTTI has also developed a framework for measuring impact of this work on the clinical trials enterprise.

<u>Contributors</u>

PRACTICAL METHODS FOR UTILIZATION OF BIOMARKERS TO ASSESS CANCER PROGNOSIS

RANI JAYSWAL

MARKEY CANCER CENTER, UNIVERSITY OF KENTUCKY

The role of individual biomarkers, whether during the early stages of basic science discovery or in the latter phases of clinical application after discovery and validation of hundreds of markers, continues to an important area of investigation in cancer research. In order for these biomarkers to have practical application in clinical decision making to guide patient treatment and management, it is still common to categorize and more commonly, dichotomize biomarker levels. We utilize and compare cut-off determination methods which are independent of or dependent on clinical outcomes of interest. Specifically, utilization of independent normal versus cancer biomarker levels allows evaluation of performance of area under the curve (AUC) of the receiver opera ting characteristic (ROC) curves, Youden's index, and odds ratio of the logistic regression model. Utilization of cancer biomarker levels only to determine association with clinical outcome allows evaluation of performance of hazard ratio and odds ratio from survival and logistic regression models, respectively. We then utilize predictiveness curves for survival endpoints to assess and guide the clinical implication of biomarker cut-offs based on these different methods. Specifically, the type of treatment/intervention and level of patient risks that are of interest play a role in the choice of cut-offs based on the predictiveness curves. An application to assess appropriate cut-off levels of several individual biomarkers related to redox signaling in prostate cancer is presented.

Contributors

Heidi Weiss
MARKEY CANCER CENTER, UNIVERSITY OF KENTUCKY

Li Chen
MARKEY CANCER CENTER, UNIVERSITY OF KENTUCKY

Stacey Slone
MARKEY CANCER CENTER, UNIVERSITY OF KENTUCKY

Luksana Chaiswing
Toxicology and Cancer Biology, College of Medicine, University of Kentucky

Daret StClair
Toxicology and Cancer Biology, College of Medicine, University of Kentucky

SURVEILLANCE OF SERIOUS ADVERSE EVENTS IN A MULTISITE CLINICAL TRIAL UTILIZING VA HEALTHCARE

ADMINISTRATIVE DATA

EILEEN STOCK

DEPARTMENT OF VETERANS AFFAIRS

Clinical trials provide the evidentiary basis for regulatory approvals of safe and effective treatments. The purpose of the International Conference on Harmonization (ICH) Good Clinical Practice (GCP) Guidelines is to protect the rights of participants in clinical trials and ensure the scientific validity and integrity of the data collected. Accordingly, the ICH-GCP requires all serious adverse events (SAEs) in an interventional trial be immediately reported and followed promptly by detailed, written reports. Per ICH, a SAE is defined as any untoward medical occurrence that results in death, is life-threatening, requires hospitalization or prolongs an existing hospitalization, causes persistent or significant disability or incapacity, or requires medical intervention to prevent one of these outcomes. A typical approach to the oversight and conduct of clinical trials involves designated trial team members reviewing and ensuring primary data are collected and recorded properly. This may involve site monitoring visits, meeting periodically with research coordinators to review their study records, ensuring that the reporting of SAEs is timely, and details of the event are complete. This undertaking, however, may become costly and difficult to manage for trials involving high-risk patients or with multiple sites. In turn, data integrity and patient safety could be compromised. Consequently, the purpose of this research was to develop and assess the efficiency of an active surveillance approach to monitoring SAEs using the Department of Veterans Affairs healthcare system's (VA) administrative data.

From the VA Corporate Data Warehouse (CDW), inpatient electronic medical records for participants, enrolled in a multisite cardiac clinical trial involving 16 sites, were searched and extracted. Events identified electronically from the CDW were then compared with those identified by research personnel at respective sites. Potentially missed SAEs were determined, leading to communication with the study sites regarding whether the event was reportable or not. A program was developed to make the process of searching for SAEs from VA administrative data more efficient and effective. This program is run quarterly, in conjunction with regular SAE monitoring.

