Abstracts

Abstracts appearing in this SCT final program book are printed as submitted by the author(s) of that abstract. Unless instructed by the author, the content of the abstract was not edited or reformatted.

A01

DO ADVANCE TELEPHONE CALLS AHEAD OF REMINDER QUESTIONNAIRES INCREASE RESPONSE RATE IN NON-RESPONDERS COMPARED TO QUESTIONNAIRE REMINDERS ONLY: THE RECORD PHONE TRIAL

Graeme MacLennan, Alison McDonald, Gladys McPherson and Alison Avenell for the RECORD Trial

Health Services Research Unit, University of Aberdeen

Background: Postal questionnaires are simple and economical for collecting outcome data for RCTs but are prone to non-response. In the RECORD trial (a large pragmatic publicly funded RCT in UK) 40% of non-responders returned reminder questionnaires at 1-year. In subsequent years we investigated the effect of an advance telephone call to non-responders (that pre-notified delivery of the reminder questionnaire, emphasised the importance of the questionnaire, thanked them for their time and congratulated them for completing the year) on response rate to reminder questionnaires and the next questionnaire four months later.

Methods: Non-responders to annual questionnaires were randomised to receive a telephone call from the trial office ahead of the reminder questionnaire (N=390) or to a control group that received a reminder questionnaire only (N = 360). The primary outcome was response to the reminder questionnaire within 21 days; secondary outcome was response to a questionnaire four months later. Results are presented as percentage difference and 95% CI for both ITT and ATT (using instrumental variables) analyses.

Results: Response was 67.8% (265/390) in the telephoned group compared to 62.5% (227/363); ITT estimate 5.4% increase (95%CI -1.4 to 12.2). Four months later percentages responding were 51.8% (202) and 42.7% (155); ITT estimate 9.1% increase (95%CI 2.0 to 16.2). In the telephoned group 12.3% (48/390) participants were not telephoned because questionnaires were returned before the scheduled telephone call. ATT estimates adjusting for this were 6.2% (95% CI -1.6 to 14.0) and 10.4% (95% CI 2.2 to 18.5) respectively.

Conclusions: The telephone call resulted in a slight increase in response to the reminder questionnaire, however at four months later the proportion in the telephoned group responding was greater. This study suggests that pre-notification telephone calls may only be worthwhile if further questionnaires are to be sent out soon after reminder questionnaires.

A02

MONITORING COMPLIANCE ON CANCER CLINICAL TRIALS: A PEER REVIEW MECHANISM FOR CONTINUOUS QUALITY IMPROVEMENT

(1) Erica E. Love (2) Yelena Novik (3) Lisa Gaynes & (4) Nick Shuman

(1) NYU Cancer Institute, NYU School of Medicine (2) NYU Cancer Institute, (3) NYU School of Medicine NYU Cancer Institute, (4) NYU School of Medicine NYU Cancer Institute, NYU School of Medicine

Internal monitoring and self-auditing on cancer clinical trials is essential in maintaining data integrity used to make therapeutic decisions, in protecting research subjects, and in ensuring proper use of resources. NYU Cancer Institute implemented an Internal Audit Committee (IAC) in 2006 to ensure compliance and improve data quality in Investigator Initiated trials (IITs). Quarterly, 10% of cases enrolled on IITs are randomly selected for auditing. Auditors consist of investigator physicians, regulatory and research coordinators, and research nurses to ensure a peer review process. Elements reviewed during audits include informed consent, screening and eligibility and protocol compliance. According to National Cancer Institute (NCI) guidelines, deficiencies are categorized as major or minor and overall results as acceptable, acceptable, needs follow up or unacceptable. On average the IAC audits 27 cases each year. Since 2006 the total number of cited deficiencies has dropped significantly. In 2007 an average of 16 major deficiencies were cited per audit. By 2009, the number of major deficiencies per audit was down to an average of 8; a 50% decrease in just two years. In 2011, the IAC implemented an additional level of monitoring entailing quality assurance (QA) reviews of the
first two cases and the third case on each IIT. In 2012, QA reviews were conducted on 5 cases. Investigators are made aware of any deficiencies from the first two cases and are asked to implement a corrective action plan (CAP). The third case is reviewed to ensure the CAP was successful in eliminating any deficiencies to ensure continuous quality improvement. Implementing these internal auditing systems resulted in consistent standards of monitoring and audits, increased participation amongst investigators and research staff in the audit process, decreased number of deviations, greater feedback from PIs to audit findings, and CAPs addressing common causes of deviations.

A03
IMPLEMENTING CLINICAL TRIALS METHODOLOGY INTO THE EVERYDAY CARE OF ABUSED CHILDREN
Chad Shenk
Cincinnati Children’s Hospital Medical Center, University of Cincinnati College of Medicine

Child maltreatment affects 700,000 children each year in the United States and increases the risk for a number of adverse health outcomes throughout the lifespan. Child advocacy centers (CACs) are often the initial stage of care for this population where multiple health services, including targeted prevention and treatment implementation, are delivered as part of the standard of care. However, bridging the gap between typical service delivery and treatment evaluation is often difficult, as CACs typically do not have the methodological expertise to evaluate clinical services formally. This presentation provides a framework for implementing state-of-the-art clinical trials methodology into the everyday practice of CACs to identify the optimal course of treatment for children exposed to abuse. The use of sequential multiple assignment randomized trials (SMARTs) will demonstrate how multiple levels of care can be evaluated within the context of ongoing healthcare delivery at a CAC. SMARTs are the ideal methodology for evaluating treatment services in clinic settings as they model the treatment decisions made by families and practitioners at various stages of care. For instance, targeted prevention is often the first line treatment for abused children. However, up to 70% of abused children do not respond to targeted prevention and require additional treatment. Additional treatment often involves behavioral or pharmacological interventions or both to achieve a response. Implementing a SMART in such a setting would identify: 1) the most effective initial intervention for abused children, 2) the optimal dose of the initial intervention, and 3) the most effective second-line intervention for those children not responding to targeted prevention. Specific measurement strategies will be reviewed to promote the assessment of predictors of treatment response as well as mediators of long-term treatment response.

A04
FROM LATE PHASE TO EARLY PHASE: EXPERIENCES OF ESTABLISHING A DISEASE-SPECIFIC EARLY PHASE TRIALS UNIT
(1) Sarah Brown (2) Louise Flanagan (3) Fiona Collinson (3) Sarah Flynn (4) Nina Thenagaran (5) Deborah Sherratt (5) Avie-Lee Coney
Clinical Trials Research Unit, University of Leeds

In November 2009, the Clinical Trials Research Unit at the University of Leeds successfully became the Myeloma UK (MUK) Clinical Trials Coordinating Office (CTCO) for the design, coordination and analysis of a network of early phase trials in myeloma. In an initiative developed by MUK, a dedicated team of clinicians, trialists and statisticians were brought together to streamline the clinical trial development process, from trial concept to first patient entered. Facilitating knowledge retention and specialization, the ultimate aim of the MUK CTCO is to promote bench to bedside research and allow more patients faster access to novel therapies. With an international track record in late phase trials, the Clinical Trials Research Unit initiated a strategy to transfer these skills to the early phase setting. Key areas associated with the successful development of the MUK CTCO included: Establishing a core team of researchers, Initiating a training program, Producing standard operating procedures, Standardization of case report forms and databases, Statistical design, Identifying independent umbrella committees, Pharmaceutical collaboration Developing trials of novel therapies, as well as providing initial evidence for use of licensed therapies in different settings, the initiation of the MUK CTCO has led to the successful implementation of five phase I and II trials.
over three years (1 in follow-up, 1 in recruitment and 3 in set-up), with a further two currently in development. We will discuss the key areas associated with the development of the MUK CTCO, from initiation to the successful set-up of five early phase clinical trials, and examine the different challenges associated with setting up and running phase I/II trials compared with large randomized controlled phase III trials and the impact this has on workload, timelines and the establishment of new processes.

A05
PULLING THE TRIGGER ON RISK-BASED MONITORING
Rick Morrison
Comprehend Systems

Due to increasing industry pressures, drug and device makers are continually looking for ways to save money and run more efficiently. On-site monitoring can account for anywhere from 25 to 30 percent of the overall cost of a clinical trial. A shift to risk-based monitoring is a critical strategy to alleviating this strain. Modern analytics tools and technologies are driving the emergence of risk-based monitoring because they enable powerful insights into data that haven’t been possible in the past. Modern alerting systems can incorporate rules from analytics systems specifically looking for known or anticipated anomalies. When alerts are triggered, processes can be implemented to react, even with on-site monitoring if necessary. As more and more anomalies are found, a library of alerts can be developed to categorize and respond to unique risks immediately. Over time, the collection can become sophisticated and hugely effective in identifying problems proactively greatly reducing the need for on-site monitoring. In this presentation, Rick Morrison, CEO and co-founder of Comprehend Systems, will discuss how sponsors can transition to risk-based monitoring practices using advanced technologies and analytics. Attendees will gain an understanding of how to get started with risk-based monitoring, including: Which types of trials are appropriate for centralized monitoring practices. How to determine and apply the effective trigger factors. What processes, procedures and tools need to be put into place. Where risk-based monitoring strategies are heading in the future. With these underlying tools and strategies in place, clinical trial sponsors can automate and streamline complex criteria and processes to enhance the efficiency of the entire site monitoring process leading to significant cost reduction of each trial.

A06
THE ROLE OF THE DATA MANAGER IN RISK-BASED MONITORING
Aaron Perlmutter
Data Coordination Unit, Division of Biostatistics and Epidemiology, Medical University of South Carolina

The high cost of on-site monitoring is well documented in the literature, but the return on this investment is less clear (Duley et al, 2008; Korieth, 2011). The development of electronic data capture and clinical trial management systems provides instant access to data and real-time quality checks, and a mechanism to perform monitoring centrally and remotely (KAI, Feb 2012). In recent years, a greater emphasis has been placed on a risk-based monitoring approach that incorporates centralized monitoring as a way to monitor data quality, and the FDA proposed a greater reliance on risk-based monitoring in its 2011 Draft Guidance for Industry, Oversight of Clinical Investigations - A Risk-Based Approach to Monitoring. Centralized monitoring can be used to assess adherence to study procedures and to identify problematic data and sites (Baigent et al, 2008). While centralized monitoring, when combined with on-site monitoring, can increase a trial’s data quality, the concept of risk-based monitoring is a relatively recent development in clinical trials, and guidance is limited regarding the role of the data manager in centralized monitoring. Data managers are in a unique position to assist in the process as they likely perform some centralized monitoring already, including checks for range and logic (e.g. correct time sequence, no contradictory data points), as well as completeness and timeliness of data entry. With the appropriate tools, the data manager can identify poorly performing sites and suspicious data patterns. When this information is shared with the study team and acted upon, the integrity of the trial can be improved. Centralized monitoring is as an important component in risk-based monitoring strategies (Brosteanua et al, 2009). However, beyond basic checklists, its methodologies are not clearly defined. This presentation attempts to define the data manager’s role in risk-based monitoring and to demonstrate tools to perform centralized monitoring in a CTMS.
A07
IDENTIFYING ATYPICAL SITES WITH CENTRAL STATISTICAL MONITORING
Erik Doffagne
CluePoints Inc.

Regulatory agencies are moving forward by encouraging alternative approaches that streamline the clinical trials costs without compromising the quality or the scientific validity. A recent draft guidance from FDA suggests that risk based approaches should be used for clinical trials monitoring. A similar position has been taken by the EMA in a reflection paper. The industry is now considering the solutions for implementing novel monitoring strategies driven by indicators of the investigational sites quality. A common way to achieve the evaluation of the site quality or performance is to look on predefined metrics, often called Key Risk Indicators (KRI). In this talk we introduce a complementary layer to KRI: a Central Statistical Monitoring (CSM) approach based on advanced statistical methods and data mining tools aiming at detecting sites with atypical patterns in the data. Atypical patterns might reflect different kinds of problems from fraud (e.g. invented patients) to quality issues (i.e. underreporting). We demonstrate a concrete CSM implementation and illustrate how the CSM processes can be designed. The requirements and the challenges of the proposed approach are discussed.

A08
RISK-BASED MONITORING APPROACH IN PRACTICE – COMBINATION OF REAL-TIME CENTRAL MONITORING AND ON-SITE SOURCE DOCUMENT VERIFICATION
Wenle Zhao
Medical University of South Carolina

About two thirds of clinical trials conducted in US are funded by federal agencies and other non-profit organizations. While 100% source document verification for industry sponsored studies has been the standard for decades, it is rarely the case for federal funded trials managed by academic research institutions due to funding limitations. The use of web-based clinical trial management system and the risk-based monitoring strategies proposed by the FDA’s Draft Guidance for Industry, “Oversight of Clinical Investigations – A Risk-Based Approach to Monitoring,” promote a hybrid approach to monitoring. By combining a real-time central monitoring function with risk-based selective on-site source document verification, higher trial operation and clinical data quality can be made more affordable. This presentation covers the system design of the risk-based monitoring approach and its incorporation into a CTMS for the Neurological Emergencies Treatment Trials Network funded by NINDS. The risk-based monitoring module consists of three components, a central monitoring component, an on-site monitoring visit planner, and an on-site monitoring report. A complete monitoring check list is developed by the study team based upon regulatory requirements and the study protocol. All answers to check points that can be derived based upon the data in the study database is propagated into the central monitoring component, which is available to authorized users in real-time. The monitoring visit planner provides site trial performance risk assessment information for monitors to better prioritize on-site monitoring tasks. The monitoring report records all site performance based findings, identified by both on-site and central monitoring.

A09
INTERACTIVE STATISTICAL SOFTWARE FOR ESTIMATING THE OPERATING CHARACTERISTICS OF CANCER PHASE I CLINICAL TRIAL DESIGNED WITH STANDARD 3+3 ALGORITHM
(1,2) Zhibo Wang, (3) Taofeek K Owonikoko, (1,2) Jeanne Kowalski, (3) Fadlo R Khuri, (1,2) Zhengjia Chen
(1) Department of Biostatistics and Bioinformatics, Emory University (2) Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute, Emory University. (3) Department of Hematology and Medical Oncology, Emory University

Standard 3+3 with or without dose de-escalation designs are still most widely used in cancer Phase I clinical trials for their simplicity and robustness. In clinical practice, it is necessary to provide the operating characteristics of clinical trials at the planning stage. Lin and Shih (2001) developed the formula for estimating the operating characteristics of the Phase I clinical trials with Standard 3+3 designs. Unfortunately, the formula
might be difficult for clinical practitioners without advanced statistics knowledge to understand and implement. Therefore, we develop an interactive and user friendly software called Standard_3+3_design©TM which can estimate the trial’s operating characteristics, including the probability of a dose being chosen as the maximum tolerated dose (MTD); the expected number of patients treated at each dose level; target toxicity level (i.e. the expected dose limiting toxicity (DLT) incidences at the MTD); expected DLT incidences at each dose level; and expected overall DLT incidences in the trial. The software has the standalone version for free download and the web-based version for direct use online at Emory University website: http://sisyphus.emory.edu/Software_clinical.html.

A10

ESCALATION WITH OVERDOSE CONTROL USING ALL TOXICITIES AND TIME TO TOXICITY DATA FOR CANCER PHASE I CLINICAL TRIALS

(1) Ye Cui (2) Taofeek K Owonikoko (3) Zhibo Wang (4) Fadlo R Khuri (5) Jeanne Kowalski (6) Zhengjia Chen

(1) Georgia State University (2) Department of Hematology and Medical Oncology, Emory University (3) Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute. Department of Biostatistics and Bioinformatics, Emory University (4) Department of Hematology and Medical Oncology, Emory University (5) Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute. Department of Biostatistics and Bioinformatics, Emory University (6) Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute. Department of Biostatistics and Bioinformatics, Emory University

The primary purpose of cancer Phase I clinical trial which is a critical step in development of new drug against cancer is to determine the maximum tolerated dose (MTD) and schedule of new drug. It is usually a small study with limited data so that fully utilisations of all toxicities and time to toxicity data are essential to improve the trial efficiency and accuracy of MTD estimation. Chen et al. (2010) proposed a novel normalized the equivalent toxicity score (NETS) system which can fully utilize multiple toxicities per patient instead of a binary indicator of dose limiting toxicity (DLT). Cheung et al. (2000) developed the time of event (TITE) approach to incorporate time to toxicity data. Escalation with Overdose Control (EWOC) is an adaptive Bayesian Phase I design which can allow rapid dose escalation while controlling the probability of overdosing patients. In this study, we use EWOC as a framework and integrate it with the NETS system and TITE approach to develop an advanced Phase I design entitled EWOC-NETS-TITE. Simulation studies have been conducted to compare its operating characteristics with those of EWOC and standard 3+3 design. Simulation results demonstrate that EWOC-NETS-TITE can not only substantially improve the trial efficiency and MTD accuracy, but also allow patients to be entered in a staggered fashion and shorten trial length. A user-friendly software of EWOC-NETS-TITE is under development and will be available in the future.

A11

CONTINUOUS TUMOR SIZE CHANGE PERCENTAGE AND PROGRESSION FREE SURVIVAL AS ENDPOINT OF THE FIRST AND SECOND STAGE RESPECTIVELY IN A NOVEL DOUBLE SCREENING PHASE II DESIGN

(1) Zhengjia Chen, (2) Ye Cui, (3) Taofeek K. Owonikoko, (4) Zhibo Wang, (5) Dong M. Shin, (6) Fadlo Khuri (7) Jeanne Kowalski

(1) Department of Biostatistics and Bioinformatics, Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute, Emory University. (2) Department of Mathematics and Statistics, Georgia State University. (3) Department of Biostatistics and Bioinformatics, Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute, Emory University. (4) Department of Hematology and Medical Oncology, Emory University. (5) Department of Biostatistics and Bioinformatics, Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute, Emory University. (6) Department of Hematology and Medical Oncology, Emory University. (7) Department of Biostatistics and Bioinformatics, Biostatistics and Bioinformatics Shared Resource at Winship Cancer Institute, Emory University

A phase II trial is an expedite and low cost trial to screen potentially effective agents for the following phase III trial. Unfortunately, the positive rate of Phase III trials is still low although agents have been determined to be effective in proceeding Phase II trials mainly because the different endpoints are used in Phase II (tumor response) and III (survival) trials. Good disease response often leads to, but can NOT guarantee, better
survival. From statistical consideration, transformation of continuous tumor size change into a categorical tumor response (complete response, partial response, stable disease, or progressive disease) according to World Health Organization (WHO) or Response Evaluation Criteria In Solid Tumors (RECIST) will result in a loss of study power. Tumor size change can be obtained rapidly, but survival estimation requires a long time follow up. We propose a novel double screening phase II design in which tumor size change percentage is used in the first stage to select potentially effective agents rapidly for second stage in which progression free or overall survival is estimated to confirm the efficacy of agents. The first screening can fully utilize all tumor size change data and minimize cost and length of trial by stopping it when agents are determined to be ineffective based on low standard and the second screening can substantially increase the success rate of following Phase III trial by using similar or same outcomes and a high standard. Simulation studies are performed to optimize the significant levels of the two screening stages in the design and compare its operating characteristics with Simon’s two stage design. A real phase II clinical trial is provided to demonstrate the substantial improvement in trial efficiency and success rate of Phase III clinical trial as well as significant reduction in the cost and length of trial.

A12
IDENTIFYING CONTINUOUS PHASE II ENDPOINTS USING RECIST 1.1 DATA WAREHOUSE MEASUREMENTS
(1) Ming-Wen An, (2) Sumithra J. Mandrekar, (3) Xinxin Dong, (4) Axel Grothey, (5) Jan Bogaerts, (6) Daniel J. Sargent
(1) Vassar College, (2) Mayo Clinic, (3) University of Pittsburgh, (4) Mayo Clinic, (5) European Organization for Research and Treatment of Cancer Headquarters, (6) Mayo Clinic

Background: The high failure rate (e.g. 50-60%) of Phase III oncology trials represents a major obstacle to therapeutic development. We previously reported that alternative cut-points to RECIST as well as alternate categorical metrics provided no meaningful improvement in overall survival (OS) prediction. In this work, we examined continuous RECIST measurement based metrics as alternative phase II oncology trial endpoints.

Methods: We assessed the predictive ability of multiple candidate continuous metrics (first and last slope; first and last % change; longest duration of stability; and indicator of >10% increase from the previous assessment) based on longitudinal cycle-by-cycle tumor measurements from 1039 patients based on the RECIST 1.1 data warehouse. Data from 13 trials were randomly split (60:40) into training and validation sets. For each set of metrics, Cox models of OS were fit separately for breast, lung, and colorectal, adjusting for baseline tumor burden, stratifying by number of lesions (\(<=3\) vs. \(>3\)) and study. Data were landmarked at 12-weeks. Predictive ability was assessed via c-index, Hosmer-Lemeshow goodness-of-fit-type statistics, and hazard ratios/p-values.

Results: Metric predictive ability for OS differed across disease group, with poorest prediction in Breast (c-index range for the various metrics, training set: 0.51-0.61 Breast; 0.56-0.64 Lung; and 0.58-.72 Colorectal). While most metrics demonstrated a statistically significant association with OS, none of the continuous metrics performed better than the current categorical RECIST classification of CR/PR versus SD versus progression (c-indices: 0.58 Breast, 0.60 Lung, 0.65 Colon).

Conclusions: The continuous metrics we considered perform no better than RECIST in predicting patient survival across multiple tumor types. Ongoing work is focusing on issues of missing data and multicollinearity. At the present time, no metrics have been demonstrated to out-perform RECIST, indicating its use continues to be appropriate due to both its simplicity and its relative predictive ability.

A13
POWER ANALYSIS FOR SEAMLESS PHASE II/III DESIGNS WITH SUBPOPULATION SELECTION AND A CHANGE OF SURVIVAL, BINARY AND NORMAL ENDPOINTS
(1) Qian Wu (2) Qiang Zhang
(1) Department of Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
(2) Radiation Therapy Oncology Group, American College of Radiology, Philadelphia, PA

Seamless designs provide a powerful way to demonstrate the effect of a new therapy with less sample size and shorter duration. In a phase II/III design, the decision of stopping or continuing to confirmative phase III is
based on the results of efficacy or futility from the learning phase II, such as the subpopulation selection design of Jenkins et al. (Pharmaceutical Statistics, 10:347-356, 2011). In recent oncology trials, progression free survival (PFS) is usually the primary endpoint in phase II and overall survival in phase III study. These two survival endpoints are correlated. Usually we assume exponential distribution for these survival times. However, we often observe non-constant hazard rate, so in this research we consider the more flexible Weibull distribution. In addition, other correlated novel endpoints have been proposed recently for phase II oncology trials, such as tumor size measured by imaging tools as continuous endpoint and/or the typical binary response rate. In order to account for the correlation between survival and binary, normal endpoints, Gaussian Copula distributions are used to simulate phase II/III data of different types. Based on the subpopulation selection methods by Jenkins (2011) for seamless phase II/III design, we conducted the power analysis and compared the results between designs using Weibull and exponential distributions. Similar designs and analyses were also done for normal and binary cases. We demonstrated that the type-I error rate is controlled at design level, such as 5% for each type of phase II endpoint with various possible correlations to overall survival, by Monte Carlo simulations. Multiple comparison issues are also discussed. Results show that consideration of correlation between different early endpoints and overall survival is important to protect type I and type II error rates and designs proposed here can be applied to a wide variety of settings including oncology trials.