With the growing trend toward global research, clinical trials are becoming larger and more complex. More participants are being enrolled and multisite trials are becoming more common. This research highlights the unique opportunity to use big data on a national scale and may inspire other groups managing clinical trials to use similar approaches with their available datasets. Monitoring of SAEs electronically, in additional to traditional monitoring, may relieve some of the increased burden being placed on study personnel. This surveillance method also has the potential to increase trial performance by helping to ensure data validity, consistency in reporting across sites and clinicians, and most importantly, ensuring safety and well-being of study participants. Early safety signal detection not only leads to better patient protection, but also has the potential to save time and costs.

Contributors

Ian Craig
Department of Veterans Affairs

Ellen DeMatt
Department of Veterans Affairs

Jeanie Kim
Department of Veterans Affairs

Kousick Biswas
Department of Veterans Affairs

TELEPHONIC INFORMED CONSENT FOR POINT OF CARE (POC) CLINICAL TRIALS

JADE FIOTTO

DEPARTMENT OF VETERANS AFFAIRS

Introduction/Objective

Point of Care Clinical Trials (POC-CTs) are pragmatic trials designed to bridge the gap between research and clinical care by embedding the study into routine clinical care. The primary goal of these studies is to make clinical research more accessible by removing, as much as possible, the burden of research for both participants and providers. To accomplish this in the Diuretic Comparison Project (DCP), participants are consented into the study through telephonic informed consent. It is necessary to determine who has better success at enrolling participants into trials, call center operators or study researchers, to create the most feasible trial.

Background

Very little information is available to instruct the performance of telephone based informed consent, so the study employed VA call center operators. These are individuals that have had previous VA Call Center experience and informed consent training, but are not study researchers. Due to high participant volume, the study team began calling participants concurrently.

Methods

Call center operators call participants on 8 hour shifts 5 days a week. Study researchers call participants when time is available outside regular study duties. Consent rates were compared over a month period of high participant influx. Metrics are kept on all calls and quality control is maintained with a robust audit program.

Results

Study researchers had higher call rates and higher consent rates than call center operators. Study researchers were able to place 18.2 calls per hour as compared to the call center operators that placed 8.0 calls per hour. The study researcher telephone consent rate (number of participants that consented to participate / total number of calls dialed) was 6.1%. The call center operators, however, only had a telephone consent rate of 2.3%. Furthermore, study researchers consented 1.1 participants per hour as opposed to call center operators that consented 0.2 participants per hour. The telephone decline rate was the same for both study researchers and call center operators at 7.3%.

Conclusions

Study researchers may be more successful at eliciting informed consent than others without the same background. While it may initially appear economical and more effective to employ an external call center, the study may lose participants that would otherwise consent when speaking with an engaged researcher. Call center operators are provided with an extensive list of frequently asked questions, however, they do not have the knowledge to answer questions outside of the given

information. It is more cost effective to employ a study researcher at a higher pay scale that can achieve more study participation than numerous call center operators. Additionally, the number of participants enrolled determines a study's feasibility. Because studies are asking important clinical questions, it is important to ensure they are possible to complete in statistically rigorous and efficient manners.

Telephone informed consent is an important tool to embed research into clinical care and reduce the infrastructure burden of research, but it must be done carefully by engaged individuals. Further study will be necessary to define "engagement" and provide better parameters for researchers conducting telephone based informed consent.

Contributors

Sarah Leatherman
Department of Veterans Affairs

Alison Majkut
Department of Veterans Affairs

Patricia Woods
Department of Veterans Affairs

Ryan Ferguson
Department of Veterans Affairs

TIME-COURSE GRAPHICS FOR INDIVIDUAL SUBJECTS IN DMC REPORTS

BRIAN TRAVER

UNIVERSITY OF WISCONSIN-MADISON

A Data Monitoring Committee (DMC) is responsible for the monitoring of patient safety and treatment efficacy of ongoing clinical trials. Interim DMC reports provide analyses in the form of tables and graphics that are frequently aggregated by treatment group.

The Statistical Data Analysis Center (SDAC) at the University of Wisconsin–Madison specializes in producing interim reports and analyses for DMCs. Our reports are graphically based, allowing DMC members to easily identify differences between treatment groups, and/or changes over time, and to review a large amount of information in a short amount of time.

DMCs may also require subject-level presentation of the data in order to look more closely at particular cases of concern. For example, an aggregate graphical view of laboratory data, such as box plots, may reveal outliers suggesting hepatotoxicity but provide no additional information. Plots displaying multiple types of data over time, such as laboratory results, adverse events and dosing, for individual subjects may provide additional insight into the possible underlying pathology. Concurrent elevated liver enzymes could indicate drug-induced liver injury. However, elevated liver enzymes could also occur in the presence of viral hepatitis. These patterns, lost in aggregate analysis, are revealed in the subject level graphical presentation.

In our presentation we will show examples of graphics for individual subject data over time with features that we have found useful for DMC reports. Types of data presented may include demographics, subject disposition (randomization, withdrawal from treatment or study, death, completion, etc.), dosing, event data (adverse events, positive antibody results), and continuous data (laboratory measures, vital signs, bone density). Relevant text, such as a listing of the subject's adverse events, may also be presented adjacent to each graphic panel.

Contributors

Melissa Schultz
University of Wisconsin-Madison

UNDERSTANDING PROGRESSION FROM AN INTERNAL PILOT TO A MAIN TRIAL: EXPLORING VIEWS AND PERCEPTIONS OF KEY STAKEHOLDERS.

KATHERINE FAIRHURST

CENTRE OF SURGICAL RESEARCH & MEDICAL RESEARCH COUNCIL (MRC) CONDUCT-II HUB FOR TRIALS METHODOLOGY RESEARCH (COLLABORATION AND INNOVATION FOR DIFFICULT AND COMPLEX RANDOMISED CONTROLLED TRIALS IN INVASIVE PROCEDURES), BRISTOL MEDICAL SCHOOL, DEPARTMENT OF

Introduction

Progression criteria are typically used to evaluate viability of main trials. Funders review whether targets for internal pilot phases have been met (e.g. recruitment/retention), and if it is appropriate for studies to proceed. Little is known about how decisions to proceed, amend or abandon main trials are made. In-depth qualitative work is required to explore and understand the practices and views of funding body representatives regarding the design and conduct of pilot/feasibility work and the selection, review and implementation of progression criteria for main trials with internal pilot phases.

Methods

Funding body panel representatives, with current or recent experience of holding a position on a UK national research funding panel (for minimum 2 years) were identified. Maximum variation, purposive sampling was used to ensure inclusion of a range of participant characteristics, including experience on different funding panels, of working in different UK areas, and in different methodological/research roles. Semi-structured interviews informed by a topic guide were conducted to explore participants' experiences, views and perceptions regarding how pilot/feasibility work is designed and conducted and how the selection and implementation of progression criteria for internal pilot work occurs. Analyses were undertaken, using Braun and Clarke's approach to thematic analysis, in an iterative and cyclical process as further interviews took place. Interviews continued until no new themes emerged and established themes ceased to evolve.

Results

Of 27 participants contacted, n=19/27 (70%) consented and were interviewed. All participants had current or recent experience as a chair/deputy chair/funding panel member. Interviews were undertaken over 7 months, in 3 iterative phases, mostly by telephone (n=15/19) with the remainder (n=4/19) face-to face, at a mean length of 58 minutes (range 30 to 88). Key issues identified by participants included persistent misguided design and reporting of pilot/feasibility work despite extensive methodological work to define the nuances of definitions and reporting. Explanations included inaccessibility of methodological information to the wider clinical/research community. Participants perceived the current system of determining pre-agreed progression criteria, suggested by trial teams, potentially adjusted by funders and applied flexibly but stringently using judgment and experience, as imperfect, though a transparent flexible system to manage risk for all involved was considered vital. Key suggestions for improvement included

assessing trial progression at multiple time points, running more trials with an internal pilot phase (as opposed to standalone pilot/feasibility work), and developing guidance tools, including greater emphasis on public reporting, to assess if, when and what type of pilot/feasibility work is needed for clinical trials.