A14
SEAMLESS PHASE II/III DESIGN FOR ONCOLOGY TRIALS WITH TREATMENT SELECTION AND A CHANGE OF SURVIVAL ENDPOINTS
(1) Qiang Zhang; (2) Qian Wu
(1) Radiation Therapy Oncology Group, American College of Radiology, Philadelphia, Pennsylvania, USA
(2) Department of Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Seamless phase II/III designs have been proposed in recent years and these designs are applied to different disease settings in oncology area. Advantages of these methods include great savings on sample size and trial duration. The goal of the phase II trial is to select the most effective regimen when there are multiple experimental arms. This is done through comparing each of the new treatments to a common control arm using a chosen endpoint, such as progression free survival. The winner will be tested further for overall survival in phase III study. We propose a new seamless design in which there are two phases, at the end of phase II, only one arm will be selected to go on to phase III. Patients accrued during phase II will be included in the definitive testing. We use group sequential method to design each component. Separate interim analyses are built in such that each phase can be stopped early for either efficacy or futility. Once sample sizes are derived, operating characteristics for the seamless II/III design are evaluated through simulations under the null and various alternative hypotheses. For these analyses, we account for possible correlations between two survival times. Issues such as multiplicity and accrual suspension are discussed. Control of the overall error rate can also be achieved based on asymptotic joint distributions of test statistics in addition to simulations. We will present expected sample sizes and stopping times using different design error rates and power. Savings on sample size and time will be compared to typical separate designs. A recently approved RTOG Head and Neck cancer study will be used as an example to illustrate how to design such trials. Results show that the newly proposed design has desired properties, which offer cost effectiveness, operational efficiency and most importantly, scientific innovation.

A15
APPLICATION OF A SEAMLESS PHASE II/III, GROUP-SEQUENTIAL DESIGN IN THE DEVELOPMENT OF A S. AUREUS VACCINE
(1) Jonathan Hartzel (2) Ivan Chan
(1) Merck & Co., Inc. (2) Merck & Co., Inc.

Deep sternal wound infection (DSWI) following cardiac surgery occurs in 0.5-3% of patients, is frequently attributable to Staphylococcus aureus, and is associated with increased morbidity/mortality. A novel vaccine candidate (V710) against S. aureus was shown to be immunogenic and generally well-tolerated in Phase I studies. Due to the low incidence of infection, a large efficacy study was required to assess the efficacy and
safety of V710 in preventing S. aureus bacteremia and/or DSWI in patients undergoing cardiac surgery. An innovative event-driven, group-sequential design was employed to conduct a seamless Phase II/III double-blind, placebo-controlled efficacy study. Details of the study design will be discussed including its advantages over conducting separate Phase II and III studies and the logistical complications it can introduce. Additionally, the surprising results of the study, which was stopped early for low efficacy and for a potential safety signal, will be reviewed.

A16

PRAGMATIC SEAMLESS DESIGN FOR AN EFFICACY TRIAL OF ASTHMA MANAGEMENT WITH REDUCED COST

(1) Mei Lu, (1) Edward Zoratti, (1) Dennis Ownby, (2) Douglas Roblin, (3) Christine Johnson, (1) Christine Joseph

(1) Department of Public Health Sciences, Henry Ford Health System (2) Department of Pediatrics and Internal Medicine, Medical College of Georgia (3) Kaiser Permanente Center for Health Research Southeast

Clinical trials are critical for medical decision-making but are fraught with problems, including low enrollment, poor compliance and costly implementation. Translation of results into practice can take years. Low cost, efficient strategies are needed for conducting trials in order to accelerate implementation. For this project we will use as a prototype a computer-tailored intervention for urban teens with asthma in the clinic that has already shown success as a school-based program. A seamless multi-site, open-label, randomized phase II/III trial has been designed to evaluate the intervention using a pragmatic approach and a comparative effectiveness design. Computerized methods will be implemented for patient identification, consent, enrollment and tracking. A centralized database with remote data capture is planned for case report form completion with a highly integrated electronic medical record (EMR). Routine clinic visits will serve as trial data collection points. Adolescents, aged 15-19 years, diagnosed with asthma (n=354), will be randomized to control or intervention arms consisting of 4 weekly computer sessions a minimum of 1 week apart, followed by a 6 and 12 month follow-up survey. The phase II trial will include the first 250 patients with primary assessment of study feasibility and reliability of EMR data. The primary endpoint is asthma control, as determined by the Asthma Control Test. Secondary outcomes include asthma exacerbations, symptom-days, symptom-nights, days of restricted activity and school/work days missed. Two interim looks are proposed when 40% and 70% (Phase II) of patients are enrolled. Phase II of the trial is funded through NHLBI. We will seek additional funding for a Phase III component if there is an indication of treatment effect at interim analysis. Currently, we are in the planning phase, have received IRB approval, and will begin patient enrollment in March, 2013. Study results will inform strategies for program dissemination.

A17

REVISITING THE QUESTION: TO ADJUST OR NOT TO ADJUST FOR BASELINE COVARIATES IN RANDOMIZATION AND/OR IN ANALYSIS OF CLINICAL TRIALS?

(1) Yuko Palesch (2) Yanqiu Weng (3) Wenle Zhao

(all) Medical University of South Carolina

Baseline covariate adjustment in randomization and/or in analysis has yielded a vast literature in clinical trials. The contentious issues involve statistical operating characteristics (SOC), randomization performance metrics, and clinical interpretation. Through simulation using bootstrap technique on data from a moderate-sized clinical trial (N=400) with two treatment arms, we evaluate six scenarios that are combinations of each of the three randomization methods (simple (SR), stratified permuted block (SPB), and a recently developed minimal sufficient balance algorithm (MSB)) with each of the two analysis approaches (with and without adjustment of baseline covariates with strong prognostic values). The results show that, as expected, the type I error probability for the adjusted analysis is slightly conservative for both SPB (0.0493) and MSB (0.0456), and more so if the analysis is not adjusted for the covariates (0.0319 and 0.0305, respectively). Power is similar among the three randomization methods for adjusted analyses, but the unadjusted analyses yield substantially reduced power. For randomization metrics, SR and MSB are ideal in the probabilities of deterministic assignments (0) and correct guesses (0.5 and 0.54, respectively), but the MSB best controls the probability of significant imbalances in the covariates between the treatment arms. We conclude that SR
with covariate-adjusted analysis may be the best approach due to its SOC and simplicity of implementation, but with the caveat that it may be more vulnerable to challenges in trial result interpretations due to potential covariate imbalances between the treatment groups. MSB, also with covariate-adjusted analysis, may alleviate that concern while maintaining similar SOC. One should weigh the benefit in the SOC and randomization metrics against the effect of potential covariate imbalance on clinical interpretation. Additional scenarios (under different Ns and other SOC parameters) also will be presented.

A18

ESTIMATING HETEROGENEOUS TREATMENT EFFECTS FROM DATA GENERATED IN RANDOMIZED CONTROLLED TRIALS BASED ON THE PARALLEL GROUP DESIGN USING PREDICTIVE BASELINE COVARIATES

(1) Ruediger Paul Laubender (2) Ulrich Mansmann

(1) Institute of Medical Informatics, Biometry, and Epidemiology (IBE), Ludwig-Maximilians-University Munich
(2) Institute of Medical Informatics, Biometry, and Epidemiology (IBE), Ludwig-Maximilians-University Munich

One of the aims of ‘individualized medicine’ is to allocate the optimal therapy to a patient when several treatment options are available and when it is known that patients differently react to these treatments. The presence of heterogeneous treatment effects (also known as subject-treatment interactions) can be estimated by study designs which allow for replication of the treatment effects like the repeated crossover design (Senn 2001). However, randomized controlled trials (RCTs) allowing for replications are not possible when treatments irreversibly change the patients or where carryover is an issue (e.g. treatments used in oncology, psychiatry and infectiology). In such situations only RCTs based on the parallel group designs can be used. From a statistical point of view, we only observe the marginal distributions of the responses to the treatments of interest from RCTs based on the parallel group design but we cannot observe the joint distribution of these responses in contrast to designs allowing for replications. For normally distributed responses, it might be still possible to reconstruct the joint distribution and hence to identify heterogeneous treatment effects from a RCT based on the parallel group design by redefining a predictive baseline covariate as an indicator for the individual treatment effect (ITE) and by assuming a multivariate normal distribution. Given these assumptions we derive estimators for reconstructing the joint distribution, for the ITE and for the probability that an individual will profit from treatment A compared to treatment B given his or her value of the predictive covariate. This is an estimator easily understandable by physicians and patients and cannot be derived from a usual regression model with a treatment-covariate interaction. A simulation study was performed to judge the performance of the newly proposed estimators. Additionally, our methods are applied to a data examples used by Gadbury et al. (2000, 2001).

A19

LUMPING AND SPLITTING APPROACH TO ANALYZING HIGH DIMENSIONAL SPARSE COUNTS DATA

(1) Zhibao Mi (2) Joseph Collins

(1) VA Cooperative Studies Program (2) VA Cooperative Studies Program

It is not unusual for hundreds or even thousands of clinical endpoints to be collected from individual subjects to assess the safety or efficacy of a drug/device in clinical trials/studies, but events from the majority of clinical endpoints are rare. The challenge to analyzing high dimensional sparse data is how to balance analytical consideration for statistical inference and clinical interpretation with precise meaningful outcome of interest at intra-categorical and inter-categorical levels. Lumping or grouping similar clinical events into a composite category has the statistical advantage of increasing testing power, reducing multiplicity size, and avoiding competing risk problem; however, too much or inappropriate lumping would jeopardize the clinical meaning of interest and external validity. Whereas splitting or keeping each individual event at its basic clinical meaningful category can overcome the drawbacks of lumping, this practice may create analytical issues of increasing type II error, multiplicity size, and competing risks and having a large proportion of endpoints with rare events. It seems that lumping and splitting are diametrically opposite approaches, but in fact, they are complementary. Both are essential for high dimensional data analysis. In this report, a lumping method is proposed based on the correlations between a treatment/exposure and clinical endpoints.
and biological information to reduce data dimension. It is combined with a splitting method using Poisson and/or negative binomial models to adjust for over-dispersion and using zero inflated Poisson models for those endpoints with substantial number of zeros. The multiplicity is adjusted based on the results from lumping and splitting analyses by controlling the false discovery rate. The clinical relevance of the proposed analytical approach was tested and validated using an example with 372 safety binary endpoints from a VA double-blinded randomized multi-center clinical trial. The safety evaluation was based on the results of lumping and splitting analyses.

A20
EVIDENCE INFORMED MISSING DATA IMPUTATION FOR COGNITIVE FUNCTION TESTS WITHIN RCTS
(1) Janice Pogue, (2) Martin O'Donnell, (3) Salim Yusuf
(1) McMaster University, (2) McMaster University and Galway University Hospital, (3) McMaster University

In order to avoid bias in trial results, we need to ensure complete data collection of outcomes for all participants. Unfortunately in trials that measure cognitive function, some participants who have already experienced disability may be unable or unwilling to complete follow up assessments. These assessments are not missing at random, yet trials frequently use statistical models that assume they are. Many authors have suggested that we use the reason that the assessment is missing to select the analytic approach for trials (Simes, Stat Med 1998; Little, NEJM 2012; Schott, NEJM 2012), but few examples of this methodology exist within published trials. We suggest that reasons for missing data be systematically collected in a form that can be directly related to each trial participant’s probability of cognitive decline. Reasons for missing data are provided by the site in a format that can be directly related to the probability (confirmed, probable, possible, or unrelated) that each individual has experienced cognitive decline. This information, together with published norms for the scale in disabled populations, is then used within the method of multiple imputation to make evidence informed assumptions for each trial participant. We provide simulation data for a trial measuring cognitive decline using the Digital Substitution Scale and demonstrate biased trial results for the common methods that assume random missingness. The systematic collection of reasons for missing data can provide a scientific rationale for handling missing data to directly inform the imputation methods we use to analyze trial results, not as a sensitivity analysis, but within the primary models. This method should decrease the likelihood of biased trial results due to informative missing data.

A21
STANDARDIZING EARLY PHASE TRIAL DOCUMENTATION AND PROCESSES
(1) Debbie Sherratt (2) Sarah Flynn (3) Louise Flanagan (4) Sarah Brown
(5) Nina Thenagaran (6) Avie-Lee Coney
Clinical Trials Research Unit, University of Leeds

In 2009 Myeloma UK set out an innovative research model to rapidly and systematically address the critical challenges that are slowing down research, the development of, and access to new drugs, developing an Early Phase Clinical Trials Network. Subsequently, in November 2009, the Clinical Trials Research Unit (CTRU) at the University of Leeds successfully became the Myeloma UK (MUK) Clinical Trials Coordinating Office (CTCO) for these early phase clinical trials. In the three years the network has been running, we have developed a number of phase I and phase II processes, and associated documents, which may be transferable between early phase trials. To date we have set-up five early phase trials, 1 in follow-up, 1 in recruitment, and 3 in trial set-up - each building on the standardization of the early phase processes developed. Standardized processes and documentation that have been developed include: consenting and registering participants, case report forms, database specifications, and data reports, dose escalation documentation, trial endpoints and safety reporting. The differences in processes between phase I and phase II trials will be discussed as well as the advantages and challenges of standardizing processes in each phase and across the two phases. The standardization of processes within the MUK CTCO has improved the efficiency of the set up of trials, reduced the need for multiple document reviews and allows familiarity with processes and paperwork within the trial team. Furthermore, it is hoped that standardization will facilitate expedited center set up and aid
the conduct of trials by allowing centers to be familiar with the processes required and the data collection
documentation to be used.

A22
SOCIAL MEDIA IN CLINICAL TRIAL RECRUITMENT AND RETENTION
(1) Evan Hempel (2) Stephanie Millin
(1) KAI Research, Inc. an Altarum Company (2) KAI Research, Inc. an Altarum Company

One may argue study participants are the most important part of clinical research; without them there would
be no clinical trials. Sponsors spend a significant portion of study budgets on participant recruitment: develop-
ning flyers, radio/television advertisements and notices in local newspapers/newsletters; yet these costly
techniques are no guarantee of success. In the ever changing world of the internet and social media many
individuals turn to the Web and virtual networks for answers to medical concerns. KAI-Research, Inc. (KAI)
has firsthand experience demonstrating how social media can be an inexpensive and effective tool for
quickly meeting recruitment goals as well as improving study retention. In a recent multi-site neonatal
clinical trial, one hospital utilized their Facebook page to post information about the study. This site easily
enrolled twice as many infants as all the other sites combined and it cost $1,000 less per subject than the
second most successful site. The rapid enrollment enabled the trial to conclude six months ahead of sched-
ule, creating noteworthy cost savings. In a longitudinal study of young adults, some sites found participants
to be more responsive to a personal Facebook message then phone calls or e-mails in study follow-up. Thus,
Facebook greatly improved the trial’s retention rate. Using social media in clinical trials also has its chal-
genges. At present there are no FDA regulations or GCP guidelines to assist with implementing a social media
strategy while maintaining privacy and HIPAA compliance. IRBs differ in their opinion of how social media
should be used in trials. Being informed about the issues and concerns as well as taking steps to address
ethical considerations are essential in the successful implementation of social media in clinical trials.

A23
SYSTEMATIC REVIEW OF STRATEGIES TO REDUCE ATTRITION IN RANDOMISED TRIALS
(1) Valerie Brueton, (1) J. Tierney, (1) S. Stenning, (2) I. Nazareth, (1) S. Meredith, (3) S. Harding, (2) G. Rait
(1) UK Medical Research Council, Clinical Trials Unit (2) University College London, Research Department of
Primary Care and Population Health (3) UK Medical Research Council, Social and Public Health Sciences Unit

Background: Attrition from randomised controlled trials (RCTs) can introduce bias and reduce study power
affecting the generalisability, validity, and reliability of results. Objectives To quantify the effect of strategies
to reduce attrition from RCTs.

Methods: Eligible trials were randomised evaluations of strategies to reduce attrition embedded within RCTs.
We searched bibliographic databases, trial registers, Society for Clinical Trials meeting abstracts and trial
report and review reference lists. We also surveyed UK clinical trials units. Heterogeneity in strategies was
explored with subgroup analyses, and the fixed-effect model used to pool results.

Results: 39 RCTs were eligible with data available from 37 of these. Six broad types of strategies to reduce
attrition were evaluated. Most aimed to improve postal or electronic questionnaire response. Response was
increased by: adding a monetary incentive (RR 1.18; 95% CI 1.09-1.28), higher value incentives (RR 1.12;1.04-
1.22), offering monetary incentives (RR 1.25;1.14-1.38), shorter questionnaires (RR 1.04;1.00-1.08) and dis-
ease-relevant questionnaires (RR 1.07;1.01-1.14). Based on results of single trials, recorded delivery of
questionnaires (RR 2.08;1.11-3.87); a “package” of postal strategies (RR 1.43;1.22-1.67); open trial design (RR
1.37;1.16 -1.63) increase response/retention and, but interviews were inferior to postal questionnaires
(RR=0.90;0.88-0.92). There is no clear evidence that questionnaire response/participant retention is improved
by the following: a charity donation, adding or offering gifts, “enhanced” letters, priority post, extra remind-
ers to participants, questionnaire question order, site reminders, sending questionnaires early, long and clear
questionnaires, behavioural or case management strategies, or that monetary incentives are better than an
offer of entry to a prize draw, or telephone surveys are better than a questionnaire plus monetary incentive.
Conclusion Offering and giving monetary incentives and short and relevant questionnaires improve
response. Other promising strategies need further evaluation. Application of these results depends on con-
text and existing trial follow-up procedures.
A24

Q-QAT (QUANTI-QUALITATIVE APPOINTMENT TIMING): A SIMPLE TECHNIQUE TO ELUCIDATE KEY RECRUITMENT ISSUES IN RCTS

Jenny Donovan, Sangeetha Paramasivan, Sean Strong, Caroline Wilson
University of Bristol

Background: Recruitment to pragmatic randomized controlled trials (RCTs) is acknowledged to be difficult, and very few interventions to improve recruitment have proved to be effective. We present a simple technique employing mixed research methods that can identify key recruitment issues in RCTs that can then be fed-back as part of a complex intervention to improve RCT recruitment.

Method: The outline of the technique emerged from a programme of qualitative research undertaken to investigate the process of recruitment in six RCTs with recruitment difficulties. Recruitment appointments were audio-recorded and transcribed, and a simple qualitative content analysis was performed. The Q-QAT technique involved coding recruitment appointment transcripts for time devoted to explaining each of the RCT arms and the RCT itself. The technique was applied to two RCTs with very different clinical contexts, organizational issues and recruitment processes to refine its development. Comparisons of Q-QAT data were made between clinical centres and individual recruiters. Findings were fed-back to recruiters and recruitment rates re-assessed.

Results: The Q-QAT technique was applied to two RCTs, and coding methods were adapted to take into account the differences in design and organization. In both RCTs there were considerable variations in the quantity of time spent by different recruiters on the RCT arms and RCT itself. In RCT1, Q-QAT data showed that similar proportions of time were devoted to explaining details of each treatment arm (supporting qualitative findings that the arms were presented in a balanced way), but very little time was spent explaining the RCT itself. In RCT2, treatments were presented very differently by specialists in different centres. Findings were presented to recruiters and discussed. Subsequent changes in Q-QAT patterns, qualitative findings and recruitment rates were seen.

Conclusion: The Q-QAT method is relatively simple to apply and appears to offer opportunities to encourage improvement in recruitment practice.

A25

THE USE OF SITE ENGAGEMENT TECHNOLOGY (SET) TO INCREASE PATIENT RECRUITMENT AND RETENTION IN CLINICAL TRIALS: A CASE STUDY

Vivek Murthy
Brigham and Womens’ Hospital/Harvard Medical School

Site engagement technologies (SETs) represent a novel approach to solving a critical problem in clinical trials: low recruitment and retention of patients. This problem has implications for cost, study validity, and the speed by which new therapeutics are made available to patients. The technology we have developed enables sponsors and/or CROs to motivate and engage study sites to improve recruitment and retention of patients - and the overall success of the clinical trial. In contrast to traditional trial investigator meetings (IMs), which connect investigators and the sponsor for a limited period of time, site engagement technology generates a sense of community and support from trial activation to closeout. SET tools focus on four key areas of the investigator site-sponsor/CRO interaction: education (on-demand video training and e-learning strategies such as quizzes), motivation (leader boards, peer and sponsor recognition awards, site ranking for milestones), communication (customized email, kudos recognition features, document change notifications), and documentation (repository for memos, FAQ, training forms, protocol, amendments, and more). Additionally, our cloud-based innovation platform allows trials to create specialized tools that can increase the rate of protocol adherence (such as the patient visitation guide and decision tracking features). We intend to highlight multiple case studies in which industry-sponsored trials implemented site engagement technologies. The results demonstrate increased rates of patient enrollment and retention and a >90% rate of technology adoption by study sites.

A26

STATISTICAL MODELING OF RECRUITMENT FOR MULTI-SITE CLINICAL TRIALS – CURRENT STATUS

(1) Kousick Biswas (2) Tyson Holmes
(1) Cooperative Studies Program, US Department of Veterans Affairs, (2) Dept. of Psychiatry and Behavioral Sciences, Stanford University School of Medicine

VA Cooperative Studies Program (CSP) conducts multi-site, randomized clinical trials using the VA medical centers around the country as the participating sites. In general, a typical CSP conducted clinical trial involves
10 to 200 medical centers as sites to recruit patients in the range of 200 – 10000. So, timely recruitment is a very important factor for timely completion of these studies within the allocated budget. This has been a chronic issue for the CSP studies. During the planning stages, as a normal practice, the length of the study and required budget are estimated based on an assumed fixed or temporally invariant recruitment rate throughout the recruitment period. From experience it is evident that recruitment rate in a multi-site clinical trial is not fixed or temporally invariant so that use of this simple model causes errors in prediction of study length. This paper provides a summary of currently used modeling techniques that account for the various uncertainties in rate of recruitment during study conduct to predict realistic recruitment trajectories.

A27
DATA MONITORING COMMITTEES IN CHINA
(1) Tingting Li (2) Kent Koprowicz (3) Xiaohui Huang (4) Ping Xu
(1) Axio Research, LLC. (2) Axio Research, LLC. (3) Axio Research, LLC. (4) Axio Research, LLC

Data monitoring Committees (DMCs), also known as Data and Safety Monitoring Boards (DSMBs) are commonly used in clinical trials where major health outcomes are expected (e.g. death, progression of disease). DMCs periodically review interim data while a trial is ongoing in an effort to ensure the ethical conduct and scientific integrity of the study. DMCs have been an important component of many government and industry sponsored US and European clinical trials for decades. While there has been a recent large increase of clinical research in China and other Asian countries, the use of DMCs remains novel. In the past year we have served as the independent statistical group for several DMCs monitoring oncology trials conducted mainly in China. In this talk we share our experience of working with China-based DMCs. In particular we note the characteristics and challenges for data monitoring in China in regard to constituting and implementing government guidelines, educating local DMC members on responsibilities and process, and meeting planning and logistics.

A28
MORBIDITY ADJUDICATION AND TRACKING DURING DCCT/EDIC
Jye-Yu C. Backlund, Stephan Villavicencio, Wanyu Hsu, Lisa Diminick, Patricia A. Cleary
and the DCCT/EDIC research group
The George Washington University, The Biostatistics Center

Concerns about the cardiovascular (CV) events for the treatment of diabetes have served to emphasize the need for accurate assessments of CV primary endpoints in clinical trials. The Epidemiology of Diabetes Interventions and Complications (EDIC), the observational follow-up of the Diabetes Control and Complications Trial (DCCT) cohort, established a surveillance and adjudication process to classify major CV outcomes in 1994. During annual visits, EDIC subjects were asked to report any CV events and the study sites notified the coordinating center of all reported events. Events were adjudicated by the Mortality and Morbidity Review Committee (MMRC), which was masked to DCCT treatment group and glycosylated hemoglobin levels. Events were adjudicated using medical records, ECG findings, and cardiac enzyme levels. Among 670 events reported between 1994-2012, 66 (9.9%) were non-fatal myocardial infarction (MI), 158 (23.6%) revascularization, 22 (3.3%) stroke, 8 (1.2%) congestive heart failure (CHF), 197 (29.4%) angina, 192 (28.7%) arrhythmia, and 27 (4.0%) transient ischemic attack (TIA). Only 280 (41.8%) of the 670 reported events were classified as meeting adjudication criteria, while 125 (18.7%) were not. Another 247 (36.9%) reported events could not be adjudicated due to insufficient documentation. As shown in the Figure, completed adjudication was highest among reported MI, revascularization, and CHF (all > 60%) and lowest among angina, arrhythmia and TIA (all

A29
HOW TO GET THE MOST OUT OF YOUR COURTSHIP WITH FDA - AN FDA MEDICAL DEVICE REVIEWER'S PERSPECTIVE
Phyllis M. Silverman
Center for Devices and Radiological Health, FDA Office of Surveillance and Biometrics Division of Biostatistics

Efficient and effective communication between FDA and Industry for medical device approval or clearance is crucial at all points along the application timeline. This talk will explore the different options available to
Industry in interacting with FDA both before and during the review process, including the different types of meetings and their purpose. It will also present specific elements of study design and statistical analyses that industry sponsors should be prepared to discuss and provide answers to both at meetings and by way of formal written review correspondence. The top 10 communication “crimes” will be revealed.

A30

HAS THE VOLUNTARY HARMONISATION PROCEDURE BEEN A SUCCESSFUL ATTEMPT TO STREAMLINE THE APPROVAL PROCESS FOR MULTINATIONAL CLINICAL TRIALS IN THE EU?

EXAMINING APPLICANTS’ EXPERIENCE TO DATE WITH THE GTFG’S INITIATIVE

David Smith
Verius Limited

Contrary to the intentions of the Clinical Trials Directive, divergent practices persist between member states reviewing applications for clinical trials in the EU. A complex network of diverging requirements and practices presents a complicated and unfavourable approval process, particularly for multinational clinical trials (MN-CTs). It is these issues that the Clinical Trials Facilitation Group was targeting in 2009 when it introduced a voluntary harmonisation procedure (VHP) for MN-CTs. Three years on, use of the VHP is far from widespread and published information relating to applicants’ experience with the procedure is negligible. The purpose of this research was to determine the level of success that the VHP has achieved in fulfilling its objectives, by targeting those who have used the procedure. An online questionnaire targeted at MN-CT applicants was distributed to regulatory professionals, and the responses obtained highlight the positive aspects of the VHP, but also the need for improvements. Applicants’ experience with the VHP has been generally positive, with a low occurrence of divergent decisions, and a reduced regulatory burden associated with the harmonised approval process. Timelines for clinical trial approval have not been markedly reduced, however, and significant issues remain that are likely to continue to dissuade prospective applicants from utilising the VHP. The VHP is clearly a step in the right direction, and applicants’ experience with the procedure hints at the likely impact of a Regulation on clinical trials, as proposed by the European Commission in July 2012. The results of this research suggest that the revised legal framework (expected in 2016) will have a positive impact on the clinical research environment in the EU, with applicants first in line to benefit through a faster, more flexible and fully harmonised approval process for clinical trials.

A31

MINIMIZING RECONCILIATION BETWEEN SAFETY AND CLINICAL DATABASES

Sean Neal
Medidata Solutions

With patient safety in mind, global regulatory agencies enforce stringent guidelines for timely reporting of safety events by investigator sites and sponsors. This process has been estimated at ~2.15 FTE days/SAE just in sponsor effort and does not even take into account the effort required by site personnel. Estimates from the Tufts Center for the Study of Drug Development suggest study coordinators spend between 10 and 20% of their total work effort on trials just on AE/SAE reporting. Besides consuming valuable safety resources in re-keying data that is already captured in EDC into the safety system, the reconciliation of the data is a huge burden. By way of example, an analysis of just SAE rates from completed studies reporting results over the past 5 years, 2007-2011, from Clinicaltrial.gov (n=4,221) demonstrates the magnitude of the reconciliation effort that would be needed. Overall while the majority of trials, 59%, exhibit SAE rates of 10% or less, 10% of trials have SAE rates between 10-and 20%, 17% of trials have SAE rates between 20 and 50%, and 6% of trials exhibited SAE rates of greater than 50%. As regulatory environment continuous to evolve, the new pharmacovigilance (PV) legislation that takes effect across the European Union (EU) in July 2012 further challenges the manual, paper-based practices. The major overhaul of post-marketing expedited reporting requires Marketing Authorization Holders (MAHs) starts reporting not just SAEs but also suspected non-serious adverse reactions within 90 days of the legislation’s commencement. This presentation will discuss how new technologies in clinical trials help researchers to minimize or fully eliminate such reconciliation between two databases, thus streamlining safety reporting process, improving regulation compliance and enabling safety team focus on what they do best - triage case and detect safety signals as early as possible.
A32

SYSTEMATIC REVIEWS OR RELEVANT RANDOMISED CONTROL TRIALS, WHICH SHOULD I BELIEVE? A DILEMMA FOR POLICY MAKERS

Martin Tickle

School of Dentistry, University of Manchester

Systematic reviews sit at the top of the evidence-based hierarchy but their findings can sometimes be at odds to the results of individual trials. This lack of agreement can confuse policy-makers when deciding how best to deploy resources.

Objectives: To compare the evidence presented by a systematic review and a large pragmatic trial and make suggestions for how policy-makers should approach evidence-based decision-making.

Methods: A case study that compares the findings of a cluster RCT of fluoride varnish delivered as a public health intervention in schools with those of a Cochrane systematic review of fluoride varnish.

Results: The systematic review reported that a 46% preventive fraction for dental caries resulted from two or more fluoride varnish applications a year. The cluster RCT randomised 190 schools into test and control groups and 1473 (test) and 1494 (control) 7-8-year-old children participated in the trial. Children in the test group were offered 3 applications of fluoride varnish each year over a 3-year follow-up period; >60% of children received all 9 applications and >90% received >6 applications. The trial could find no statistically or clinically significant difference between test and control groups. Reasons for the differences in the findings between the review and the RCT are identified and discussed, these include fluoride exposure and caries risk in different historical and geographical populations; the setting in which trials are conducted; difference between pragmatic and effectiveness trials; quality of older trials and the potential for publication bias.

Conclusions: Systematic reviews should be frequently updated and where possible sub-group analysis should be undertaken according to population risk and trial design. Policy makers need access to relevant methodological expertise to support their decision-making. A recent, adequately powered, well-designed and well-conducted, pragmatic trial conducted in a relevant population is more useful to policy-makers than a systematic review.

A33

SWITCHING FROM NON-INFERIORITY TO SUPERIORITY IN AN ADAPTIVE SETTING: SOME NOTES

(1) Dario Gregori, (2) Giada Morpurgo, (3) Federica Zobec

(1) University of Padova, (2) ZETA Research Srl

In the presence of competing primary targets, it is customary to first determine the sample size w.r.t the primary endpoint and then make a power analysis of the remaining target, to make sure that enough power analysis is warranted. EMEA and FDA allow both switching from non-inferiority to superiority in both traditional and adaptive designs. However, the issue of the case highlighted above is not well addressed in literature, i.e.: how we should re-estimate sample size after each interim analysis when both primary and non-primary endpoints are evaluated. Ours paper aimed at showing practical consequences in the design of a safety device when either non inferiority and superiority should be verified in order to move further on to next step.

A34

USING ADAPTIVE DESIGNS TO BUILD A BEHAVIOURAL CHANGE INTERVENTION IN A COMPLEX INTERVENTION RANDOMISED CONTROLLED TRIAL

John Norrie, Jill Francis, Heather Waterman, Jen Burr, Ruth Thomas, Mark Pilling

University of Aberdeen, City University London, University of St. Andrews, University of Manchester

It is routinely challenging picking the optimal intervention when designing complex intervention trials changing behaviours to improve patient reported outcomes, from many different interventions with competing evidence of varying quality and substance, backed by different underlying theories of behaviour change. To separately trial each one isn’t usually feasible. To select one or two to take forward to trial runs high risk of missing the most effective. And what works for some may not work for others. Although adaptive designs are well developed for drug trials, they are less common in non drug complex intervention...
settings. We were faced with building the optimal intervention to increase compliance (adherence) with eye drops for managing intra ocular pressure (IOP) in glaucoma. This talk will review the design for an explanatory trial using a mixture of two adaptive type designs to find the best package of measures to change behaviour and improve compliance with the eye drops, which is often poor. The design mixes a standard Adaptive Experimental Design (AED) to build the best combination of 5 generic interventions (including educational information, self efficacy, and barriers to change), delivered either individually or in groups, with real time assessment of effectiveness via a study web portal leading to rapidly evolving stages of randomisation, using a play the winner algorithm, to identify at a population level what intervention works for most people. Alongside this, we integrate a Sequential Multistage Adaptive Randomised Trial (SMART) design to pick up those participants still non compliant and then identify individualised interventions that work for them. This novel application of two adaptive designs promises to provide a useful tool to improve trial design in challenging areas such as behavioural change, and so increase the likelihood of finding the best interventions, uncovering benefit at both population and individual level.

A35
DESIGN OF SEQUENTIALLY RANDOMIZED TRIALS FOR TESTING ADAPTIVE TREATMENT STRATEGIES
Semhar Oghagaber, Abdus Wahed
University of Pittsburgh Department of Biostatistics

An adaptive treatment strategy (ATS) is an outcome-guided treatment algorithm approach that allows personalized treatment of complex diseases based on patients’ need, disease and treatment history. Complex diseases such as AIDS, depression, cancer involve several stages of treatment due to multifaceted disease paths. Sequential multiple assignment randomized (SMAR) designs are generally used to draw simultaneous inference about multiple ATS’s, where patients are randomized to available treatment options at different stages as they become eligible to receive those treatments. However, design issues such as sample size and power for such designs have not been addressed adequately in the literature. In this article we propose a sample size formula based on Wald-type statistic for comparing multiple ATS’s based on a continuous endpoint through a SMAR trial. We show via simulation that the proposed sample size formula maintains the nominal power and hence provides a useful tool for practitioners to design a SMAR trial ensuring adequate power.

A36
ADAPTIVE ADJUSTMENT OF THE RANDOMIZATION RATIO USING HISTORICAL CONTROL DATA
(1) Brian P. Hobbs (2) Bradley P. Carlin (3) Daniel J. Sargent
(1) Department of Biostatistics, The University of Texas MD Anderson Cancer Center (2) Division of Biostatistics, University of Minnesota (3) Department of Health Sciences, Mayo Clinic

Prospective trial design often occurs in the presence of “acceptable” (Pocock, 1976) historical control data. Typically this data is only utilized for treatment comparison in a posteriori, retrospective analysis to estimate population-averaged effects. We propose an adaptive trial design in the context of an actual randomized controlled colorectal cancer trial. The proposed trial implements an adaptive randomization procedure for allocating patients aimed at balancing total information (concurrent and historical) among the study arms. This is accomplished by assigning more patients to receive the novel therapy in the absence of strong evidence for heterogeneity among the concurrent and historical controls. Allocation probabilities adapt as a function of the effective historical sample size (EHSS), which characterizes the relative informativeness defined in the context of a piecewise exponential model for evaluating time to disease progression. A Bayesian hierarchical model is used to assess historical and concurrent heterogeneity at interim analyses and to borrow strength from the historical data in the final analysis. Balancing total information with the adaptive randomization procedure leads to trials that on average assign more new patients to the novel treatment when the historical controls are unbiased or slightly biased compared to the concurrent controls. Large magnitudes of bias lead to approximately equal allocation of patients among the treatment arms. Using the proposed hierarchical model to borrow strength from the historical data, after balancing total information with the adaptive randomization procedure, provides preposterior admissible estimators of the novel treatment effect with desirable bias-variance trade-offs.
A37
IMPLEMENTING UNBALANCED ASSIGNMENT BY MINIMISATION
(1) Colin Everett (2) Samantha Hinsley (3) Sarah Brown
(1) Clinical Trials Research Unit, University of Leeds, UK. (2) Clinical Trials Research Unit, University of Leeds, UK. (3) Clinical Trials Research Unit, University of Leeds, UK

Minimisation, a dynamic method of assigning patients to a study group in a randomised controlled trial, ensuring the treatment groups are similar with respect to a number of specified factors, has often been used for trials where an equal balance between one or more groups is desired. Although equal allocation between two or more groups is common and efficient from a statistical view, there are good reasons for desiring an unequal allocation: one or more interventions may be considerably more expensive, have more significant associated risks, be more inconvenient - leading to larger numbers of patients withdrawing, or may be more limited in its supply or those trained to deliver it. Unbalanced assignment can be easily achieved using simple randomisation and through permuted block randomisation, but the method for achieving unbalanced assignment through minimisation may not be so straightforward. Motivated by the desire to use minimisation to assign patients to interventions with unequal allocation in our trials, we review the methods proposed in the literature for assigning patients to their groups by minimisation which aims to achieve unbalanced assignment. We briefly lay out the underlying methods of the different techniques proposed, and discuss what is involved, along with the challenges faced, when putting these methods into practice, including the additional work undertaken when testing what may be a novel process. Using simulation, we consider whether or not the targeted assignment ratio is reached in both the short and long term, the chances of predicting the next assignment, and impact of using and varying the random assignment probability. We also investigate the effect of changing the number of minimisation factors and their corresponding levels on the reliability of the methods, including looking at the effect of requiring the preservation of the assignment ratio within each centre of a multi-centre trial.

A38
A LIKELIHOOD-BASED APPROACH TO SELECTING DOSES WITH ACCEPTABLE TOXICITY IN A
STANDARD ALGORITHM-BASED PHASE I CANCER TRIAL
(1) Cody Chiuzan (2) Elizabeth Garrett-Mayer
Medical University of South Carolina Hollings Cancer Center

In oncology, the standard ‘3+3’ design has passed the test of time and survived various proposals to adjust the sample size, or other dose-escalation dynamics. The objective of algorithm-based designs is to identify the maximum tolerated dose (MTD), assuming that both efficacy and toxicity increase with dose. While this assumption is reasonable for cytotoxic agents, new cancer treatments such as immunotherapies may not follow this principle. For such treatments, where it is anticipated that lower doses may be as efficacious as higher doses, the objective of the phase I trial might instead be to identify doses with acceptable toxicity, which can be moved forward for further testing. We offer a likelihood-based approach for identifying doses with acceptable toxicity consistent with the decision rules used in the standard algorithm. Without incorporating any prior beliefs, our proposed method is based on the evidential paradigm (Royall 1997) that uses likelihood ratios to measure the strength of statistical evidence for one hypothesis over the other. Once toxicity responses at a certain dose level have been observed, the likelihood-ratio (LR) under the null and alternative hypotheses is calculated and compared to a certain k threshold (level of evidence). Given a true toxicity scenario, the following statistical characteristics are computed for different choices of k: i) probability of weak evidence, ii) probability of favoring the null under the null (analogous to 1-alpha), iii) probability of favoring the alternative under the alternative (analogous to 1-beta), and iv) probability of non-escalation under the null. The k value that displays the best operating characteristics is chosen to be the threshold for the likelihood-ratio. If ln(LR) > ln(k), then the dose is considered acceptably safe; otherwise the dose is thought to be unacceptably toxic. Simulation results are presented for different combinations of testing hypotheses, k thresholds, and various true toxicity rates.
A39

DESIGN AND ANALYSIS OF CLUSTER RANDOMIZED CROSSOVER TRIALS WITH BINARY OUTCOMES WITH APPLICATION TO INTENSIVE CARE RESEARCH
(1) Andrew Forbes (2) Muhammad Akram (3) Rinaldo Bellomo (4) Michael Bailey
(1) Monash University, (2) Monash University, (3) Monash University and the Australia and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation, (4) Monash University and the Australia and New Zealand Intensive Care Society Centre for Outcome and Resource Evaluation

There are a number of near-universal interventions applied in the intensive care unit (ICU) setting that may have an effect on mortality, such as intravenous caloric delivery in patients receiving nutrition or oxygen level targeting in mechanically ventilated patients. However, assessment of the effects of these interventions, which are likely to be small (eg 2% absolute reduction in mortality), are hindered by the need for the interventions to be applied at intensive care unit level together with the limited number of intensive care units in Australia. As a result, parallel-arm cluster randomized trials to detect small intervention effects cannot be sized appropriately. Efficiency can be gained by incorporating treatment crossover into the design of these trials, leading to so-called cluster randomized crossover trials. However, the development and assessment of these designs in the literature has been limited, and has been primarily focussed on continuous outcomes with associated linear mixed models rather than for binary outcomes. In this presentation we report on our recent analytical and simulation work evaluating estimation methods for cluster randomized crossover trials with binary outcomes. Using an underlying marginal model, we provide modifications of sample size formulae to accommodate binary outcomes and period effects, and variance expressions incorporating unbalanced sample sizes across clusters and across periods within clusters. We report on the adequacy of these expressions using a simulation study based on the cluster size variability and within- and between-period correlations observed in the Australian adult patient intensive care database. Our results indicate that cluster summary and marginal model estimators provide generally appropriate size, power and confidence interval coverage, however marked differences in efficiency occur in certain settings, and we provide rationale for these differences. We also discuss the potential for extension to multi-period designs and their feasibility in the intensive care research setting.

A40

A NEW PARADIGM FOR LARGE, SIMPLE TRIALS BASED ON ELECTRONIC HEALTH RECORDS
KyungMann Kim
University of Wisconsin-Madison

In March 2012, the National Institutes of Health issued a funding opportunity announcement, RFA-RM-12-002, “NIH Health Care Systems Research Collaboratory - Pragmatic Clinical Trials Demonstration Projects (UH2/UH3),” under a Common Fund (Roadmap) initiative and then in September 2012 announced its funding decisions via NIH News with the caption “NIH funds will strengthen national capacity for cost-effective, large-scale clinical studies.” Independently of this initiative, the Institute of Medicine of the National Academies hosted a workshop entitled “Large Simple Trials and Knowledge Generation in a Learning Health System” during November 26-27, 2012. Both these initiatives were made possible thanks to the investment in EHRs by the health care systems and to the commitment made by the US Federal Government through the Health Information Technology for Economic and Clinical Health Act of 2009. Besides improving patient care, EHRs hold much promise of making the conduct of large-scale, pragmatic clinical trials very cost-efficient and even clinical effectiveness research possible in the health care systems. However, it poses regulatory, human subject and information technology challenges to the clinical trials community. In this presentation, I will discuss promise and challenges of conducting pragmatic, large, simple trials and clinical effectiveness research that rely on access to EHRs for obtaining patient follow-up and outcome data, rather than on traditional protocol-specified patient visits and procedures.
PATIENT AND PROVIDER ACCEPTANCE OF A PRAGMATIC POINT-OF-CARE PILOT STUDY

(1) Ryan Ferguson (2) Patricia Woods (3) Katherine Riley (4) John Hermos (5) Thomas Sabin (6) Mary Brophy (7) Robert Lew (8) Leonard D’Avolio (9) Louis Fiore (10) Philip Lavoir

(1) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System, Department of Epidemiology Boston University School of Public Health (2) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System (3) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System (4) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System, Department of Medicine Boston University School of Medicine (5) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System (6) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System, Department of Medicine Boston University School of Medicine (7) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System, Department of Biostatistics Boston University School of Public Health (8) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System (9) VA Cooperative Studies Program Coordinating Center VA Boston Healthcare System, Department of Medicine Boston University School of Medicine (10) Department of Health Research and Policy Stanford University

The Point of Care (POC) initiative within the Department of Veterans Affairs (VA) Cooperative Studies Program is designed to conduct randomized comparative effectiveness clinical trials during the delivery of routine clinical care. The objective of the program is to create a learning healthcare system that uses pragmatic trials and the national VA electronic medical record (EMR) infrastructure to answer questions that are vitally important to the health and care of the nation’s veterans. A pilot study comparing inpatient insulin regimens was conducted to assess the feasibility of the initiative at two VA Medical Centers. Medical house staff were offered the option to randomize the insulin regimen at the time of placing insulin orders in the EMR insulin order menu modified for the study. Clinicians declined the randomization option in only 19% of patients eligible for the study. Of the 119 eligible patients approached, 102 (85.7%) agreed to randomization, 8 (6.7%) declined participation, 9 (7.6%) were not enrolled due to use of prohibited medications or because consent was not possible within 24 hours of referral to the study. To date, there have been no safety events and no losses to follow-up. The refusal rate of 6.7% of patients is markedly lower (3 times lower) than the rate experienced in the traditional clinical trials conducted within this Cooperative Studies Program Coordinating Center. Early results of this pilot study demonstrate the feasibility and acceptance by both the patients and the providers in integrating the pragmatic design of the POC trial into clinical care with no added study related procedures for clinicians and patients. The result has created an environment where the research infrastructure and the clinical apparatus of the VA are operating synchronously, thereby reducing barriers to participation in studies that are typically seen in traditional trials.

CONTINUOUS SAFETY MONITORING FOR RANDOMIZED CONTROLLED CLINICAL TRIALS WITH BLINDED TREATMENT INFORMATION

Greg Ball
Astellas Pharma Global Development

Data monitoring committees (DMCs) evaluate accumulating unblinded data and make recommendations about the continuing safe conduct of a trial. However, it is the trial leadership who make the ethical decisions about stopping a trial and they could benefit from objective statistical rules that help them judge the strength of evidence contained in the blinded data. We design safety signals for randomized controlled clinical trials with blinded treatment information, which could be used by anyone, including trial leadership. A Bayesian framework, with emphasis on the likelihood function, is used to allow for continuous monitoring, using all of the available information, without adjusting for multiple comparisons. Close collaboration between Statistics, Medical Science and Pharmacovigilance will be needed in order to design safety signals with good operating characteristics. It is imperative that emerging trends in the data are frequently and carefully evaluated to protect the safety and welfare of patients. While investigators must conscientiously maintain their blind, DMCs must have unblinded access to all of the available information, so that together they can fully protect patients from unsafe treatments, without compromising trial...
integrity or otherwise interfering with the strength of evidence. Safety signals will provide a full measure of the blinded data that the trial leadership can use, in combination with open information from the DMC, to evaluate the strength of evidence in the available data in order to make fully informed decisions that protect patients from unnecessary harm while allowing the trials to lead to conclusive results. Trial leadership need these safety signals to help them effectively monitor the ongoing safe conduct of clinical trials with blinded data.

A43

PRECIS-2: A TOOL TO IMPROVE THE APPLICABILITY OF RANDOMISED CONTROLLED TRIALS
Shaun Treweek, Kirsty Loudon, Frank Sullivan, Merrick Zwarenstein
University of Aberdeen; University of Dundee; University of Dundee; Western University

Background: RCTs are the best design to evaluate the effect of different interventions but randomisation does not itself promote the applicability of the results to situations other than the one in which the trial was done. PRECIS is a tool that helps trialists consider the effects of their design decisions on applicability.

Aim: To produce an improved and validated version of PRECIS.

Methods: The study has three phases. Phase 1 involves brainstorming and a two-round Delphi survey of authors who cited PRECIS. The Delphi results will then be discussed and alternative versions of PRECIS-2 developed and user-tested by experienced trialists. Phase 3 will evaluate the validity and reliability of the most promising PRECIS-2 candidate using a sample of 15-20 trials rated by 15 other trialists.