Conclusions

This work forms part of a larger project which is reviewing the decision-making process around completed internal pilots funded by the NIHR HTA programme in the UK. The qualitative work illustrates on going confusion about the place and purpose of pilot and feasibility work. The current system for selecting and implementing progression criteria for internal pilots could be further refined. Overall, this project aims to develop recommendations and guidance on pilot/feasibility work for widespread implementation amongst those designing, funding and conducting clinical trials in the UK.

Contributors

Kerry Avery
Centre of Surgical Research & Medical Research Council (MRC) ConDuCT-II Hub for Trials Methodology Research (Collaboration and innovation for Difficult and Complex randomised controlled Trials In Invasive procedures), Bristol Medical School, Department of

Alicia O'Cathain
School of Health Related Research (ScHARR), University of Sheffield, UK.

Pat Hoddinott
NMAHP-RU (Nursing, Midwifery and Allied Health Professions Research Unit), University of Stirling & Medical Research Council (MRC) ConDuCT II Hub for Trials Methodology Research, UK.

Jane Blazeby
Centre of Surgical Research & Medical Research Council (MRC) ConDuCT-II Hub for Trials Methodology Research (Collaboration and innovation for Difficult and Complex randomised controlled Trials In Invasive procedures), Bristol Medical School, Department of

USING ELECTRONIC DATABASES TO INCREASE ACCRUAL TO GASTROINTESTINAL CANCER CLINICAL TRIALS

PABLO SERRANO

MCMASTER UNIVERSITY

Introduction: In a given jurisdiction, percent accrual to gastrointestinal randomized clinical trials is often in the single digits. We tested if such rates could be improved through the use of population-based electronic databases. The ePATH database, maintained by Cancer Care Ontario, collects and codes pathology reports related to cancers from across Ontario (population 14.5M), including hospitals and clinics in the Local Health Integration Network 4 (LHIN4) (population 1.4M). OneView is an electronic repository of diagnostic imaging (computed tomography and magnetic resonance imaging scans) performed in Southwestern Ontario, funded by participating Hospitals and the federal government. Pathology reports and radiological images are uploaded in "real-time". The aim of our study was to compare traditional methods of patient accrual, including letters addressed to clinicians (surgeons, medical and radiation oncologists as well as gastroenterologists in the LHIN4 region) and prospective collection of cases discussed at multidisciplinary tumour boards versus use of ePATH and OneView to identify potential participants to the "Resection of Colorectal Cancer with Synchronous liver metastases" (RESECT) study, a trial evaluating postoperative complications and quality of life of patients following simultaneous resection of colon cancer and liver metastases.

Methods: ePATH was prospectively reviewed on a biweekly basis by a research assistant to identify patients diagnosed with primary colorectal cancer from January to November 2017. Cases were then linked to OneView files. Actual images from the latter were reviewed by a hepatobiliary surgeon to identify patients with potentially resectable liver metastases and absence of extrahepatic disease, as specified in the RESECT protocol. A priori we decided that an identification ratio ? 1.3 would be clinically significant – that is, the ratio of cases identified as potentially eligible for RESECT using electronic databases versus traditional methods.

Results: Of 235 patients with a diagnosis of primary colorectal cancer abstracted from ePATH, 25 patients were identified as potentially eligible for the RESECT trial based on OneView imaging. Traditional methods identified 16 patients, resulting in a population-level identification ratio of 1.6 (25/16).

Conclusion: We identified more cases for the RESECT trial using population-based electronic databases versus traditional methods. The use of electronic databases may be an efficient and valuable method of increasing accrual to gastrointestinal randomized clinical trials.