Results: Brainstorming sessions identified the PRECIS presentation (a wheel), lack of a scoring system and domain weighting as issues for exploration in the Delphi. Ninety individuals were invited to complete round 1 of the Delphi; completed responses were received from 34. Many respondents wanted a scoring system, were unhappy about presentation format and suggested several additional domains for PRECIS-2. Based on these responses, round 2 of the Delphi tested specific scoring alternatives, tabular vs a wheel presentation and new domains. Early responses to round 2 support using both wheel and tabulated forms of PRECIS, several new domains (eg. recruitment setting), 1-7 scoring and show continued uncertainty over weighting. The next stage will use the Delphi suggestions to create alternative versions of PRECIS-2 for user-testing in spring 2013.

Conclusions: We have concrete suggestions for improving PRECIS and a growing list of enthusiastic individuals interested in contributing to this work. By the end of 2013 we expect to have a validated PRECIS-2.

A44

DESIGNING A SERIES OF HYBRID TRIALS WITH CONTINUOUS OUTCOMES
(1) Siew Wan Hee (2) Nigel Stallard
(1) Warwick Medical School, University of Warwick (2) Warwick Medical School, University of Warwick

Background: Clinical trials are commonly designed individually. The number of patients available and other resources required to conduct the trial are essentially finite. A decision concerning the design of one trial will therefore inevitably affect the design and allocation of resources of subsequent trials.

Purpose: To design a trial in the context of other trials so that the clinical development plan is more cohesive.

Method: The development plan is considered to consist of a series of independent trials. The responses observed from each trial are analysed using the classical frequentist analysis. Thus, in the design stage the type I error rate is fixed but the unknown parameter of the treatment efficacy is assumed to be random and follows a prior distribution. A sample size that optimizes the Bayesian expected power (assurance) is identified as the optimal sample size. The design is then extended to include the start-up cost spent on planning, designing and so on.

Results: If the costs of conducting a trial are ignored, the optimal strategy is to make trials as small as possible. This is unsurprising as it maximises the number of trials that can be conducted and consequently the expected total number of trials that reject the null hypothesis. Taking costs into account, there is a range of optimal solutions for different combinations of cost, parameter of likelihood function and a priori knowledge of the unknown parameter. As expected, if the cost is very low, the optimal sample size is small but if the cost is on the other extreme, that is, very high, the optimal strategy may be not to conduct any trial at all.
A45

INCORPORATING EMERGING THERAPIES INTO ONGOING RANDOMISED CLINICAL TRIALS
(1) Dena Cohen (2) Walter Gregory (3) Sue Todd (4) Julia Brown

(1) Clinical Trials Research Unit, University of Leeds (2) Clinical Trials Research Unit, University of Leeds (3) Department of Mathematics and Statistics, University of Reading (4) Clinical Trials Research Unit, University of Leeds

The incorporation of an emerging therapy as a new randomisation arm to be included in a clinical trial that is already open to recruitment is a challenging question, both practically and statistically, yet is increasingly a reality in studies. It may take many years to run a clinical trial from concept to reporting within a rapidly changing drug development environment. In order for trials to be most useful to inform policy and practice it is advantageous for them to be able to adapt to the changing environment. Currently there is only scarce literature describing statistical methodology or practical considerations when adapting a trial by adding a new treatment midway through recruitment. Yet, it is desirable to researchers, regulators and patients to allow this adaptation within a clinical trial to reduce overall time and costs for determining optimal therapies compared to running separate trials, facilitate selection of the best of the emerging treatments and keep the outcomes current. It is essential to ensure that the research integrity is not compromised when adapting a trial in this way, because otherwise the outcomes may be uncertain, which would be unethical and an unacceptable waste of resources. This presentation will summarise the findings of a systematic literature review investigating what methodology has already been considered for this scenario, and identifying any trials where this amendment has been implemented. The statistical or design issues that were faced and the validity of the trial outcomes following the amendment will be reported. Considerations for future research as identified from the literature and initial research stages will be discussed. These will include conservation of error, appropriate use of control data from each stage, optimal allocation ratios for randomisation and plausible ranges of key parameters such as time to amendment.

A46

RANDOMIZED PHASE II TRIALS: ARE THEY ENTERING RESEARCH PRACTICE?
Catherine Klersy, Luigia Scudeller, Annalisa De Silvestri, Cristina Fiocchi*, Carmine Tinelli
Service of Biometry and Statistics & *Ethical Board Scientific Secretariat, Research Department. IRCCS Fondazione Policlinico San Matteo

Phase II trials should be viewed as proof-of-concept trials with the aim of ruling-out/ruling-in novel treatments (drug or medical devices), for further consideration of their efficacy in phase III studies. Historically, phase II trials have been based on the assessment of treatment efficacy in a one-arm design, as compared to historical controls. More recently, randomized phase II trials have been advocated to reduce the bias of using non concomitant controls. They have been considered for the analysis either as embedded one-sample phase II designs within each treatment arm, or as comparative designs with less stringent criteria for type I error than in phase III studies. Several reviews of the pertinent literature have shown an increase in randomized phase II trials published, particularly in the oncologic literature. It is not clear whether this increase in publication of randomized phase II studies is paralleled by an increase in the number of protocols designed. The aim of our study is to assess the prevalence and characterization of randomized phase II protocols submitted to a large university hospital Ethical Board (EB) over the last 5 years. From June 2008 to November 2012, 1130 protocols were submitted to the EB of our Institution. Six-hundred and fifteen were clinical trials and 205 were classified as phase II studies. The prevalence of one-arm designs decreased over time from 60% in 2008 to 45% in 2012. The association of the choice of a randomized design with the type of study sponsor (profit or no-profit), the clinical setting (oncology, infectivology, or other), the type of treatment (drug/device), the number of centres involved (multicenter/one center) and the type of outcome and follow-up duration will be explored. Among randomized trials, the type of analysis (independent one-arm or comparative design) will be assessed. How to increase methodological knowledge will be discussed.
A47

COVARIATE EFFECT ON CONSTANCY ASSUMPTION IN ACTIVE CONTROLLED CLINICAL TRIALS
(1) Siyan Xu (2) Ralph D’Agostino (3) Kerry Barker (4) Sandeep Menon
(1) Boston University (2) Boston University (3) Pfizer, Inc. (4) Pfizer, Inc. and Boston University

Biopharmaceutical companies are heading towards demonstrating a proposed therapeutic protein product is biosimilar to a licensed reference product after the Patient Protection and Affordable Care Act signed into law in 2010. Therefore, instead of answering the question of whether two treatments are different, we are more interested in question like whether a new treatment is equivalent or non-inferior to an existed treatment. In other words, the focus is no longer statistical significance difference, but clinical meaningful comparability. Equivalence or Non-Inferiority (NI) tests are used in late phase of biosimilar studies at efficacy endpoints to show we are equivalent or not inferior to the reference product. If we have a better NI test, it should be easy to apply it to the other direction and get a better equivalence test. In addition to establishment of equivalence/ non-inferiority for efficacy endpoint, if the new treatment also shows other benefits, such as cheaper in price, easy to administration, less toxicity and so on, there will be an improvement in clinical care. To apply NI test, standard statistical methods (Fixed margin method and Synthesis method) requires constancy assumption. However, this assumption may be violated due to covariate effect. My work will focus on comparing four different NI methods: Fixed margin method, Synthesis method, Covariate adjustment method, and Adaptive 2-stage method. The first two are standard methods, but the last two may be useful when constancy assumption is violated due to covariate. We will study the impact of 1) strength of covariate 2) degree of interaction between covariate and treatment arm 3) change in distribution of covariate between historical and current trials on type I error rate and power of each method, and using three different metrics (difference, log relative risk and log odds ratio).

A48

THE CONSEQUENCES OF PROPORTIONAL-HAZARDS BASED MODEL SELECTION
(1) Harlan Campbell, (2) Charmaine Dean
(1) EMMES Canada, (2) The University of Western Ontario

For testing the efficacy of a treatment in a clinical trial, the Cox proportional hazards model is the well-accepted, conventional tool. When using this model, one must confirm that the required proportional hazards (PH) assumption holds true. If the PH assumption fails to hold, it may occur that upon examining a Kaplan-Meier (KM) plot, the survival curves appear to cross, suggesting long-term survival is higher among one group of patients. In this situation - given that the PH assumption does not hold, and given that the KM survival curves are observed to cross - there are options available, proposed as alternatives to the Cox PH model. An important question which arises is whether the potential bias introduced by such a sequential model fitting procedure merits concern and, if so, what are effective mechanisms for correction. We investigate by means of simulation study and draw attention to the considerable drawbacks, with regards to power, of a simple resampling technique, the permutation adjustment, a natural recourse for addressing such challenges. We also consider the recently proposed two-stage testing strategy of Qiu & Sheng (2008) for ameliorating these effects.

A49

ASSESSING THE ‘SUCCESS’ OF THE BLIND IN SHAM-CONTROLLED RANDOMIZED CLINICAL TRIALS AND THE IMPACT ON TREATMENT EFFECT
Valerie Durkalski
Medical University of South Carolina

A critical component to randomized-controlled clinical trials is the inclusion of adequate treatment blinding to help ensure unbiased estimates of treatment effects. Although a common design feature, several trials, particularly device and surgical trials, are challenged to develop adequate blinding procedures. When feasible, these trials attempt to preserve the blind by developing a ‘sham’ control that mimics the experimental treatment. In these cases, it is important to capture the quality of blinding throughout the trial period and to assess the impact on the treatment estimates. Options for assessment include questionnaires of 'best'
treatment guess and confidence in the guess which should be collected at multiple timepoints throughout the trial period. We examine the use of these questionnaires in two recently conducted sham-control trials (one device and one surgical trial) and the relationship between blinding quality, the ‘hunch’ theory and treatment effect.

A50
RESPOSNER ANALYSIS OUTCOMES IN THE ANALYSIS OF ACUTE STROKE CLINICAL TRIALS
(1) Kyra M. Garofolo (2) Sharon D. Yeatts (3) Viswanathan Ramakrishnan (4) Edward C. Jauch (5) Karen C. Johnston (6) Valerie L. Durkalski
(1,2,3,6) Division of Biostatistics and Epidemiology, Department of Medicine, Medical University of South Carolina (4) Division of Emergency Medicine, Department of Medicine, Medical University of South Carolina (5) Department of Neurology, University of Virginia School of Medicine

Background: Traditionally in acute stroke trials, the primary outcome is a dichotomized modified Rankin Scale (mRS). New statistical methods, such as responder analysis, are evolving to address the concern that baseline severity impacts the likelihood of a successful outcome. Responder analysis tailors the definition of success according to baseline severity, producing a more clinically relevant insight into the actual effect of investigational treatments. Whether statistical analyses should adjust for baseline severity when responder analysis is used is unclear, as the outcome already takes into account baseline severity.

Objective: To investigate the effect of covariate adjustment on statistical operating characteristics in the responder analysis framework.

Methods: Using a current stroke clinical trial and its pilot studies to guide simulation parameters, 1000 clinical trials were simulated at varying sample sizes under several treatment effects to assess power and type I error. Covariate-adjusted and unadjusted logistic regression were used to estimate the treatment effect.

Results: Under various settings, the operating characteristics of the unadjusted and adjusted analyses do not substantially differ. Power and type I error are preserved for both the unadjusted and adjusted analyses. Conclusions: Our results suggest that, in the responder analysis framework, the decision to adjust for baseline severity should be guided by the needs of the study, as type I error rates and power do not appear to vary largely between the methods. These findings extend to any trial that uses an ordinal scale as a primary outcome measure and has a baseline prognostic variable.

Trial Registration: This research is part of the Stroke Hyperglycemia Insulin Network Effort (SHINE) Trial, Identification Number NCT01369069, conducted in conjunction with the Neurological Emergencies Treatment Trial (NETT) Network (U01 NS059041, U01 NS056975).

A51
THE BENEFITS AND CHALLENGES OF PERFORMING A CENTRAL REVIEW OF RESPONSE IN A CHRONIC LYMPHOCYTIC LEUKAEMIA TRIAL
(1) Lucy McParland (2) Peter Hillmen (3) Andy Rawstron (4) Alexandra Smith (5) Anna Chalmers (6) Corinne Collett (7) David Phillips (8) Dena Cohen
(1) Clinical Trials Research Unit (CTRU), University of Leeds (2) St James’s University Hospital (3) St James’s University Hospital (4) Clinical Trials Research Unit (CTRU), University of Leeds (5) Clinical Trials Research Unit (CTRU), University of Leeds (6) Clinical Trials Research Unit (CTRU), University of Leeds (7) Clinical Trials Research Unit (CTRU), University of Leeds (8) Clinical Trials Research Unit (CTRU), University of Leeds

Central reviews of response, where responses to therapy are assessed by an independent panel of experts, are beneficial in clinical trials to: increase the accuracy and reliability of the endpoint data; ensure consistency in reporting, particularly in multicenter trials; and reduce potential bias in reporting. Such reviews can be extremely time consuming and expensive. The logistics and appropriateness of conducting a central review requires careful consideration to ensure robust processes and funds are in place prior to any formal analyses. A central assessment on the primary endpoint of our trial in Chronic Lymphocytic Leukaemia (CLL) was recently carried out by a team of independent CLL clinicians prior to an interim analysis. Data required to assess the primary endpoint, rate of complete response (CR) according to iwCLL guidelines, was collated, blinded and sent to two independent assessors, with a third independent assessment requested for discordant reviews. 98 cases were centrally reviewed. In 76(77.6%) cases the initial two assessors gave concordant
results, with a third review required in 22(22.4%) cases. The concordance of the local assessments of response compared with the independent central assessments was evaluated for 94 available cases. 48(51.1%) assessments disagreed, with 28(29.8%) differing to the extent of changing the primary endpoint result. In 20(71.4%) of these cases, local reviewers disagreed with the final central assessment and classified participants as having achieved a CR. The central review process proved challenging in this trial, with difficulties including: time taken to attain the required data items and collate the central reviews; impact on the research sites, the central laboratory who process the biological samples and the clinical trial team; and handling of missing data. However, performing a central review of response proved to be vital to ensure the highest levels of accuracy in reporting the primary endpoint.

A52
CHARACTERISTICS OF PARTICIPANTS AGREEING TO LONG-TERM (UP TO TEN YEARS) FOLLOW-UP IN A LARGE RANDOMIZED CLINICAL TRIAL
(1) Alice J Sheffet, (2) Susan E Hughes, (3) MeeLee Tom, (4) Ariane Mackey, (5) William Brooks, (6) Wayne M Clark, (7) Michael D Hill, (8) Vito Mantese, (9) Jenifer H Voeks, (10) Mary E Longbottom, (11) Thomas G Brott for the CREST Investigators
(1) UMDNJ-New Jersey Medical School, (2) UMDNJ-New Jersey Medical School (3) UMDNJ-New Jersey Medical School, (4) CHA Hop de L’Enfant Jesus, (5) Central Baptist Hospital, (6) Oregon Health Science Univ, (7) University of Calgary, (8) Mercy Hospital St. Louis, (9) Medical University of South Carolina, (10) Mayo Clinic Jacksonville, (11) Mayo Clinic Jacksonville

Background: Retaining participants in clinical trials is a recognized challenge. Long-term data are needed to assess durability of carotid artery stenting (CAS) and carotid endarterectomy because most patients live a decade or longer post-procedure. Identifying patient and site characteristics of those consenting to continue in long-term follow-up is important for study design and operation.

Methods: The Carotid Revascularization Endarterectomy vs. Stenting Trial (CREST) randomized 2502 patients at 117 sites. Characteristics of surviving active participants and their enrolling sites were investigated to compare those who consented to participate long-term (5-10 years) with those who refused continuation past 4 years. Chi-square and t-tests were among those used to test for differences between those who did and did not consent to extended follow-up.

Results: Of 1854 surviving active participants, 1650 (89%) consented to 5-10 years follow-up, and 204 (11%) refused. Sites randomizing >30 patients had a significantly higher proportion consented to long-term follow-up than sites randomizing.

A53
LOGISTICAL CONSIDERATIONS FOR INDEPENDENT REVIEW OF PATIENT IMAGING FOR QUALITY ASSURANCE OR REAL TIME REPORTING IN ONCOLOGY CLINICAL TRIALS
(1) Catherine Lowe (2) Julie Croft (3) Sarah R Brown
(1) Clinical Trials Research Unit, University of Leeds (2) Clinical Trials Research Unit, University of Leeds (3) Clinical Trials Research Unit, University of Leeds

The Response Evaluation Criteria in Solid Tumours (RECIST) guideline recommends that tumour responses are reviewed by an independent expert in clinical trials utilising response rate as a primary endpoint, therefore oncology trials often employ central collection of patient imaging for quality assurance or real time reporting. Here the authors reflect on their experiences of imaging collection in two large UK multicentre randomised phase III Clinical Trials of Investigational Medicinal Product (CTIMPs). PICCOLO compared Irinotecan alone (Ir) with Irinotecan plus Ciclosporin (IrCs) or Irinotecan plus panitumumab (IrPan) in colorectal cancer. Treatment response was assessed locally and independent confirmation of RECIST response was planned on all baseline and 12 week CT scans for quality assurance purposes only. STAR compares conventional continuation strategy (CCS) with an experimental drug-free interval strategy (DFIS) using sunitinib in metastatic renal cell carcinoma. Baseline and all 12 weekly CT scans are assessed by an independent central radiologist and management of trial treatment is based this rather than the local assessment. Both trials required scans to be copied to disc and forwarded to the data centre. Although logistical issues with this process were experienced by both trials, these problems had greater implications to STAR as patient
treatment decisions are reliant on timely assessment of response. Logistical problems experienced included; scans not sent within required timeframe, data centre never receiving discs, anonymisation issues (identifiable information either not obscured or conversely too much detail removed from the disc including scan date, scales and callipers), incompatibility between local and central software, incorrect delivery address used, discs lost in the post and hard copies rather than discs sent to data centre. Further detail on the issues experienced along with possible solutions will be presented, in addition to considerations for future trial design dependant on type of independent review being employed.

A54
MEASURING THE IMPACT OF METHODOLOGICAL RESEARCH AT THE UK MEDICAL RESEARCH COUNCIL CLINICAL TRIALS UNIT
Valerie Brueton, Claire Vale, Rachel Jinks, Babak Oskooei, Jayne Tierney
Medical Research Council Clinical Trials Unit

Background: Evidence of research impact is increasingly requested by research funders. This commonly relies on citation analysis, however other indicators may be more informative. We aimed to measure indicators of impact for our methodological research in the design, conduct and analysis of clinical studies.

Methods: We used methodological research projects conducted or completed at the UK Medical Research Council Clinical Trials Unit (CTU) from 2009-2012. Publication citation rates and related metrics were obtained using ISI Web of Science. In addition, we recorded email queries generated by these projects retrospectively. Semi-structured interviews were conducted with CTU methodologists to obtain further data on the uptake of methods. The content of the interviews and emails was analysed, along with a descriptive analysis of citations.

Results: Preliminary results are based on 38 publications from 2009-2010 and 21 research projects. Seven publications had >10 citations/year in both general and specialised medical articles. Approximately 70% of citations came from original research articles. We were unable to obtain email queries for all projects, but based on data available for 6 projects there were 194 queries from 170 individuals across 23 countries, showing widespread use of these methodological developments and associated software. The 13 interviews revealed numerous impacts including application of methods in new research, further development and dissemination of statistical software, establishment of new national and international collaborations and the uptake of research findings in trial governance guidelines. Updated results will be presented.

Conclusions: We have quantified the impacts of our methodological research via citation rates and research queries. Through the interviews we obtained a range of less obvious but important impacts. These results will be used to inform prospective monitoring of impact and strategies to increase awareness and uptake of our methodological research, providing a framework that others might also use.

A55
TRANSITIONING DATA MANAGEMENT SYSTEMS DURING AN ONGOING STUDY: CHALLENGES, LIMITATIONS AND SUCCESSES
Holly Battenhouse
Medical University of South Carolina

The transition of data management systems from one data management center (DMC) to a new one is a complex process, particularly when studies within the database are active. Accepting data from a previous DMC requires detailed communication, data cleaning, a full understanding of the process of data collection and entry, and documentation of all variables and transfer steps. Challenges to the data transfer process include potential changes to the case report forms and/or system functions, availability of staff with knowledge of the previous database, and system downtime. In order to facilitate a seamless transfer, all aspects of the process must be understood and planned in advance. This presentation is based on the experience gained from the successful completion of the transfer between two DMCs of over 10 years worth of registry data from the NIDDK-funded Acute Liver Failure Study Group. The transfer was from an access database to a new web-based system with revised case report forms and continuing data collection. The presentation will focus on the process as well as lessons learned and recommendations for improvements to the process.
FROM THE LAB TO THE NETWORK: MANAGEMENT OF EXPERIMENTAL, CLINICAL AND SPECIMEN DATA USING THE OPEN SOURCE LABKEY SERVER PLATFORM
(1) Elizabeth K. Nelson, (2) Peter Hussey, (3) Matthew Bellew, (4) Mark Igra, (5) Josh Eckels, (6) Britt Piehler, (7) Adam Rauch
(1-7) LabKey Software

Existing software systems strain to meet the evolving needs of large-scale clinical research networks. To gain insight into difficult diseases, network researchers must make sense of diverse types of data while working as part of distributed teams and leveraging a range of expertise, methods, and data sources. Few software systems provide robust support for managing, analyzing, and sharing not just clinical and specimen information, but also data from complex, high-throughput assays. Systems that provide significant assay support are often organization-specific, assay-specific, and/or separate from systems that provide access to clinical and specimen information. Ad hoc solutions for integrating assay, clinical, and specimen data are often spreadsheet- and email-based, so they are not as robust, scalable, efficient, reliable, or secure as networks need. LabKey Server stands out as an open source system that bridges the data management needs of network labs performing cutting-edge research assays, central project managers tracking study progress, and network scientists analyzing data from many sources. The system supports managing, integrating, analyzing, visualizing, and securely sharing the diverse data types produced by clinical research networks, from clinical records to specimens to cutting-edge assays. Notable differentiators of LabKey Server include easy extensibility, tools for managing newly invented assay data types, built-in support for complex experimental data types (including flow cytometry, proteomics, various plate-based assays, and DNA sequencing), specimen request tracking tools, basic support for ancillary studies, and a built-in R interface. LabKey Server readily integrates with and extends existing systems that are based on databases and/or spreadsheets. Since its launch in 2005, this open source system has been adopted and customized by organizations across the globe, including the Immune Tolerance Network and consortia within the Global HIV Enterprise. Source code, compiled binaries, documentation and tutorials are professionally maintained and freely available under the Apache 2.0 license at http://www.labkey.org.

AN ANALYSIS OF DATA CHANGES IN RESPONSE TO REAL-TIME PROTOCOL VIOLATION (PV) ALERTS IN AN ELECTRONIC DATA CAPTURE (EDC) SYSTEM
(1) Cassidy Conner, (2) Tomoko Goddard, (3) Aaron Perlmutter, (4) Catherine Dillon, (5) Karen Briggs, (6) Keith Pauls, (7) Wenle Zhao, (8) Valerie Durkalski, (9) Yuko Palesch
Medical University of South Carolina

Real-time protocol violation (PV) alerts programmed in an Electronic Data Capture (EDC) system are triggered at the time of data entry when data deviates from what is allowed by the protocol. These alerts can serve many purposes, including blocking ineligible subjects from randomization and providing a mechanism for sites to explain the reasons for violations. However, real-time PV alerts may lead to inappropriate data changes in order to bypass a violation. To determine the extent of inappropriate data changes in response to PV alerts, the audit trail of an EDC system for an NIH/NINDS funded phase III multicenter trial was reviewed. We found that the majority of data changes in response to PV alerts were made in order to correct data entry errors due to the complexity of the programmed data checks. When real-time PV alerts are programmed based on the responses to several data points on a form (e.g. the logical progression of a time/date sequence) or based on responses across multiple forms, this complexity resulted in higher rates of data entry errors. Based on the results of this analysis, several key improvements were made to the EDC system and case report forms to help reduce data entry errors.

HOW MUCH IS ENOUGH: A RISK BASED APPROACH TO ANALYSIS VERIFICATION
Deborah Toyota
Boston Biomedical Associates

Verification is typically the most time consuming part of producing an analysis. This is especially true for interim analyses where “dirty” data makes verification all the more difficult and all the more necessary.
There are different approaches to verification, including manually counting and calculating as well as independent programming with manual or automated compare. All approaches have their share of advantages and their share of disadvantages. And, all approaches typically take more time to verify than they did to produce. One sure way to reduce verification time is to reduce the amount of items checked during verification. The question is, though, how to determine what to check and what not to check. It is also important to consider how to document that decision. A formal risk analysis is the key to scaling analysis verification in a way that provides documented evidence of the decision. FDA has published several guidance documents discussing risk based approaches to different processes used in clinical trials. Using these documents as reference, BBA developed an analysis verification risk assessment. This assessment has been utilized on several projects over the past eighteen months. This presentation will share the risk-based strategies for analysis verification implemented by BBA. The impact on verification processes will be presented and lessons learned will be discussed.

A59

PHASE II/III SEAMLESS ADAPTIVE DOSE SELECTION DESIGN FOR LONGITUDINAL PATIENT DATA
(1) Caitlyn Ellerbe (2) Jordan Elm (3) Viswanathan Ramakrishnan (4) Bruce Turnbull
(5) Stacia DeSantis (6) Edward Jauch (7) Valerie Durkalski
(1) Division of Biostatistics and Epidemiology, Medical University of South Carolina (2) Division of Biostatistics and Epidemiology, Medical University of South Carolina (3) Division of Biostatistics and Epidemiology, Medical University of South Carolina (4) Departments of Statistical Science and of Operations Research, Cornell University (5) Division of Biostatistics and Epidemiology, Medical University of South Carolina (6) Division of Emergency Medicine, Medical University of South Carolina (7) Division of Biostatistics and Epidemiology, Medical University of South Carolina

In response to the increase of randomized controlled trials (RCT) that fail during late stages of clinical development, adaptive design trials offer investigators the ability to modify trial parameters to promote safety and trial efficiency. Despite interest in these designs, implementation is rare in the confirmatory setting given uncertainty over whether they maintain RCT standards. Consider a trial comparing three doses of interest and a placebo control, where the primary endpoint is repeatedly measured at several follow-up visits and a clinician preferred dose exists. In a non-adaptive setting, a Phase II study selects the optimal dose based on a short-term endpoint before testing this against a control during a Phase III RCT using a long-term definitive endpoint. Downsides of this approach include the lack of long-term primary endpoint and safety information available for dose selection (Phase II) as well as the downtime associated with pre-study activities between phases. Alternatively, several seamless adaptive dose finding designs exist. When longitudinal data exist, these designs are limited in their ability to use all available information because of correlation between measures. At the interim analysis (dose selection), investigators can either use a short-term surrogate endpoint for all available individuals or a long-term definitive outcome for a subset of the recruited population depending on the selected design. The final analysis (efficacy) must incorporate the remaining data such that the type I error (alpha) is not inflated due to analyzing a single point multiple times. We propose a design using all available information (longitudinal data) at each analysis thus providing an intuitive approach which increases efficiency (decreased N). The design demonstrates the following features: a) ability to control alpha; b) theoretically sound combination function; c) closed testing procedure for correlated hypotheses; d) physician guided dose selection.

A60

INTERNAL PILOT DESIGN FOR CLUSTER RANDOMIZED TRIALS WITH UNEQUAL CLUSTER SIZES
(1) Ashutosh Ranjan (2) Christopher Coffey (3) Leslie McClure
(1) Department of Biostatistics, University of Alabama at Birmingham (2) Department of Biostatistics, University of Iowa (3) Department of Biostatistics, University of Alabama at Birmingham

Background: Cluster randomized trials are popular in health care research and community based studies. Many of these trials, while planned for equal cluster sizes, in reality have unequal cluster sizes which may lead to an underpowered study. Internal pilot designs can be used to adjust the sample size midway through a study to account for variation in cluster size to help ensure sufficient power to detect planned treatment differences.
Methods: We compare two methods for re-estimating sample size midway through a cluster randomized trial, for varying levels of imbalance in cluster size: 1) sample size is re-estimated assuming equal cluster sizes followed by a cluster level analysis and 2) sample size is re-estimated considering variability in clusters followed by a cluster level analysis weighted by cluster size. The operating characteristics of the two methods are investigated for a hypothetical but typical community based trial.

Results: As long as the initial planning variance overestimates the true variance, both the methods maintain desired study characteristics. Method 2 maintains power close to the target level with trivial bias in type I error across all the scenarios of imbalances considered when the initial planning variance underestimates the true variance. In comparison, there is substantial loss in power when Method 1 is used, when the true variance is greater than the initial variance estimate.

Conclusion: Sample size re-estimation using a formula that takes into account the variability in cluster sizes for a trial planned assuming equal cluster sizes, followed by a cluster-level analysis weighting by cluster sizes, is recommended for trials with unequal cluster sizes. This method requires a larger number of clusters be recruited, but researchers employing this procedure will have a better chance of realizing the aims of their study with adequate power.

A61
LIKELIHOOD-BASED METHODS FOR THE PREDICTION OF ENDPOINT OCCURRENCES IN CLINICAL TRIALS
(1) Megan Smith (2) Ying Qing Chen
(1) University of Washington (2) Fred Hutchinson Cancer Research Center

The successful conduct of a clinical trial requires careful coordination of the sponsor, clinicians, statisticians, Data Safety Monitoring Board, and many others. Tools to aid in the logistical management of trial operations are essential and help facilitate best adherence to preset interim analysis and monitoring plans. In a clinical trial with censored time-to-event data, the statistical information is governed by the number of observed events. Once the trial is underway, it is useful to have methods to predict the timing of future landmark numbers of events. Unforeseen trends in accrual, health outcomes, or patient dropout may influence the accuracy of design-phase event time predictions. Thus a sensible strategy is to use accumulating data from the trial itself to refine preliminary projections. Here we present flexible, computationally inexpensive likelihood-based approaches for the real-time prediction of endpoint occurrences. We describe two methods to obtain point estimates and prediction intervals for the timing of landmark events. The first is a marginal method that produces an overall estimate of the observed and anticipated study progress. The second method predicts the timing of future events conditional on the exact number and timing of events already observed in the trial. Our data-driven approaches achieve favorable accuracy and precision, are simple to implement, and require neither prior assumptions about model parameter distributions nor the simulation of additional data. Simulation studies demonstrate the accuracy and uncertainty measures for these methods under various distributional assumptions for accrual, event, and dropout times. We also compare our model predictions to actual event times observed in data from the HIV Prevention Trials Network 052 Study to assess the effectiveness of immediate versus delayed antiretroviral therapy strategies on sexual transmission of HIV-1.

A62
CHARACTERIZATION OF TWO-STAGE LIKELIHOOD CONTINUAL REASSESSMENT METHOD
(1) Xiaoyu Jia (2) Shing M Lee (3) Ying Kuen Cheung
(1) Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center; Department of Biostatistics, Columbia University. (2) Department of Biostatistics, Columbia University. (3) Department of Biostatistics, Columbia University

The continual reassessment method (CRM) is an increasingly popular model-based method for dose finding clinical trials among clinicians. A common practice is to use the CRM in a two-stage design, whereby the model-based CRM is activated only after an initial sequence of patients are tested. While there are practical appeals of the two-stage CRM approach, the theoretical framework is lacking in the literature. As a result, it is often unclear how the CRM design components (such as the initial dose sequence and the dose toxicity
model) can be properly chosen in practice. This paper studies a theoretical framework that characterizes the design components of a two-stage CRM, and proposes a calibration process. A real trial example is used to demonstrate that the proposed process can be implemented in a timely and reproducible manner, and yet offers competitive operating characteristics when compared to a labor-intensive ad hoc calibration process. We also illustrate using the proposed framework that the performance of the CRM is insensitive to the choice of the dose-toxicity model.

A63

FUTILITY DESIGN INCORPORATING CONCURRENT CONTROLS: A VARIATION ON THE PHASE II TRIAL DESIGN

(1) Sharon D Yeatts (2) Magdy Selim (3) Yuko Y Palesch
(1,2) Division of Biostatistics and Epidemiology, Department of Medicine, Medical University of South Carolina (1,2) Data Coordination Unit, Medical University of South Carolina (3) Department of Neurology, Stroke Division, Beth Israel Deaconess Medical Center, Harvard Medical School

Background: There is a rising need for more efficient designs to discard ineffective therapies. The objective of the single-arm futility design is to reject treatments without promise. Use of a pre-specified threshold, rather than a concurrent control, as the comparator, reduces the sample size of the traditional concurrently controlled Phase II design. The threshold can be determined based on a clinically relevant treatment effect and historical control data. Potential pitfalls associated with historical control data, including temporal changes and protocol variations across trials, raise concerns over its use. If the historical control data is no longer relevant, results may not accurately reflect the futility of the treatment. Calibration of the threshold via a small control cohort has been suggested. If this cohort is too small, its usefulness is limited; if too large, the trial resembles an underpowered Phase III.

Objective: To propose an appropriately powered concurrently controlled futility design and describe its application in the phase II trial of deferoxamine in ICH.

Methods: The futility hypothesis is based on a direct comparison of the arms, in which the treatment effect is compared to a pre-specified futility threshold. Where the primary endpoint is favorable outcome, an absolute treatment effect less than/in favor of the experimental treatment leads to rejection of the treatment as futile. Failure to conclude futility is evidence in favor of a phase III trial.

Conclusions: Despite the increase in sample size over the single arm design, the proposed is not an alternative to Phase III testing. The objective remains to establish futility, rather than to demonstrate efficacy, of the treatment. A concurrent control arm avoids the drawbacks of historical control data and allows for a direct comparison of treatment arms.

A64

PERSONALIZED AND ADAPTIVE TRIALS IN PHASE 2: COMPARISON OF DESIGNS

(1) Russell Reeve (2) Bradley Ferguson
(1) Quintiles, Inc. (2) North Carolina State University

Interest in adaptive trials is increasing, but there is still considerable debate about their value in clinical research. This debate extends from how efficient they are, to how to implement these trials in practice. In this talk, we develop an optimal trial design based on the Hill model, and compare its efficiency compared with standard trials. We present this design in the context of a trial to assess the advantages of a novel rheumatoid arthritis agent in the presence of methotrexate background therapy. In this design, we use ACR20 – a binary endpoint that measures improvement over baseline – as the primary endpoint. In addition to the adaptive design, we also introduce a personalized design for this case. We show that both of these designs yield improvement in the precision to estimate the ED70 versus a fixed design. A discussion of using adaptive designs based on the Hill model versus the more flexible categorical models (e.g., NDLM) will be discussed, and the Hill model design will be shown to be at least as efficient in estimated the ED70 as the NDLM design. And in fact, the personalized design is shown to have superior properties over the adaptive design, which is consistent with efficiency of randomized concentration-controlled trials. Finally, a discussion of some challenges in implementing these highly adaptive designs will be discussed, with solutions based on experience.
A65
RESPONSE-ADAPTIVE DOSE-FINDING FOR CORRELATED BIVARIATE DATA WITH APPLICATIONS TO COMPLEMENT SYSTEM INHIBITION STUDIES
Mitchell Thomann
University of Iowa

The complement system is an important part of the innate human immune system. In particular, the C5 complement system plays a role in inflammatory responses throughout the body. Inhibition of this system has been shown to have therapeutic effects in a variety of diseases, including arthritis and sepsis. An important outcome that measures the effectiveness of drugs that work via C5 inhibition is the percent of the complement system that is inhibited. These efficacy outcomes are continuous and bounded. Furthermore, there is little benefit expected when treating patients at higher doses once the maximum level of inhibition has been achieved. Thus, when performing dose-finding trials for drugs that inhibit this system, it is important to consider both drug-related toxicities and complement system inhibition. This paper presents a response-adaptive method for dose-finding that jointly models subjects' binary toxicity and bounded continuous efficacy outcomes. There are existing dose-finding methods for modeling binary toxicity and normal efficacy outcomes; this method expands their utility to bounded efficacy outcomes. It is based upon the factorization of the joint distribution of the outcomes using the product of the distribution of toxicity outcome and the conditional distribution of the efficacy outcome. Parameter estimates are obtained using Markov Chain Monte Carlo sampling. Groups of subjects are sequentially enrolled to the trial at dose levels based upon updated estimates of efficacy and toxicity. A simulation study is also presented to demonstrate the utility of this method under a variety of conditions. The simulation study's results are summarized and discussed. The simulations demonstrate that this method performs well when the data are generated from a truncated normal distribution. Similar models could be derived for bounded data from other truncated continuous distribution functions.

A66
UNINTENTIONAL OPERATION MISTAKES IN EMERGENCY TREATMENT TRIALS AND THEIR IMPACTS ON TRIAL RESULTS
Wenle Zhao, Yuko Palesch
Medical University of South Carolina

Unintentional operation mistakes could occur in clinical trial practices in various ways. Trials treating medical emergencies like strokes and traumatic brain injuries are even more vulnerable. While some operation mistakes may have trivial impacts on the final results of the trial, those mistakes associated with subject eligibility assessments, subject randomization, study drug dispatch, treatment administrate, and primary outcome assessment could affect the trial results seriously. Typically, in clinical trial design, a sample size inflation factor is used to cover the power lost caused by noncompliers. However, the proportion of noncompliers could be under estimated for emergency treatment trials, and different types of operation mistakes could have different impacts. This presentation will quantify the impacts of commonly observed unintentional operation mistakes on the power and type I error under both superiority and on-inferiority trial scenarios.

A67
INVESTIGATING TREATMENT EFFECT VARIABILITY IN RANDOMIZED LONGITUDINAL CLINICAL TRIALS
Joseph R. Rausch
Cincinnati Children’s Hospital Medical Center University of Cincinnati, College of Medicine, Department of Pediatrics

The standard approach to investigating the effects of intervention/prevention regimens on continuous outcomes in randomized clinical trials is via the average (i.e., mean) causal effect. However, there is reason to believe that averages can hide important variation in the treatment effect in many real-world applications and even can potentially mislead researchers and clinicians alike when determining the appropriate
Abstracts

S31

http://ctj.sagepub.com  Clinical Trials  2013; 10: S1 –S88

treatment for a particular individual. Statistical methods which directly address the possibility of variance in
the treatment effect should be employed to better understand treatment effects at the individual level. After
reviewing statistical challenges associated with examining treatment effect variability, I discuss one particu-
larly worthwhile approach for investigating treatment effect variability in randomized longitudinal clinical
trials using multilevel modeling. Hypothesis testing and estimation, in conjunction with their correspond-
ing assumptions, are discussed to assess scenarios in which these approaches are appropriate in applied
research. The practical ramifications of utilizing a statistical approach which directly assesses treatment
effect variability are delineated. Finally, an illustrative example is provided to facilitate dissemination of this
methodological framework more generally within the context of randomized longitudinal clinical trials.

A68
DICHOTOMIZING PARTIAL COMPLIANCE AND INCREASED PARTICIPANT BURDEN IN
FACTORIAL DESIGNS: THE PERFORMANCE OF FOUR NON-COMPLIANCE METHODS
(1) Peter Merrill (2) Leslie McClure
(1) University of Birmingham at Alabama (2) University of Birmingham at Alabama

Non-compliance to treatment assignment is an inevitable occurrence in randomized clinical trials (RCTs).
Intention to treat (ITT) is generally considered the best method for addressing non-compliance in RCTs.
Several alternatives to ITT exist, including: per protocol (PP), as treated (AT), and instrumental variables (IV).
These three methods define participant compliance as a binary variable, but partial compliance is a very
common occurrence in RCTs. By defining a threshold, above which a participant is called a complier, PP, AT
and IV can, but the resulting loss of information may affect their performance. Additionally, trials with fac-
torial designs may experience higher rates of non-compliance due to a heavier burden participants experi-
ence by being assigned to multiple experimental treatments. Using simulations, we assessed the performance
of ITT, PP, AT, and IV in both the partial compliance setting and in a 2-by-2 factorial study with increased
participant burden for those randomized to both active treatments. The bias, mean square error, and type I
error rates of the IV method after dichotomizing partial compliance was heavily inflated. PP and AT improved
on the bias and power of ITT without inflating type I error beyond acceptable limits. The use of PP and AT
cannot readily be recommended because of the possibilities of selection bias inherent in the use of these
methods. The performance of all four methods depended heavily on the level non-compliance present, with
higher average non-compliance leading to poorer performance. The results suggest the need for a method of
estimating treatment effects that can fully utilize partial compliance information.

A69
USING CENTRAL REVIEW DATA IN ANALYSES OF PHASE III TRIALS
Sarah Brown, Helen Marshall, Catherine Lowe, Julie Croft
Clinical Trials Research Unit, University of Leeds

PICCOLO is a multicentre, phase III, advanced colorectal cancer trial, comparing Irinotecan alone (Ir) with
Irinotecan plus Ciclosporin (IrCs) for patients with KRAS mutated or undefined tumours, and with Irinotecan
plus panitumumab (IrPan) for patients with KRAS wildtype tumours. The trial was designed as two parallel
phase III trials, accounting for patients’ KRAS status. For the IrCs comparison, a non-inferiority hypothesis
was proposed for the primary endpoint of the proportion of patients alive and progression-free at 12 weeks
post-randomisation, based on RECIST criteria. A central, independent review of progression assessments at
this time point was planned. During the review process, a number of logistical issues arose. Difficulties were
encountered with obtaining both baseline and follow-up scan images from centres; with software issues
restricting access to images received; and with ability to read and combine images to enable assessment of
progression-free status. These logistical issues resulted in independent review of only a subset of patients
being possible. Primary analysis was pre-specified to use local clinical assessment data, with independently
reviewed data providing a quality assurance assessment only. Data summaries of the proportion of patients
for which there was disagreement in the primary endpoint of progression-free status were performed. Final
analysis of the IrCs comparison failed to demonstrate non-inferiority of IrCs compared to Ir on the basis of
the locally assessed primary endpoint. Of 241/672 patients for whom a central assessment of progression-
free status was performed, discrepancies were apparent for 54 patients (22%). In an extrapolated ad-hoc
analysis taking into account these discrepancies, the primary endpoint result was unchanged and
non-inferiority was still not demonstrated. We will present our approach to incorporating such review data into analyses, and discuss the purpose of these types of review data.

A70
CURRENT PRACTICES IN REMOTE MONITORING VISITS
Ellen Peskin, Kathleen McWilliams, Maureen Maguire
University of Pennsylvania Department of Ophthalmology Clinical Trials Coordinating Center, Center for Preventive Ophthalmology & Biostatistics

Effective monitoring of study activities in multi-center clinical trials is vital to ensure the protection of human subjects, adherence to study protocols and the collection of accurate data. The gold standard for monitoring has been in-clinic visits, which are effective, and allow the monitor to directly observe implementation of study protocols. Disadvantages of in-person visits are cost and travel time - key issues in this resource-tight environment. The FDA has recently issued draft guidance for risk-based monitoring, including adoption of alternative monitoring methods. Similarly, ICH E6 advises developing monitoring plans that consider several factors in determining the nature of monitoring for a given trial. The development of new technologies allows for monitoring to occur remotely, using the Internet and VOIP services. Currently there are no specific guidelines pertaining to when remote monitoring of clinical trials is appropriate (or not), no guidance on the best methods for conducting them and the resources necessary to implement remote monitoring visits. This leaves Coordinating Centers, sponsors and CROs with wide discretion but less certainty whether a remote monitoring program is compliant with GCP, particularly in trials seeking FDA marketing approval. For this study we will conduct interviews with leaders of 25 Coordinating Centers and CROs conducting both NIH and FDA-regulated trials to ascertain current remote monitoring practices. Interviewees will be queried on the frequency of remote site monitoring, the factors driving the decision to employ them, the equipment required for effective ‘visits’, monitor training, lessons learned and the benefits and drawbacks relative to in-person visits. Interviews will be also conducted with 10 Clinic Coordinators who have experienced remote monitoring visits for their assessment of the relative strength and weaknesses of such visits. The study aims to identify current remote monitoring practices, strategies for successful implementation and guidance for Coordinating Centers/CROs considering adopting them.

A71
DESIGN OF AN INTERNATIONAL CLUSTER-RANDOMIZED TRIAL COMPARING TWO DATA MONITORING PRACTICES
(1) Katherine Huppler Hullsiek, (2) Nicole Wyman, (3) Jonathan Kagan, (4) Jesper Grarup, (5) Catherine Carey, (6) Fleur Hudson, (7) Elizabeth Finley, (8) Waldo Belloso
(1) Division of Biostatistics, University of Minnesota (2) Division of Biostatistics, University of Minnesota (3) National Institute of Allergy & Infectious Diseases, National Institutes of Health (4) Copenhagen HIV Programme, University of Copenhagen (5) The Kirby Institute, University of New South Wales (6) Medical Research Council (7) Veterans Affairs Medical Center (8) Hospital Italiano de Buenos Aires (8) Hospital Italiano de Buenos Aires

Monitoring of clinical research studies protects human subjects and the integrity of trial data. While required by regulatory agencies, specific monitoring activities are at the discretion of the sponsor. Research is needed on the effectiveness and efficiency of risk-based monitoring. As part of the START study the International Network for Strategic Initiatives in Global HIV Trials (INSIGHT) clinical trials network is conducting a cluster randomized Monitoring Substudy. Eligible sites are randomized to one of two strategies: (1) central and local monitoring or (2) central and local and on-site monitoring. The purpose is to evaluate if site performance is improved with on-site monitoring in addition to central and local monitoring. The START study is suitable for evaluation of on-site monitoring because it is a strategic trial on the timing of antiretroviral therapy and not a pivotal or early phase safety study. All sites are centrally (remotely) monitored by the statistical/data center and international coordinating centers for data quality issues and FWA lapses in real-time; local (internal) monitoring is conducted at all sites by the semi-annual completion of forms designed to assess site quality assurance adherence; sites randomized to on-site monitoring also have annual visits conducted by a person independent of the site. A challenge for the Monitoring Substudy design was crafting a primary outcome that could be captured both with and without on-site monitoring. The primary composite outcome
includes components related to major eligibility and informed consent violations, use of drugs not permitted by protocol, START primary endpoints and serious events not reported within 6 months of occurrence, and data alteration or fraud. Currently the ongoing Monitoring Substudy has randomized 215 sites (enrolling 3013 people) in 33 different countries. We describe in detail the central, local and on-site monitoring activities along with the design and implementation challenges that have been encountered.

**A72**

**CORRELATION OF A FORM-BASED VERSUS ITEM-BASED METRIC OF DATA ACCURACY IN CASE REPORT FORM COMPLETION IN A CLINICAL TRIAL NETWORK**

Erin Bengelink, Tess Bonham, Cassidy Connor, Andrace Deyampert, Catherine Riley Dillon, Valerie Durkalski, Shirley Frederiksen, Deneil Harney, Donna Harsh, Samkeliso Mawocha, Keith Pauls, Joy Pinkerton, Arthi Ramakrishnan, Valerie Stevenson, Wenle Zhao, Robert Silbergleit

*Neurological Emergencies Treatment Trials (NETT) Network Clinical Coordinating Center, University of Michigan Neurological Emergencies Treatment Trials (NETT) Statistical Data Management Center, Medical University of South Carolina*

**Background:** Performance improvement in a clinical trial network requires metrics that are both valid and practical, but there can be tension between these two qualities. Granular metrics with face validity can be more burdensome than metrics using aggregated data. To measure data collection accuracy, we compared two alternatives. One was an item-based metric (IBM) that required added coding by site monitors but had evident face validity. The other was a form-based metric (FBM) that was readily accessible from typically available information, but with less evident validity.