Contributors

Sameer Parpia
McMaster University

Donna McCarty

McMaster University

Natalie Solis
McMaster University

Marlie Valencia
McMaster University

Sinan Jibrael
McMaster University

Alice Wei

University of Toronto

Steven Gallinger

University of Toronto

Marko Simunovic

McMaster University

VIDEO ADJUDICATION IN CLINICAL TRIALS: ENABLING DISTRIBUTION OF EXPERTISE AND ACCURATE SCORING

ALAN STEVENSON

ROBERTSON CENTRE FOR BIOSTATISTICS

Background

Within stroke trials, there is a need to provide fair, unbiased scoring of participant assessments, based on their recovery over specified time points. In order to provide this, a key stumbling block can be assembling a reliable team of adjudicators, and coordinating that adjudication. If relying on a more distributed group, the issues then include providing access to paper notes, and tracking their comments out with a centralised meeting.

Methods

By capturing assessments on video and developing a system to manage those videos, we can allow study coordinators to manage these assessments, receiving them from multiple centres and distributing them to adjudicators regardless of geographic location.

By allowing adjudicators to actually view the assessments, rather than rely on written notes, we remove a layer of obfuscation from the adjudication, with adjudicators being able to directly judge participant reactions for themselves.

We can provide a toolset that can assign algorithms to the selection of adjudicators by criteria such as location, or provide translations of foreign language assessments, simplifying the work in distributing the assessments. We can also provide integration with study eCRFs to report adjudicated scores back to the main study datastore, while maintaining metrics on the differences in scoring, leading to quality checks on the assessments being done in centres.

Conclusion

We will discuss the approach of designing a generic, multi-trial system to accomplish these goals, and real life experience from its operation in multiple concurrent trials.

<u>Contributors</u>

Sharon Kean
Robertson Centre for Biostatistics

WHY NOT BUCKLEY-JAMES? A SIMULATION STUDY

KRISTEN WROBLEWSKI

UNIVERSITY OF CHICAGO

Buckley and James (1979) introduced a method for estimating parameters in a linear regression model with censored data. The procedure and its theoretical properties have been investigated by Ritov (1990), Lai and Ying (1991), Jin, Lin, and Ying (2006), and Zhou and Li (2008), who also introduced various extensions and modifications. An attractive feature of the Buckley-James (BJ) procedure is that it admits an accelerated failure-time (AFT) interpretation and is therefore a useful alternative to the Cox regression model (Wei, 1992). The method is nonparametric, requiring no assumptions about the underlying error distribution other than zero mean and finite variance. Buckley and James provided a formula for the standard error of the estimated parameter vector based on the non-censored observations. Weissfeld and Schneider (1987) described an alternative variance estimator, as well as two bootstrap procedures for drawing inferences.

In this work we consider two aspects of the BJ procedure. The first is the reliability of the BJ variance estimate— previous numerical studies suggest that it can be unstable (Hillis, 1993). Second is the bias and efficiency of the BJ estimator relative to parametric methods for fitting the AFT model. We found that the BJ variance estimator tends to underestimate the variance, leading to confidence intervals with less than nominal coverage probability. The use of the bootstrap for deriving standard errors and confidence intervals is therefore advocated. With regard to efficiency, we found that when the underlying distribution is normal the BJ estimator has relatively high efficiency. For other error distributions (exponential and Weibull), however, there was a 25%-40% reduction in efficiency compared to parametric modelling. On the other hand, the BJ procedure always produced unbiased estimates, whereas parametric modelling yielded biased estimates if the wrong model was fit. Lack of convergence (oscillating solutions) is known to be a problem with the BJ estimator but was infrequent in our simulations (approximately 1% or less). The BJ method has seldom been applied in the analysis and reporting of results from clinical trials; its use may warrant further consideration.

Contributors

Theodore Karrison
University of Chicago

Contributors