**Objective:** To determine the correlation between a pragmatic FBM of data accuracy in case report form (CRF) completion and an item-based metric IBM with greater face validity, and to compare the characteristics of each.

**Methods:** The IBM described the number of data items corrected / number of monitored CRFs after source document verification and data clarification request (DCR) by a site monitor. This required coding every individual DCR as to whether it did or did not result in a correction. The FBM described the number of CRFs with 1 site monitor DCR's / the number of CRFs monitored. The metrics were calculated for 16 enrollment Hubs participating in an ongoing data-intensive clinical trial of acute traumatic brain injury. Spearman Rank Correlation by Hub was determined. The distributions of the two metrics were characterized.

**Results:** There were 10,693 monitored CRFs for 663 subjects containing 2,297 DCRs. 1,344 DCRs led to data corrections. The Spearman rank correlation between the IBM and the FBM was 0.77 (p=0.00074). Distributions of the two alternative metrics were similar.

**Conclusions:** Development of performance metrics may involve tradeoffs between apparent validity and pragmatic implementation. In this application of CRF completion data accuracy metrics we found that a more pragmatic FBM closely approximates but does not completely reproduce the results of the IBM.

**A73**

**QUANTUM LEAP - LEVERAGING REAL-TIME (DIRECT) DATA-ENTRY TO INCREASE SPEED, IMPROVE QUALITY, AND DRAMATICALLY REDUCE COSTS**

(1) Dean Gittleman, (2) Jules Mitchell, (3) Dario Carrara, (4) Judith Schloss Markowitz

(1) Target Health Inc. (2) Target Health Inc. (3) Ferring Pharmaceuticals A/S (4) Target Health Inc.

Electronic Data Capture (EDC) has been in routine use in clinical trials for over a decade. Its promise of better data faster, though, continues to go largely unrealized. In part this is due to layering its use on top of processes developed for paper-based trials. In part this is due to the timidity of sponsor organizations to take advantage of Regulatory gifts (e.g., Risk-based Monitoring, eSource). This resistance to change manifests itself in several ways, but particularly in the continued reliance on site-based source data verification (SDV) by monitors, a practice which we view as both high cost and low value. This paper describes the lessons learned from recently completed trials and one ongoing phase III trial. As such, it presents real-world experience in the use of processes optimized around direct data-entry, using an EDC tool specifically tooled to support real-time data-entry by investigative sites. The authors present data that demonstrate the value inherent in capturing trial data at the time of subject visits -- specifically that this enables near real-time monitoring, improved data quality, and improved ability to make evidence-based decisions much faster, while at the same time dramatically reduced monitoring costs, and delivering an improved site experience.
Further, the authors describe lessons learned from the implementation of risk-based monitoring plans, centralized (or remote) monitoring, and regular quality reviews, and how these are enabled through the use of technology-enabled, real-time capturing of data at investigative sites. Finally, the paper demonstrates that real-time (direct) data-entry is not only feasible, but preferable to stakeholders, and lays out the challenges and pay-offs for moving in this direction.

A74

GRAPHICAL METHODS FOR MONITORING CLINICAL TRIAL DATA
(1) Elizabeth Sugar, (2) Lea Drye, (3) Thomas Louis, (4) Janet Holbrook
(1-4) Bloomberg School of Public Health, Johns Hopkins University

Monitoring the trade off between risks and benefits as well as the integrity of the data are important components of clinical research. One of the roles of a data coordinating center (DCC) is to provide the data safety monitoring board (DSMB) with the information necessary to adequately evaluate the study. Summary statistics and pre-defined modeling make up one component of DSMB reports, however these summaries usually examine separately adverse events and efficacy outcomes, or are focused on very specific questions (e.g., visual acuity at six months after cataract surgery). Graphical representations of the data often provide greater flexibility to explore the individual components and simultaneously to examine multiple factors. To this end, we describe our experience as a DCC using graphical methods for monitoring data. Understanding the longitudinal patterns in data is one of the most important tasks, but it is also one of the most difficult to accomplish. Standard techniques such as Kaplan-Meier (KM) survival curves, sequential box-plots, scatter-plots and spaghetti plots are all regularly used, but do not provide a complete picture. KM curves are useful for identifying the first event but ignore the initial event duration and all subsequent events. Spaghetti plots have the ability to track efficacy outcomes and adverse events simultaneously, but often become unreadable for moderate to large samples. Additional techniques are being developed (e.g., time-lines and Lasagna plots), and these can be adapted to provide information about both risk and benefit over time. Careful selection of sorting variables is especially important when using techniques that display individuals on a single line over time. Development and evaluation of graphical techniques for displaying data is an important component of trial monitoring. We provide examples of their role in monitoring in asthma and ophthalmology trials, demonstrating both the strengths and limitations of several methods.

A75

IDENTIFYING POTENTIAL ADVERSE EFFECTS BY PATIENTS’ RATINGS: A PROOF-OF-CONCEPT STUDY OF A NOVEL APPROACH
(1) Rajnish Mago, (2) Angelica Kloos, (3) Constantine Daskalakis, (4) Michelle Shwarz, (5) Barry W. Rovner
(1) Thomas Jefferson University, (2) Department of Child and Adolescent Psychiatry, Children’s National Medical Center, (3) Thomas Jefferson University, (4) Thomas Jefferson University, (5) Thomas Jefferson University

Objective: Methods to evaluate adverse effects in clinical trials are significantly underdeveloped compared to those for evaluating efficacy. In this pilot proof-of-concept study, we preliminarily compared a novel approach ‘the Symptom Assessment Tool (SAT)’ to a systematic and detailed assessment by a physician for identifying symptoms that were potentially adverse effects (sensitivity) and excluding symptoms that were unlikely to be adverse effects (specificity).

Methods: The SAT consisted of a) a symptom inventory and rating of symptom severity were completed before starting a psychotropic medication (or increasing its dose) and again 2 weeks later, and b) an algorithm based on questions answered by patients that supported or argued against the symptom being an adverse effect. Each symptom was systematically assessed by both the patient-rated SAT and by a physician and was classified as either a potential or an unlikely adverse effect. The primary analysis compared the classification of symptoms by the SAT to that by the physician. Potential adverse effects were subcategorized as possible or probable adverse effects.

Results: A sample of 193 symptoms from 15 adults was evaluated, only 37.3% of which were considered potential adverse effects by the physician. Approximately half of the “treatment-emergent” symptoms were considered unlikely adverse effects by the physician, most commonly due to a clear alternative explanation. Sensitivity of the SAT compared to physician’s assessment was 90.3% for potential adverse effects and
97.5% for the subgroup of probable adverse effects. The SAT correctly identified 63.6% of the symptoms as unlikely adverse effects (specificity), and its negative predictive value was 91.7%.

**Conclusions:** The SAT, appropriate for its intended use as an initial patient-rated tool in clinical trials, had high sensitivity and moderate specificity and could present the investigator with a limited number of potential adverse effects for further assessment and intervention. Further evaluation and refinement of this approach is warranted.

---

**A76**

**THE RISKS AND BENEFITS ASSOCIATED WITH RUNNING INVESTIGATOR LED EARLY PHASE TRIALS IN COLLABORATION WITH PHARMACEUTICAL AND CHARITY PARTNERS**

(1) Nina Thenagaran (2) Louise Flanagan (3) Sarah Brown (4) Sarah Flynn (5) Deborah Sherratt (5) Avie-Lee Coney

*Clinical Trials Research Unit, University of Leeds*

The Clinical Trial Coordinating Office (CTCO) based at the University of Leeds manages a portfolio of early phase multiple myeloma trials in collaboration with the charity Myeloma UK and their Early Phase Clinical Trial Network, and a number of pharmaceutical companies who supply the compounds of interest. These symbiotic collaborations between charity, Clinical Trials Unit (CTU) and the pharmaceutical industry are becoming common practice in clinical research, especially in the wake of recent government reductions to the scientific research budget in the United Kingdom. Now more than ever, these parties have a greater mutual interest in sharing costs and experiences, pooling of risks and ultimately providing high quality scientific research to improve treatment options for patients. The vast experience offered by the pharmaceutical industry, CTCO and Active Trial Centers (ATCs) allows the relatively inexperienced charity to drive forward clinical research in their particular disease area. There are of course, disadvantages to these collaborations, in that each party has different goals and missions which may cause conflict and subsequent time delays. Based on the experiences of the CTCO, we identify the risks and benefits of such collaborations, including the impact of funding on the prioritization of and recruitment in to investigator led trials; the autonomy to test regimens that may not ordinarily be considered and the issues surrounding new collaborations with inexperienced partners.

---

**A77**

**MANAGING PERFORMANCE OF FIELD STAFF IN THE GULF LONG-TERM FOLLOW-UP STUDY (GULF STUDY)**

(1) Steven K. Ramsey, (1) Edward E. Gaunt, (1) Brian D. Blackmon, (1) Matthew D. Curry, (1) Christopher K. Treseder, (1) David A. Johndrow, (2) Richard K. Kwok, (2,3) Larry S. Engel, (2) Dale P. Sandler

(1) SRA International, Inc.; (2) National Institute of Environmental Health Sciences; (3) University of North Carolina

Standardization in data collection activities is critical in assuring the validity of health studies, especially in large-scale, field-based studies examining a wide range of outcomes. This presentation describes our approach to managing performance of field staff to achieve standardization in the GuLF STUDY, a large-scale epidemiological study investigating the potential health effects associated with exposure to oil and dispersants among clean-up workers who responded to the 2010 Deepwater Horizon oil spill. More than 31,000 participants have enrolled to date, including 8,600 who completed a home examination that included anthropometric, clinical, and environmental measurements. Home examinations have been conducted by 57 home-based certified medical assistants (CMAs) who live in the five Gulf states. CMAs are required to complete a four-day, in-person training session provided by the study coordinating center, which includes practice and certification activities. At the time of training, CMAs receive a variety of job-aides to ensure adherence to a standard procedures. Following training, state-based managers oversee the CMAs in collaboration with the coordinating center field operations manager. The operations manager circulates weekly reports to managers that include performance metrics by state and CMA and over time (e.g. procedural completion rates, conditions of specimens) and holds weekly manager teleconferences to review the reports. State managers use these reports to identify areas where additional coaching or re-training is needed. In addition, state managers periodically observe visits and record their observations on standardized quality control forms that are
transmitted to the coordinating center. The findings from these visits are also used to identify areas where additional training is needed. Our approach may be of interest to other coordinating centers managing large-scale, field-based investigations.

A78

ESTIMATING SITE COSTS PRIOR TO CONDUCTING CLINICAL TRIALS – A STUDY SITE BUDGETING TOOL

(1) Dorothee Arenz, (2) Barbara Hero, (3) Barbara F. Eichhorst, (4) Martin Langer, (5) Lars Pester, (6) Julia von Tresckow, (7) Maria J.G.T. Vehreschild, (8) Oliver A. Cornely

(1) (4) (8) Clinical Trials Center Cologne, ZKS K’in, BMBF 01KN1106, University of Cologne; (2) Department of General Pediatrics, University Hospital Cologne; (3) (5) (6) (7) 1st Department of Internal Medicine, University Hospital Cologne

Rationale: Development of a tool for calculating trial fees prior to initiating a clinical trial. Background. Conducting clinical trials is costly and time-consuming. Trial sites usually do not calculate site costs. Underestimating required resources slows enrollment and lowers data quality. It is unclear how to reliably estimate trial site costs.

Methods: To develop this tool, trial staff from sites at the University of Cologne, Germany formed a task group. Tasks within a clinical trial were itemized into single activities and basic time expenditures were assigned. Hourly rates for different occupation groups involved were derived from total labor costs. No costs invoiced to third parties, e.g., pass through costs for imaging or laboratory tests were included. Results were used to design a cost calculation tool. The resulting tool underwent round robin tests by having study coordinators from various clinical disciplines calculate time expenditures based on the same study protocol and case report forms to validate the tool. In addition, study coordinators of one site calculated time expenditure of all trials initiated and tracked time prospectively over a period of 12 months to assess the predictive value of the tool.

Results: The study site budgeting tool (STUDGET) is a web-based application determining hours of work on a trial and calculates responding hourly rates of staff, totaling them to fees required to properly conduct the trial. By improving the accuracy of STUDGET in round robin tests, we achieved a median deviation of 371.59 (range 43.58 - 1,152.08) in calculated case payments (reference 1,671.97). Comparison of predicted and actual hours showed a correlation of 105% median (range 18% - 228%). Outliers were due to unforeseeable changes in trial execution.

Conclusion: We developed the web-based tool STUDGET allowing trial sites to determine the case fee needed to conduct a clinical trial.

A79

ONLINE REQUEST SYSTEM FOR DATASETS, WRITING GROUPS, AND ANCILLARY STUDIES FOR THE SEARCH FOR DIABETES IN YOUTH STUDY

Jennifer W. Talton, John Hepler, Jeanette S. Andrews, Kenneth Wilson, Donna Kronner, Ronny A. Bell, Ralph B. D’Agostino Jr.

Wake Forest School of Medicine

Wake Forest School of Medicine serves as the Coordinating Center (CoC) for the SEARCH for Diabetes in Youth Study (SEARCH). SEARCH is a multi-center observational study that began conducting population-based ascertainment of cases of non-gestational diabetes in youth less than 20 years of age beginning in 2001 and continuing through the present. In addition to identifying cases, the project has a prospective cohort component that involves follow-up visits. Most of the roughly 60 active investigators who are looking to analyze the large amounts of data SEARCH collects choose to create writing groups, which include a statistician from the CoC to conduct the analyses. For investigators who prefer to analyze the data themselves, a CoC statistician provides them with a dataset. SEARCH also has 6 ancillary studies to the main project and new requests for ancillary studies continue to be made. In order to handle the large volume of requests (writing groups, datasets, ancillary studies), which are all made through the CoC, an online request system was designed which streamlines communication and documentation between the investigators and the CoC. The request system is a step-by-step wizard which guides the investigator through a series of questions and information screens that must be completed prior to sending the request to the CoC. For example, in the
dataset request system the investigator must have approval from the Papers and Presentations committee before the request can be submitted and they cannot request a dataset completion deadline that is less than one-month from the submission date. Once completed an e-mail is generated to both the CoC and investigator with the request attached in PDF form. This presentation will elaborate on the specifics of the system by describing additional features available; such as tracking the progress of the request (submitted, completed) which was created for documentation purposes.

A80

WHAT MAKES PEOPLE PARTICIPATE IN CLINICAL TRIALS? INSIGHTS FROM A META-ETHNOGRAPHY OF QUALITATIVE STUDIES
Sharon K McCann, Marion K Campbell, Vikki A Entwistle
Health Services Research Unit, University of Aberdeen

It is well known that recruitment to trials can be difficult. In recent years a number of in-depth qualitative studies have been published that have examined patient experiences of recruitment and participation in trials. To understand the collective insights from these studies, we undertook a meta-ethnography (a formal synthesis method which centres on the construction of interpretations cumulatively across qualitative studies) of qualitative studies published between 1996 and 2010. To be included all studies had to focus on people’s own accounts of their decisions to accept or decline trial participation. Our synthesis highlighted how key aspects of context, recruitment approach and person approached can interact to influence trial recruitment. In particular, the way potential participants were situated in terms of their health states and treatment junctures (and perceptions about these at the time of trial recruitment) was particularly salient. Their sense of their situation at the time of being approached about trial participation influenced their judgements, particularly around the implications of trial participation for the common good (their desire to help others) and their own identity in relation to participating (what their non/participation might say about them). It could also mediate the influence of their communication and relationship with trial recruiters and of the nature of the trial interventions and processes. These insights led to the development of a conceptual model of factors likely to affect trial recruitment. The results of the synthesis and model of factors likely to affect recruitment will be presented together with strategies trialists might adopt using the insights they offer.

A81

CHALLENGES OF COMPOSITE ENDPOINTS FOR TRIAL EXECUTION, STATISTICAL ANALYSIS AND INTERPRETATION
Joerg Hasford
Ludwig-Maximilians-Universitaet Muenchen Institut fuer Med. Informationsverarbeitung, Biometrie und Epidemiology

Recent years have seen an exceptional improvement of overall survival (OS) in patients with chronic myeloid leukemia. Thus classical endpoints like OS would need either huge sample sizes or very long observation time. Composite endpoints, combining OS with events like progression, not achieving complete cytogenetic or molecular remission within a certain treatment period (typically 12 or 18 months) and loss of remission, called Progression-free Survival (PFS) or Failure-free survival (FFS), were introduced to allow for a faster evaluation of new treatments. FFS subsumes ~ 10 singleton events ranging from death to loss of response. FFS data are thus difficult to interpret and provide limited value for informing clinical decisions as the relevance of the singleton endpoints is very heterogeneous. Hence drug authorities ask that CE do not mix disease outcomes with clinical management decisions. An often neglected issue is that intention to treat analysis of CE requires that all singleton events have been properly monitored and assessed at the time points specified in advance in the trial protocol. Thus in a large CML trial out of 1466 evaluable pts only 571 were available for the analysis of FFS (Lauseker M et al. ELN Frontiers Meeting 2012). Including in the statistical analysis patients with missing information on singleton endpoints may seriously bias results. Yet another problem is the common lack of widely accepted uniform definitions of CEs precluding a comparative assessment across trials. The challenging issues of CEs will be exemplarily discussed on the basis of two recently published CML-trials (Saglio G, et al. NEJM 2010;362:2251-2259, and Kantarjian HM, et al. NEJM 2010;362:2260-2270). There is an urgent need for consented CEs and alternatives to CEs.
While p-values may be the key to ultimate study success, they give very little information about the data collected. Confidence intervals are an improvement because they communicate both value and variability, but there is often much more to be learned from our data if we go beyond numerical summaries. Patterns, trends, and relationships can get lost in numeric summary tables but will jump off the page when shown with the appropriate figure. Much can be done with the basic go-to favorites of histograms, scatterplots, boxplots, and Kaplan-Meier survival curves. This presentation, designed for non-statisticians, takes a look at simple twists on these old favorites with examples of what they can offer to give a broader look at your data through exploratory data analysis. Possibilities beyond histograms, scatterplots, boxplots, and Kaplan-Meier curves are touched on as well.

Mobile computing is becoming ubiquitous, with even toddlers (http://abcnews.go.com/WNT/video/ipad-effect-toddlers-14813787) and seniors (http://www.lvrj.com/business/more-seniors-get-savvy-with-smartphones-155271575.html) using and navigating mobile devices. What does the explosion is smartphone and tablet ownership and usage mean for clinical research? Use of handheld devices to collect data from subjects is commonplace, and the use of tablet computers to capture source data, whether from investigators or patients, is becoming more common. As smartphones and tablets grow in popularity, even penetrating areas of the globe where consistent electricity is problematic, researchers have the opportunity communicate directly with patients. A recent direct-to-subject study highlighted somethe opportunities available, but also pointed out some limitations not addressed by technologies. This talk will review the technologies applicable for direct-to-subject communication - traditional ePRO, web-based communications, IVR, SMS text messaging, smartphone apps- and specialized technologies that capture biological measurements directly from patients. The appropriateness and limitations of each for various clinical research scenarios and therapeutic areas will be explored. Finally, some crystal-ball gazing will be indulged, examining opportunities and possibilities generated by mobile technologies that have been heretofore unrealistic.

The NLST recruited 53,454 participants with histories of heavy smoking and at high risk for developing lung cancer to determine whether image-screening reduces lung-cancer mortality. Participants were randomized to Computed-Tomography (CT) or chest-X-ray arms and received up to 3 screens each, at annual intervals. NLST findings (announced 2011) suggest that CT screening does reduce lung-cancer mortality while chest-X-ray screening does not. NLST non-image data are maintained by Information Management Services (IMS). The CT images (~20 million) are a collection of The Cancer Imaging Archive (TCIA), managed by Washington University (WU). Data and images are available to NLST investigators with approved research plans. In an effort to stimulate and facilitate post-trial secondary research, a WU-IMS partnership created a PostgreSQL-based Query Tool (QT) that allows investigators to quickly search NLST trial data by building their own queries using simple tree menus. QT selection-menus consist of tables (13) and variables (300+) that may be
conditioned with yes/no, numerical-range, and pick-list values. Results may be saved to files suitable for statistical analysis. Queries may be saved for later re-use. Hover-over any variable displays its tool-tip data-dictionary-paraphrase. A Help drop-down-menu offers a complete data dictionary, QT User Guide, and QT Tutorial. The QT offers the option to download the CT images associated with query results. If an investigator chooses to download images immediately, the QT launches the image-downloader of the National Biomedical Imaging Archive (NBIA) application, a key component of TCIA. Alternatively, the investigator may choose to save the image pointers as an NBIA shared-list; the investigator and any collaborator may access this shared-list at any time to download the list images. QT allows researchers to explore completed-trial data and download images for, say, retrospective studies or computer-aided-detection algorithm development. QT is open-source and easily configured for other trials, with or without images. Reference: https://biometry.nci.nih.gov/cdas/studies/nlst

A85

THE NIH HEALTH CARE SYSTEMS RESEARCH COLLABORATORY DISTRIBUTED RESEARCH NETWORK

(1) Jeffrey Brown, (2) Lesley Curtis, (3) Kimberly Lane, (4) Beth Syat, (5) Richard Platt

(1) Department of Population Medicine of Harvard Medical School and Harvard Pilgrim Health Care Institute
(2) Duke University School of Medicine
(3) Department of Population Medicine of Harvard Medical School and Harvard Pilgrim Health Care Institute
(4) Department of Population Medicine of Harvard Medical School and Harvard Pilgrim Health Care Institute
(5) Department of Population Medicine of Harvard Medical School and Harvard Pilgrim Health Care Institute

Background: In September 2012, the NIH funded the Health Care Systems Research Collaboratory to engage health care systems as research partners in conducting multi-center pragmatic and cluster randomized clinical studies. The Collaboratory consists of a Coordinating Center and, initially, seven health system-based clinical trial demonstration projects. The Collaboratory is developing broadly applicable tools, methods, and policies to support these activities. One aspect of this work involves developing infrastructure and methods for the use of multi-center electronic clinical, administrative, and research data. Scope The Collaboratory will create the NIH Distributed Research Network to facilitate collaborative clinical research by allowing organizations to quickly and easily share information with research partners without disclosing protected health information or proprietary data. The NIH Distributed Research Network will enable organizations holding clinical or research datasets to 1) make detailed information (metadata) about clinical data resources and existing data sets available and 2) enable secure distributed querying of data that remain in the physical control of the clinical system or investigator that created it. In the parlance of the Office of the National Coordinator’s (ONC) Query Health program, the NIH Distributed Research Network will enable investigators to ‘send the question to the data’ instead of creating large pooled datasets.

Conclusions: The NIH Distributed Research Network will provide a new capability for multi-site clinical investigation by enabling secure distributed querying of clinical and research data without requiring disclosure of confidential or proprietary data.

A86

USING A WEBPORTAL TO ENHANCE NETWORK COMMUNICATION

(1) Laurie McLeod, (2) Stephanie Wellford, (3) Emily Wilson

(1) Rho (2) Rho (3) Rho

In federally-funded projects, Rho is responsible for managing and coordinating communication for large, multi-center collaborative research efforts. Communication within these large networks can be a challenge due to the number of people involved, the volume of documents and versions distributed, and varying schedules and time zones. Over the last 10 years, communication in clinical research has changed dramatically with technological advances. Direct mail and facsimile have been replaced with more immediate and reliable forms of communication. Rho adapted by creating secure network webportals to disseminate current, readily available information to all network members in an efficient manner. The network webportals house general information, study documents, documents in development, and meeting materials for stakeholders to reference. Managing study and meeting materials real-time and in a central location allows Rho to easily direct stakeholders to information or documents, ensures that all stakeholders are accessing the
same version of materials, and reduces the volume and size of documents in email inboxes. Collaborative abstracts, manuscripts, and other working documents are easily edited and re-posted through a virtual check-out system, simplifying version control and streamlining the review process. The webportals are secured through a log-in system, which requires stakeholders to enter a username and password before accessing any content. Rho manages user permissions to guarantee each stakeholder can view the information that relates to their scope of work and interest. These permissions also can be used to allow for distribution of materials to an exclusive group. Our experience and feedback from stakeholders indicate that other multi-center programs could benefit from using network webportals as an information hub to provide network members with current information and documents in a structured, easily accessible form.

A87
DEVELOPMENT OF DATA DIAMOND REPORTING SYSTEM
Sherrie Whisler
Pennsylvania Academy of Family Physicians and Foundation

The Pennsylvania Academy of Family Physicians and Foundation developed a quality data reporting website with grant funds from the PA Department of Health (PADOH). Phase I development was fast-tracked, starting January 2012 and in production by June 2012. Primary care practices using the site report data on Diabetes, Asthma, IVD, HTN, and Preventative measures and were trained and using the site by August 2012. The project is called Data Diamond, and its primary design was to establish a website for data reporting in a very dynamic, highly efficient, and scalable way. Using SQL Server 2008, Microsoft .net, and LogiXML, we quickly produce high quality data reports for funders, including the PADOH, and practices. Development of the site allowed us, as an organization, to house multiple projects on one site and establish access to the site for the users based on project permissions. Based on the data model of more than 50 entities, all data entry and reports drive the system’s performance and allow the administrator to define project, practice, and measure in minutes. Subsequent web pages appear instantly or in seconds. User Roles define which areas users can access based on their role in the project. Currently we have more than 4,500 defined users, such as consultant, provider, leader, faculty, and physician champion. The new system easily produces dynamic sets of data reports. Analysis is the core of the Data Diamond website. Stakeholders get an overview of data and can drill down as far as the data allows for more specific information. Phase II development is underway and we plan to develop a patient-specific registry to track data on individuals rather than just total population. With the new Activity/Event tracker, users will be reminded of Learning Sessions, CME Events and register for events online from the system.

A88
SIMULATION-BASED EVALUATIONS OF METHODOLOGICAL ALTERNATIVES FOR RECYCLING RANDOMIZATION ASSIGNMENTS IN RANDOMIZED CLINICAL TRIALS (RCTS) OF ACUTE CONDITIONS (AC) WITH BRIEF FOLLOW UP
(1) Mohamed Mubasher*, (2) Abdullah Alangari, (1) Mohammad Al-Tannir, (1) Imad Tlejeyjeh, (1) Sameeh Ghazal
(1) Clinical Research Center, King Fahad Medical City (2) Department of Pediatrics, College of Medicine, King Saud University

Background: In RCTs of AC (pediatric asthma, flu-like symptoms, etc.) subjects are randomized and kept under brief observation (e.g., ER setting or outpatient clinic) to determine the study outcome(s). If future episodes recur, a subset of these subjects might become eligible. Considerations to re-recruit such subjects constitute a challenging methodological task from randomization schematic perspective and future data analysis.

Objectives: To present and evaluate alternatives for re-randomization schemes in AC RCTs.

Methods: Assume in a 2-arm AC RCT, n1 and n2 subjects (n1 + n2 = m) are randomized according to a 1:1 permuted block design (PBD) and then evaluated for the study outcomes. Let n11 and n22 of the n1 and n2, become eligible. We investigated 3 scenarios to re-randomize these subjects: 1) According to the ongoing PBD as independent randomization units, 2) Crossing them over to the other arm according to a preset scheme and 3) Re-enroll them according to their original assignments and treat their outcomes as repeated measures. The distribution of prognostic characteristics was compared between study groups using conditional dichotomous logistic regression to determine the adequacy of the 3 re-randomization scenarios.
100,000 simulations of randomization situations were generated; m = 300 (100) 1000; n ij; i < j =1,2 were selected randomly between n i /4 and n i /2. A non-homogeneous Poisson process was used to generate subjects’ entry. Five dichotomous prognostic factors were simulated with variable proportions; P i = 0.10 (0.20) 0.50 but equal distribution between the two groups.

**Results:** For the five prognostic factors: 1. All scenarios produced variable but adequate probabilities of between-groups balance 2. The likelihood of balance was highest under the cross-over scenario, followed by the scheme of retaining the same randomization assignment. Conclusions Cross-over schemes seem to be the most adequate for re-randomization in such RCTs.

---

**A89**

THE CAMEROON MOBILE PHONE SMS (CAMPS) TRIAL: A RANDOMIZED TRIAL OF TEXT MESSAGING VERSUS USUAL CARE FOR ADHERENCE TO ANTIRETROVIRAL THERAPY

(1) Lawrence Mbuagbaw, (2) Lehana Thabane, (3) Pierre Ongolo-Zogo, (4) Richard T Lester, (5) Edward J Mills, (6) Marek Smieja, (7) Lisa Dolovich, (8) Charles Kouanfack

(1) Centre for the Development of Best Practices in Health (CDBPH), Yaoundé Central Hospital/ Department of Clinical Epidemiology and Biostatistics, McMaster University (2) Department of Clinical Epidemiology and Biostatistics, McMaster University/ Biostatistics Unit, Father Sean O’Sullivan Research Centre, St Joseph’s Healthcare (3) Centre for the Development of Best Practices in Health (CDBPH), Yaoundé Central Hospital (4) Department of Medicine, Division of Infectious Diseases, University of British Columbia/ British Columbia Centre for Disease Control (5) Faculty of Health Sciences, University of Ottawa (6) Department of Clinical Epidemiology and Biostatistics, McMaster University/ St. Joseph’s Healthcare (7) Department of Family Medicine, McMaster University, McMaster Innovation Park (8) Accredited Treatment Centre, Yaoundé Central Hospital

**Background:** Mobile phone technology is a novel way of delivering health care and improving health outcomes. This trial investigates the use of motivational mobile phone text messages (SMS) to improve adherence to antiretroviral therapy (ART) over six months.

**Methodology/Principal findings:** CAMPS was a single-site randomized two-arm parallel design trial in Yaoundé, Cameroon. We enrolled and randomized HIV-positive adults on ART, aged 21 years and above to receive a weekly standardized motivational text message versus usual care alone. The primary outcome was adherence measured using a visual analogue scale (VAS), number of doses missed (in the week preceding the interview) and pharmacy refill data. Outcomes were measured at 3 and 6 months. Service providers and outcome assessors were blinded to allocation. Analysis was by intention-to-treat. Between November and December 2010, 200 participants were randomized, with 101 in the intervention group and 99 in the control group. At 6 months, overall retention was 81.5%. We found no significant effect on adherence by VAS > 95% (risk ratio [RR] 1.06, 95% confidence interval [CI] 0.89, 1.29; p= 0.542; reported missed doses (RR 1.01, 95% CI 0.87, 1.16; p>0.999) or number of pharmacy refills (mean difference [MD] 0.1, 95% CI: 0.23, 0.43; p=0.617. One participant in the intervention arm reported a possible disclosure of status.

**Conclusions/Significance:** Standardized motivational mobile phone text messages did not significantly improve adherence to ART in this study. This study was different from other studies in terms of its duration, sample size and duration of participants on ART. Other types of messaging or longer term studies are recommended.

---

**A90**

PRACTICAL CONSIDERATIONS ON MEDICAL DEVICE CLINICAL TRIALS IN GEORGIA, A FORMER SOVIET UNION COUNTRY

(1) Giada Morpurgo, (2) Tamar Kemoklidze, (3) Marco Ghidina, (4) Federica Zobec

(1) Zeta Research Ltd (2) CCRG Ltd (3) Zeta Research Ltd (4) Zeta Research Ltd

Budget and time restrictions often lead clinical trials to be conducted outside European borders. Two post market clinical trials on a medical device have been conducted in 2012 in Georgia, a former Soviet Union country geographically and culturally very close to Europe. The investigational medical device was a class I ophthalmic solution for allergic conjunctivitis. The trials were structured on different populations (pediatric and adult) and with different design (randomized placebo-controlled, and single arm). Many direct and indirect economic advantages has been registered, basically based on the fact that Georgia offers small and
flexible private clinical structures with diffused high speed internet and qualified medical personnel, even if slight differences in ordinary medical practices can be met. Legislation’s principles follow the GCP and international harmonization guidelines but the practical management of a clinical trial is strongly favored by quick authorization procedures (no formal approval by the MoH is required for medical device trials) and flexible session calendars of the ethical review boards, with great time saving compared to European average standards. Contracts and budget management with sites and investigators too seems to be cheaper and strongly simplified. Lack of web-available information about procedures and regulation, investigational products import custom procedures and the need for all the documentation to be translated in local Georgian language makes a local CRO support precious and strongly recommended. Cultural difference mostly with the older generations appear in the patient involvement in the study: as no reimbursement of expensive biotech and other products is provided, Georgian patients are globally looking at clinical trials as an opportunity to receive the latest medical treatments, but patients’ retention in the study seems to be a critical aspect as patients tend to fail in follow up visits after reaching benefits from the treatment.

A91

A NEW TRIAL DESIGN FULLY INTEGRATING BIOMARKER INFORMATION FOR THE EVALUATION OF TREATMENT-EFFECT MECHANISMS IN PERSONALISED MEDICINE
(1) Richard Emsley, (2) Hanhua Liu, (3) Sabine Landau, (4) Graham Dunn
(1) Centre for Biostatistics, The University of Manchester (2) Centre for Biostatistics, The University of Manchester (3) Institute of Psychiatry, King’s College London (4) Centre for Biostatistics, The University of Manchester

The development of personalised (stratified) medicine is intrinsically dependent on an understanding of treatment-effect mechanisms (effects on therapeutic targets that mediate the effect of the treatment on clinical outcomes). There is a need for novel clinical trial designs for the joint evaluation of treatment efficacy, the utility of predictive markers as indicators of treatment efficacy, and the meditational mechanisms proposed as the explanation of these effects. We review the problem of confounding (common causes) for the drawing of valid inferences concerning treatment-effect mechanisms, even when the data has been generated using a randomised controlled trial. We illustrate the potential of the predictive biomarker-stratified design, together with baseline measurement of all known prognostic markers, to enable us to evaluate both the utility of the predictive biomarker in such a stratification and, perhaps more importantly, to estimate how much of the treatment’s effect is actually explained by changes in the putative mediator. We call this a biomarker stratified efficacy and mechanisms evaluation (BS-EME) trial. The analysis strategy involves the use of instrumental variable regression, using the treatment by predictive biomarker interaction as an instrumental variable - a refined, much more powerful and subtle use of Mendelian randomisation. Stratification without corresponding mechanisms evaluation lacks credibility and in the almost certain presence of mediator-outcome confounding, mechanisms evaluation is dependent on stratification for its validity. Using Monte Carlo simulation we show that both stratification and treatment-effect mediation can be evaluated using the BS-EME trial design together with detailed baseline measurement of all known prognostic biomarkers and other prognostic covariates. The direct and indirect (mediated) effects are estimated through the use of instrumental variable methods together with adjustments for all know prognostic markers (confounders) the latter adjustments contributing to increased precision (as in a conventional analysis of treatment effects) rather than bias reduction.

A92

WEB-BASED INTERVENTION MODELLING EXPERIMENTS: A WAY OF EXPLORING PROFESSIONAL BEHAVIOUR CHANGE INTERVENTIONS BEFORE A FULL-SCALE TRIAL
Shaun Treweek, Debbie Bonetti, Karen Barnett, Martin Eccles, Jill Francis, Claire Jones, Graeme MacLennan, Nigel Pitts, Ian Ricketts, Frank Sullivan and Mark Weal
(1) University of Aberdeen; (2) University of Dundee; (3) University of Edinburgh; (4) University of Newcastle; (5) City University; (6) University of Dundee; (7) University of Aberdeen; (8) King’s College; (9) University of Dundee; (10) University of Dundee; (11) University of Southampton

Background: An intervention modelling experiment (IME) is a theory-based technique that can guide the choice of professional change interventions by simulating encounters between clinicians and patients prior
Abstracts

S43

http://ctj.sagepub.com

Clinical Trials 2013; 10: S1 – S88

To entering the intervention into a full-scale trial. To date IMEs have all been paper-based. To test the IME methodology we replicated an earlier, paper-based IME of antibiotic prescribing by family doctors for upper respiratory tract infections using web-delivery. If the IME methodology is robust, the same results should be seen in both IMEs.

**Method:** The study was a two-stage, national trial targeting family doctors. In the larger stage, Stage 1, each participating doctor received a web-based questionnaire and clinical scenarios, which were designed to identify predictors of doctors’ antibiotic prescribing behaviour. Stage 2 delivered a second questionnaire and set of scenarios to evaluate the effect of two, web-delivered interventions on doctors’ prescribing behaviour, one taken from the paper-IME, the other developed from Stage 1, compared to a ‘no information’ comparator.

**Results:** 270 doctors took part in Stage 1. Eight of ten predictors of prescribing behaviour were in agreement between the paper and web-based IMEs. From these predictors, a new theory-based intervention based on Action Planning was developed. Stage 2 involved 131 GPs and those receiving the interventions showed increases in the number of scenarios without an antibiotic prescription (Action Plan=0.82 (95% CI=0.26 to 1.37); Persuasive Communication=0.73 (95% CI=0.14 to 1.31). This was in agreement with the paper-based IME.

**Conclusions:** The IME methodology is robust and can be effectively delivered over the web to develop and explore behaviour change interventions prior to a full scale trial. This is likely to lead to interventions with a greater potential for effecting the desired behaviour change.

A93

RANDOMIZING TWO STUDY EYES FROM THE SAME PARTICIPANT IN OPHTHALMIC CLINICAL TRIALS

(1) Adam R. Glassman, (2) B. Michele Melia
(1,2) Jaeb Center for Health Research

There are many examples of ophthalmology trials enrolling only one eye or requiring enrollment of both eyes, but few allow participants with one or two study eyes depending on the number eligible. The Diabetic Retinopathy Clinical Research Network (DRCR.net) has conducted several clinical trials that have allowed but not required both eyes of participants to be randomized. This design increases the potential pool of eligible eyes. When enrollment of both eyes is allowed, eyes may be randomly assigned to the same or different treatments with advantages and disadvantages to each approach. Randomizing one vs. two eyes: Assuming positive but not perfect correlation between the participant’s two eyes, there is greater power for a fixed number of participants when participants can contribute two eyes. The increase in power depends on the correlation and the proportion of patients with two study eyes and is presented in another abstract (Melia and Glassman). This can result in substantial savings on enrollment time and study cost. However, there may be barriers to enrolling both eyes including logistical complexities, safety concerns, or low frequency rate of bilateral disease. Randomization to different or same treatment groups: When eyes are randomized to different treatment groups, potential confounding with patient-level systemic factors is controlled for, decreasing variability and increasing statistical power. This approach may not be recommended if there is expected to be a cross-over effect in the contra-lateral eye, if attributing systemic adverse events to a particular drug is important, if masking is not feasible and the outcome is subjective, or quality of life assessment is important. Careful consideration should be given when designing a study where multiple measurements per participant are possible. If supported by the trial design, including and analyzing both eyes from the same participant can save time and money.

A94

POWER FOR THE PARTIALLY-PAIRED RANDOMIZATION DESIGN AS COMPARED WITH THE UNPAIRED DESIGN IN OPHTHALMOLOGICAL CLINICAL TRIALS

Michele Melia, Adam Glassman
Diabetic Retinopathy Clinical Research Network (Jaeb Center for Health Research)

**Background:** When designing an ophthalmological trial, there frequently is a choice of designs: (1) randomize 1 eye, (2) require and randomize 2 eyes, or (3) randomize one eye when one eye is eligible and randomize both eyes when both eyes are eligible. In the latter 2 designs, eyes within a subject may be randomized to the same treatment or to different treatments. Each design has its advantages and disadvantages as
presented in another abstract (Glassman and Melia). We explored the impact of proportion of subjects with 2 eyes and correlation between eyes on power for design 3, when eyes are randomized to different treatments, and compared with design 1.

**Methods:** We simulated data for change in visual acuity, a common outcome in ophthalmology trials, for 10,000 trials of varying number of treatment arms, sample size, proportion with 2 eyes, correlation between eyes, and treatment effects. Simulated trials were analyzed using a linear mixed model to account for correlation and obtain estimates of power for treatment comparisons.

**Results:** As expected, power of the partially-paired design was higher than that of the unpaired design for fixed sample size with positive correlation between eyes. Power of the partially-paired design was lower than the unpaired design when correlation was negative. Power gain for the partially-paired design was modest for the proportion of 2 eyes (0.20) and correlation between eyes (0.25) seen most frequently in our own clinical trials, e.g. 5% for detecting 3 letter difference with 150 subjects per group, but was considerably higher when there was greater correlation between eyes, e.g. 0.5 and higher, and/or when more subjects had 2 eligible eyes.

**Conclusions:** Our results allowed us to quantify the statistical gains of using the partially-paired design over the unpaired design which can be used to make better-informed decisions when choosing a design.

---

**AN EVALUATION OF INFERENTIAL PROCEDURES FOR ADAPTIVE CLINICAL TRIAL DESIGNS WITH PRE-SPECIFIED RULES FOR MODIFYING THE SAMPLE SIZE**

(1) Gregory Levin, (2) Sarah Emerson, (3) Scott Emerson

(1) University of Washington (2) Oregon State University (3) University of Washington

Many papers have introduced adaptive clinical trial methods that allow modifications to the sample size based on interim estimates of treatment effect. There has been extensive commentary on type I error control and efficiency considerations, but little research on estimation after an adaptive hypothesis test. We evaluate the reliability and precision of different inferential procedures in the presence of an adaptive design with pre-specified rules for modifying the sampling plan. We extend group sequential orderings of the outcome space based on the stage at stopping, likelihood ratio test statistic, and sample mean to the adaptive setting in order to compute median-unbiased point estimates, exact confidence intervals, and P-values uniformly distributed under the null hypothesis. The likelihood ratio ordering is found to average shorter confidence intervals and produce higher probabilities of P-values below important thresholds than alternative approaches. The bias adjusted mean demonstrates the lowest mean squared error among candidate point estimates. A conditional error-based approach in the literature has the benefit of being the only method that accommodates unplanned adaptations. We compare the performance of this and other methods in order to quantify the cost of failing to plan ahead in settings where adaptations could realistically be pre-specified at the design stage. We find the cost to be meaningful for all designs and treatment effects considered, and to be substantial for designs frequently proposed in the literature.

---

**EVALUATING THE IMPACT OF INCREASING THE SAMPLE SIZE IN THE SECONDARY PREVENTION OF SMALL SUBCORTICAL STROKES (SPS3) TRIAL: HINDSIGHT IS 20/20**

(1) Leslie A. McClure, (2) Jeff Szychowski, (3) Christopher S. Coffey

(1) University of Alabama at Birmingham, Department of Biostatistics; (2) University of Alabama at Birmingham, Department of Biostatistics; (3) University of Iowa, Department of Biostatistics

The NIH-funded SPS3 study assessed whether combination antiplatelet therapy (AC) was superior to aspirin alone (AP), and whether lower blood pressure (BP) control was superior to ‘usual’ BP control in preventing recurrent stroke among patients with a lacunar stroke at baseline. Due to lower than expected overall rates of recurrent stroke, the sample size was increased from 2500 to 3000 to ensure adequate power to detect a 25% reduction in recurrent stroke in each arm. The antiplatelet trial was terminated early due to lack of efficacy coupled with safety concerns. Retrospectively, we assessed the gains achieved by the sample size re-estimation by comparing the difference in stroke recurrence among the first 2500 patients recruited, assuming follow-up ended when the antiplatelet arm terminated to results obtained with the complete 3020 patients. Among the first 2500 patients, we observed 248 events, over an average of 3.9 years follow-up.
The recurrent stroke rate for the AC and AP groups was 2.5% and 2.7% per year, respectively. The hazard ratio for AC vs. AP was 0.92 (95% CI: 0.72-1.18). Including the additional 520, only 15 more patients had primary events, and recurrent stroke rates (2.7%, and 2.5% year), and the HR of 0.92 (95% CI: 0.72-1.16) were similar. Despite an increase in sample size in SPS3, there was no appreciable difference in the results obtained for the antiplatelet trial had the study adjustment not been made; however, there is potential for gain to be seen in the BP arm. While sample size re-estimation is an important tool for ensuring adequate power for detecting treatment effects of interest, there is a potential to add unnecessary cost and complexity when no effect exists. More post-hoc examinations of studies using adaptive designs are warranted to fully understand their strengths and weaknesses.

A97

A SYSTEMATIC REVIEW OF METHODS FOR SPECIFYING THE TARGET DIFFERENCE IN RANDOMISED CONTROLLED TRIALS (DELTA REVIEW)
(1) Jonathan Cook, (2) Jenni Hislop, (3) Temitope Adewuyi, (4) Cynthia Fraser, (5) Kirsten Harrild,
(6) Douglas Altman, (7) Craig Ramsay, (8) Peter Fayers, (9) Andrew Briggs, (10) John Norrie,
(11) Ian Harvey, (12) Brian Buckley, (13) Luke Vale for the DELTA group
(1) University of Aberdeen (2) University of Aberdeen (3) University of Aberdeen (4) University of Aberdeen
(5) University of Aberdeen (6) University of Oxford (7) University of Aberdeen (8) University of Aberdeen
(9) University of Glasgow (10) University of Aberdeen (11) University of East Anglia (12) National University of Ireland (13) University of Newcastle

Background: Randomised controlled trials (RCTs) are widely accepted as the preferred study design for evaluating healthcare interventions. When determining sample size, a (target) difference must typically be specified. This provides reassurance that the study will be informative i.e. should such a difference exist, it is likely to be detected with the required statistical precision. From both a scientific and ethical standpoint, selecting an appropriate target difference is of crucial importance; too large or small a study is arguable unethical, wasteful and potentially misleading. While a variety of methods have been proposed in the medical field to specify a target difference, their relative merits are unclear as is whether there are methods in the non-medical field which could be used. Aim: To systematically review of medical and non-medical literature to identify methods for specifying the target difference in a randomised controlled trial.

Methods: Electronic searches of biomedical and non-medical databases were performed. Clinical Trial textbooks were also reviewed. Titles and abstracts were screened prior to full-text assessment. Studies which reported a method which could be used to specify the target difference were included. Clinical trial textbooks were also reviewed.

Results: The search identified 11,485 potentially relevant studies; 1,434 were selected for full-text assessment. Seven methods were identified: anchor, distribution, health economic, opinion-seeking, pilot study, reviews of the evidence base and standardised effect size (SES). The anchor, distribution and SES methods were most commonly used.

Discussion: A variety of methods were identified though each had important variations in application. While no single method provides a perfect solution to a difficult question, methods are available to inform specification of the target difference and should be used whenever appropriate.

A98

SAMPLE SIZE AND SCREENING SIZE TRADE OFF IN THE PRESENCE OF SUBGROUPS WITH DIFFERENT EXPECTED TREATMENT EFFECTS
(1) Kyle Rudser, (2) Edward Bendert, (3) Joseph Koopmeiners
(1) Division of Biostatistics, University of Minnesota (2) Statistics Collaborative (3) Division of Biostatistics, University of Minnesota

Statistical study design considerations typically focus on sample size, power, and a single population treatment effect given a fixed significance level (generally 0.05). Eligibility criteria is formulated to select the patient population of interest to be studied for which the magnitude of the treatment effect is expected to hold. In some instances researchers may expect a larger treatment effect in one subgroup while others exhibit an attenuated effect. Identification of these subgroups can be based on a clinical decision rule, e.g.,
Biomarker cutoff, but may not be precise, i.e., sensitivity and specificity are not simultaneously at 100%. In these settings there can be a trade-off between a smaller average treatment effect with broader enrollment criteria and a larger effect with restricted criteria but longer enrollment duration. We evaluate the impact of including attenuated subgroups on design operating characteristics and illustrate scenarios where overall trial enrollment duration, and hence time to complete the study, may be shorter by not being restrictive.

A99
IMPACT OF SUBJECT ATTRITION ON SAMPLE SIZE DETERMINATIONS FOR LONGITUDINAL CLUSTER RANDOMIZED CLINICAL TRIALS
Moonseong Heo
Albert Einstein College of Medicine

Subject attrition is a ubiquitous problem in any type of clinical trials and thus needs to be taken into consideration at the design stage particularly to secure adequate statistical power. Here, we focus on longitudinal cluster randomized clinical trials (cluster-RCT) that aim to test the hypothesis that an intervention has an effect on the rate of change in the outcome over time. In this setting, the cluster-RCT assumes a three level hierarchical data structure in which subjects are nested within a higher level unit such as clinics and are evaluated for outcome repeatedly over the study period. Furthermore, the subject-specific slopes can be modeled in terms of fixed or random coefficients in a mixed-effects linear model. Closed form sample size formulas for testing the hypothesis above have been developed under assumption of no attrition. In this paper, we propose closed form approximate samples size determinations with anticipated attrition rates by modifying those existing sample size formulas. Specifically, the two parameters in existing formulas that would be affected by subject attrition are the number and the variance of the assessment time points per subject. Our sample size determination strategy is to replace those two parameters by their corresponding expected number and variance under anticipated attrition rates. With extensive simulations, we examine performances of the modified approximate formulas considering the following factors: fixed and random slope models; different attrition rates; monotone attrition process; different types of distributions of attrition time points; and three attrition mechanisms: attrition completely at random, attrition at random and attrition not at random. In conclusion, the proposed modification is very effective under fixed slope models but yields biased, if not substantially, statistical power under random slope models.
P01
THE EVOLUTION OF THE CRF DESIGNER INTO A DATA MANAGER
Christina Tittsworth
SGS Life Science Services

The role of the Case Report Form (CRF) designer has historically been independent; only needing to extract elements defined in the protocol to create CRFs. Now, the CRF designer is more data management and data standards oriented as clinical trials move to electronic data capture (EDC) systems. In the future, the “role” of the CRF designer will be pivotal in study and data management, applying standards for data collection, analysis, and sharing. Past: Relied solely on protocol for extracting CRF elements, Understood flow of data capture. Present: The CRF designer is mindful of the complexity of translating paper CRFs to eCRFs and understands: Intuitive and simple CRF design/508 compliance, Data standards and controlled terminology, Data analysis. Future: Maintains CRF library for faster study startup/quality control, Implements data standards with analysis and sharing in mind, Cost efficient, eliminates need to create paper spec to mimic eCRF As the electronic age pervades clinical trials, the role of CRF designer will transcend into clinical trials data manager. By recognizing important elements in the protocol, and using data management and data standards to streamline analysis elements, meaningful study results will be achieved.

P02
NON-SURGICAL PERIODONTAL THERAPY REDUCES CORONARY HEART DISEASE RISK MARKERS: A RANDOMIZED CONTROLLED TRIAL
Syed Akhtar Bohkari, Ayyaz A Khan, Arshad K Butt, Mohammad Azhar, Mohammad Hanif, Mateen Izhar, Dimitris N Tatakis

The University of Faisalabad, Dentistry

Aim: Periodontal disease elevates systemic inflammatory markers strongly associated with coronary heart disease (CHD) risk. The aim of this randomized controlled trial was to investigate the effect of non-surgical periodontal therapy on systemic C-reactive protein (CRP), fibrinogen and white blood cells in CHD patients. Materials and Methods: Angiographically proven CHD patients with periodontitis (n = 317) were randomized to intervention (n = 212) or control group (n = 105). Primary outcome was reduction in serum CRP levels; secondary outcomes were reductions in fibrinogen and white blood cells. Periodontal treatment included scaling, root planing and oral hygiene instructions. Periodontal and systemic parameters were assessed at baseline and at 2-month follow-up. Intent-to-treat (ITT) analysis was performed.

Results: Study was completed by 246 subjects (intervention group = 161; control group = 85). Significant improvements in periodontal and systemic parameters were observed in intervention group. The number of subjects with CRP > 3mg/L in intervention group decreased by 38% and in control group increased by 4%. ITT analysis gave a significant ($\chi^2 = 4.381, p = 0.036$) absolute risk reduction of 12.5%.

Conclusion: In CHD patients with periodontitis, non-surgical mechanical periodontal therapy significantly reduced systemic levels of C-reactive protein, fibrinogen and white blood cells.

P03
HEART EXERCISE AND REMOTE TECHNOLOGIES (HEART): A RANDOMIZED CONTROLLED TRIAL
(1) Yannan Jiang, (2) Ralph Maddison, (3) Leila Pfaffl, (4) Robyn Whittaker, (5) Ralph Stewart, (6) Andrew Kerr, (7) Geoffrey Kira, (8) Karen Carter

(1)(2)(3)(4)(7)(8) National Institute for Health Innovation, The University of Auckland, Auckland, New Zealand (5) Auckland District Health Board, Auckland City Hospital, Auckland, New Zealand (6) Middlemore Hospital, Auckland, New Zealand

Cardiac rehabilitation (CR) can improve health behaviors to slow or reverse the progression of cardiovascular disease (CVD). Exercise is a central element of CR. Technologies such as mobile phones and the Internet (mHealth) offer potential to overcome the physical and geographical barriers associated with low levels of exercise participation. The HEART trial aimed to investigate the effects of a mHealth delivered exercise-based CR program to improve exercise capacity and functional outcomes at 24 weeks compared with usual CR care in New Zealand (NZ) adults with a diagnosis of CVD. A two-arm, single-blind, parallel randomized
controlled trial was conducted in 2011-12 [Maddison 2011]. Participants were cardiac outpatients identified from Auckland and Middlemore hospitals, who were clinically stable, could perform exercise, owned a mobile phone, and had access to the Internet. The intervention group received a personalized automated package of text messages delivered via mobile phone, and were supported with a personalized interactive website [Pfaeffli 2012]. The control group received usual care which involves encouragement to exercise and an offer to join a local cardiac club. A total of 171 eligible patients were randomized (intervention N=85; control N=86) and followed for 24 weeks. Participants were on average 60 years old (39-79 years), and were mostly never or ex-smokers (94%), NZ Europeans (77%), males (81%), married (79%), with a self-reported medical history of high blood pressure (52%), high cholesterol (74%), heart attack (74%), angina (50%), and other forms of heart disease (28%). The primary outcome was change in maximal oxygen uptake (VO2max) (ml.kg-1.min-1) from baseline. Six minute walk test, health-related quality of life (SF-36v2TM), self-efficacy, and self-reported physical activity (IPAQ) were also assessed. Treatment evaluations were performed on the principle of intention-to-treat. Multiple imputation method was employed with missing data on the primary outcome. Full trial results will be presented and discussed.

P04
WHISPERER: INTEGRATED TEAM APPROACH TO DECLARATIVE STUDY SPECIFICATION
(1) John Perkins, (2) Nicholas Best, (3) Kelly Cho

(1) VA Cooperative Studies Program, Massachusetts Veterans Epidemiology Research and Information Center, (MAVERIC), VA Boston Healthcare System (for authors 1, 2, and 3) (2) Division of Aging, Department of Medicine, Brigham and Women’s Hospital (for author 3)

Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC) supports data collection for clinical, observational, and biorepository studies. A declarative study specification includes all questions, element names, element types, repeating elements, and simple constraints plus specifications for calculations, complex constraints, question hiding, form layout, randomization, and patient calendar. The WHISpERer tool transforms a declarative data model into an operational data collection mechanism including data dictionaries, forms, databases, and sample data. Each integrated study team (program management, data management, and biostatisticians) collaborates to create and refine a platform independent declarative study specification where WHISpERer generated objects are exercised by the team on enterprise level collection, storage, reporting, and analysis mechanisms. MAVERIC has created extendable standardized forms relating to adverse events, contact information, demographics, drug accountability, medication history, military history, patient termination, protocol deviation, and vital signs. In addition, formalized study specifications exist for dozens of previous studies. The MeDItATE component of WHISpERer supports mining metadata from this study specification repository. WHISpERer encourages the integrated team to take an iterative stepwise approach to study specification. First begin study specification by reusing portions of existing specifications when practical. Second concentrate on specifying the data model that is shared by collection, storage, reporting, and analysis mechanisms. MAVERIC has created extendable standardized forms relating to adverse events, contact information, demographics, drug accountability, medication history, military history, patient termination, protocol deviation, and vital signs. In addition, formalized study specifications exist for dozens of previous studies. The MeDItATE component of WHISpERer supports mining metadata from this study specification repository. WHISpERer encourages the integrated team to take an iterative stepwise approach to study specification. First begin study specification by reusing portions of existing specifications when practical. Second concentrate on specifying the data model that is shared by collection, storage, reporting, and analysis. Thirdly have one part of the team concentrate on specifying complex constraints and calculations while other team members concentrate on specifying form layout and the conditional hiding of questions and controls. This integrated team approach to declarative study specification 1) increases the quality of the data being collected, 2) insures the data model is the one desired by the entire study team prior to deployment to the production platform, 3) allows all team members to exercise the study specification based on technologies each member knows and uses, 4) and permits each study to benefit from lessons learned from previous studies.

P05
ACCUMULATING EVIDENCE AND RESEARCH ORGANIZATION (AERO) MODEL: A NEW TOOL FOR REPRESENTING, ANALYZING, AND PLANNING A TRANSLATIONAL RESEARCH PROGRAM
(1) Spencer Hey, (2) Charles Heilig, (3) Charles Weijer

(1) Biomedical Ethics Unit, McGill University (2) US Centers for Disease Control and Prevention (3) Rotman Institute of Philosophy, University of Western Ontario

Maximizing efficiency in drug development is important for drug developers, policy-makers, and human subjects. Limited funds and the ethical imperative of risk minimization demand that researchers maximize
the knowledge gained per patient enrolled in research. Yet, despite a common perception that the current system of drug development is beset by inefficiencies, there remain few approaches to systematically representing, analyzing, and communicating about the efficiency and coordination of the research enterprise. In this paper, we present the first steps toward developing such an approach: a graph-theoretic tool for representing the Accumulating Evidence and Research Organization (AERO Model) across a translational trajectory. We then apply this tool to a case study involving the antibacterial agent, moxifloxacin, for the treatment of tuberculosis disease.

P06
THE PATH TO A GLOBAL CLINICAL TRIAL FOR ACUTE STROKE IN JAPAN
(1) Haruko Yamamoto (2) Yoko Kai (3) Kanae Hirase (4) Shoichiro Sato (5) Shoji Arihiro (6) Masatoshi Koga (7) Kazunori Toyoda (8) Yuko Y. Palesch (9) Adnan I. Qureshi (10) Kazuo Minematsu
(1)-(3) Dept. of Advanced Medical Technology Development, National Cerebral and Cardiovascular Center (4)-(7) Dept. of Cerebrovascular Medicine, National Cerebral and Cardiovascular Center (8) Dept. of Biostatistics, Medical University of South Carolina (9) Dept. of Neurology, University of Minnesota (10) Vice Director General, National Cerebral and Cardiovascular Center

Background: Stroke is a leading cause of morbidity and mortality and a burden on the healthcare system in Japan. Many domestic clinical trials in stroke have been conducted; however, international collaboration in investigator-initiated trials has been scarce (Asplund K, et al, 2012). In 2010, about 10 Japanese clinical sites decided to participate in the multi-national Antihypertensive Treatment of Cerebral Hemorrhage II (ATACH-II) Trial, which is funded by the National Institutes of Health.

Methods: To overcome the language and regulation differences between the US and Japan, which are considered to be the main barriers against participation of Japanese institutions in global trials, a local coordinating center (JCC) was established at the National Cerebral and Cardiovascular Center in February 2011. The JCC acts as the liaison between the Japanese sites and the Clinical Coordinating Center (CCC) and the Statistical and Data Coordination Center (SDCC) in the US. The JCC’s work has included: 1) identifying and discussing with the CCC/SDCC the regulatory differences between US and Japan; 2) educating the site staff about the importance of adhering to the US regulations; 3) translating the main research documents and manuals; and 4) creating the ‘help-desk’ for the regulatory documents preparation. The JCC also reviews and appropriately approves the submitted regulatory documents in Japanese.

Results: The first site in Japan was released to enroll in February 2012. The first subject was randomized within one month. As of the end of November 2012, 9 sites have been released to enroll, and 3 sites have randomized 16 subjects.

Conclusions: The Japanese sites have recruited over 15% of the total subjects in the ATACH-II to date. It attests to the successful integration of the Japanese clinical sites into an investigator-initiated global trial. Establishing the JCC has proven to be an effective way to enable and facilitate the process.

P07
MULTIPLE TESTING IN CLINICAL TRIALS: SOME NEW APPLICATIONS
Yunzhi Lin, Kefei Zhou, Jitendra Ganju
University of Wisconsin Amgen Gilead Science

The traditional paradigm of determining whether or not the efficacy objective is met in Phase 2 and 3 trials is very limited: there is a single primary endpoint and a single primary method of analysis. The interpretation of results is highly inadequate when either of these ‘singles’ is deficient in capturing a non-null treatment effect. This is a serious problem given the complexity and the soaring cost of trials. This session will focus on a more reliable way to frame questions and test hypotheses so that the methods are more robust or more powerful. Applications will include group sequential trials, determination of cut-points for biomarkers and model selection for dose ranging trials. The goal of the session is to show that the proposed methods are a better way to plan Phase 2 and 3 trials than the current approach.
Subgroup analysis of randomized clinical trial data is used to examine the relationship between a treatment, an independent covariate and a primary outcome of interest. The goal of such analysis is often to determine how a treatment effect is modified by another clinical covariate. Caution is warranted when subgroup analysis is performed and careful interpretation is required, particularly when the subgroup analysis is post-hoc. Usually, when the covariate is continuous, linearity is often assumed in the estimate of a treatment-by-covariate interaction effect. Here we propose a varying-coefficient Cox model to obtain a non-linear estimate of the hazard ratio over the range of a continuous covariate. A penalized spline approach is used for estimation and a likelihood ratio test for inference. Confidence intervals for the hazard ratio over the continuum of the covariate are constructed with a bootstrap procedure. Estimation and inference are validated via a simulation exercise. A local fitting approach is also discussed as an alternative. The proposed method is illustrated using data from randomized clinical trials in heart failure. The analysis of data from heart failure studies shows that there are important instances where the linearity assumption is violated. Results suggest that further caution may be needed when presenting subgroup analyses that involve a continuous covariate. This is especially the case when the subgroup analysis may inform inclusion and exclusion criteria for subsequent clinical trials or restrict the guidelines by which therapy is applied.

Introduction: A variety of surgical techniques are used in the caesarean section operation. Many have not been rigorously evaluated in an RCT, and it is not known which techniques are associated with better outcomes for women and babies. Five specific aspects of the caesarean section procedure were examined to determine which methods lead to optimum outcomes. Women were eligible if they were undergoing their first or second caesarean section through a transverse abdominal incision. Practical challenges The CORONIS trial was a non-regular fractional, factorial RCT conducted in Argentina, Chile, Ghana, India, Kenya, Pakistan and Sudan. This is a rarely used trial design. Five comparisons were evaluated using a was via the web (with an automated back-up telephone system). The intervention pairs were: 1. Blunt versus sharp abdominal entry 2. Exteriorisation versus intra-abdominal repair of the uterus 3. Single versus double layer closure of the uterus 4. Closure versus non-closure of the peritoneum (pelvic and parietal) 5. Chromic catgut versus Polyglactin-910 for uterine repair Primary outcome: death or maternal infectious morbidity. Sample size: 15,000 women; minimum 9,000 women per intervention pair. Organisation & management Regional Trial Offices staffed by small teams managed the local data, dealt with queries, collected SAE data, distributed trial material and monitored protocol adherence. The intelligent data entry, management and monitoring systems used were developed by the International Co-ordinating Centre (ICC) in Oxford. This model allowed “real-time” data management; routine central data monitoring at the ICC was a key element of trial management. 15,935 women were randomised and 98% were followed-up at approximately six weeks post discharge. Central monitoring and regular on-site monitoring meant that recruitment targets were met, overall and by intervention pair. Adherence to the allocation by surgeons was exceptional.

Background: When successfully implemented in randomized controlled trials (RCTs), blinding can prevent ascertainment bias. While many RCTs include procedures to enhance blinding, few report on consequent results.
**Methods**: The SOX Trial is a multicenter RCT of active (A) vs. placebo (P) elastic compression stockings (ECS) to prevent post-thrombotic syndrome (PTS) in 803 patients with deep vein thrombosis (DVT). Attempted blinding of patients, research coordinators (RC) and site investigators (SI) utilized several strategies: enrolling patients with a first episode of DVT (“ECS-naive”), shipping ECS directly to patients’ homes, and instructing patients not wear ECS on study visit days. Success of blinding was assessed at the end of follow-up by guess of treatment group. Responses were analyzed using two indices: James’ Blinding Index (BI), which assesses overall degree of disagreement between treatment allocation and guess, where BI0.2 represents unblinding.

**Results**: James’ BI was 0.71 (95% CI 0.68-0.73) for patients, 0.81 (95% 0.79-0.83) for RCs, and 0.90 (95% CI 0.88-0.91) for SIs. Bang’s BI for the A-ECS and P-ECS interventions respectively were 0.38 (95% CI 0.33-0.43) and -0.15 (95% CI -0.21,-0.09) for patients, 0.22 (95% CI 0.18-0.26) and 0.00 (95% CI -0.04-0.05) for RCs, and 0.12 (95% CI 0.09-0.15) and 0.02 (95% CI -0.01-0.05) for SIs. There was no difference in the incidence of PTS between treatment groups (HRadj 1.17; 95% CI 0.75-1.81; p=0.49).

**Conclusions**: Results suggest that RCs and SIs were blinded to the allocated intervention. For patients, James’ BI suggests that blinding was achieved, while Bang’s BI suggests those in the A-ECS group were unblinded. As overall trial results were negative, this unblinding has minimal impact on trial conclusions. Supported by Canadian Institutes for Health Research and Sigvaris Corp.

---

**P11**

FACTORS ASSOCIATED WITH RETENTION OF AFRICAN-AMERICAN AND WHITE PARTICIPANTS IN A PROSTATE CANCER PREVENTION TRIAL

(1) John A. Hermos, (2) Kathryn B. Arnold, (3) Elise D. Cook, (4) Karen B. Anderson, (5) Jeffrey L. Probstfield

(1) VA Cooperative Studies Program Coordinating Center, VA Boston Healthcare System (2) SWOG Statistical Center (3) MD Anderson Cancer Center, University of Texas (4) Cancer Research and Biostatistics (5) Division of Cardiology, University of Washington School of Medicine

**Background**: Adequate retention of African-American participants in long-term, randomized trials is important for achieving broadly applicable results.

**Purpose**: To determine the incidence of retention failure and individual and study site factors associated with retention failure among White and African-American participants from the Selenium and Vitamin E Cancer Prevention Trial (SELECT), a phase III study of selenium and vitamin E for prevention of prostate cancer.

**Methods**: SELECT randomized 35,533 participants from 427 study sites. Age eligibility was >55 years for Whites, >50 years for African-Americans; race was determined by self-report. Participants from Puerto Rico and non-African-American, non-White participants, as well as participants from sites with severe participant and data management issues, were excluded from analysis. The population analyzed was 32,440: 28,118 (87%) White and 4,322 (13%) African-American. Time to retention failure was defined as days to the earliest of either (1) the second consecutive missed in-person or phone visit or (2) refusal of future contact with study staff. Covariates included participant and study site characteristics. Cox regression was used to estimate hazard ratios.

**Results**: From August 2001 to October 2008, 26.8% of African-American men and 12.9% of White men had retention failure, with rates of 62.5 and 25.7 per 1000 person-years, respectively. This difference was driven by younger African American participants, ages 50-59, comprising 60% of African-Americans with a high retention failure rate (75.0 per 1000 person years). In preliminary analyses, non-adherence to study supplements was strongly associated with retention failure in both racial groups; smoking, living alone, participation from Veterans Affairs-associated sites, and joining SELECT not to “help others in the future” were also important factors in both groups.

**Conclusion**: Results may be useful in identifying personal and study site characteristics to guide recruitment and enhance retention of both African-American and White men in long-term prevention trials.

---

**P12**

OVERVIEW FOCUS AND DRILLDOWN – DATA REPORT INTELLIGENCE

Sherrie Whisler

Pennsylvania Academy of Family Physicians and Foundation

The Pennsylvania Academy of Family Physicians and Foundation developed a data reporting website with grant funds from the PA Department of Health (PADOH). Phase I development was fast-tracked, starting
January 2012 and in production by June 2012. Practices using the site report data on Diabetes, Asthma, IVD, HTN, and Preventative quality measures and were trained and using the site by August 2012. The primary design of the project - called Data Diamond - was to develop a dynamic, highly efficient, and scalable website for data reporting. Using SQL Server 2008, Microsoft .net, and LogiXML, we quickly produce high quality data reports for funders (the PADOH) and practices. The new website houses multiple projects and limits user access based on project permissions. Data analysis through dynamic reporting functions is critical to all projects on the website. LogiXML is the reporting layer, starting at a high project-specific level and then drilling down to specifics as determined by the report requester. Reports show the data in various views, i.e. gauge charts for a quick visual comparing outcome and process measures to an overall project goal. We tweak the concept into a line chart for disease-level analysis. Next level down is practice level and deeper still is the individual measure level analysis. Views for data analysis include gauges, charts, spark lines, and target deltas. All of this data is produced in one interactive report so users can click and drill down where they choose based on their needs. Reports are both flexible and secure, designed based on the user, their roles and permissions. Users only see what is applicable to them and their projects even if they are associated with more than one organization or more than one project. The system manages all roles and provides appropriate views.

**P13**

IMPLEMENTING A CENTRAL MONITORING PROCESS

Gary R Gensler, Zorayr Manukyan, Noga Lewin, Gaurav Sharma, Anne S Lindblad

*All affiliated with The EMMES Corporation except Zorayr Manukyan (Biogen Idec)*

In August 2011, the FDA released a draft guidance encouraging sponsors to develop “risk-based approaches to monitoring clinical investigations of human drug and biological products, medical devices and combinations thereof.” Prior to this date, the FDA Office of Scientific Investigations (OSI) contracted the EMMES Corporation to develop risk-based approaches in an attempt to determine which of four de-identified NDAs were flagged by OSI for having sites with significant good clinical practice (GCP) violations. The methods noted below helped EMMES to successfully identify the 2 NDAs with GCP violations and to further identify nearly all flagged study sites within the two NDAs. These methods included reviewing a) enrollment patterns, b) randomization sequences and treatment balance, c) timing of scheduled study visits relative to anticipated date of occurrence, d) parallel coordinate plots that display within-subject means, variability, Euclidian distance and repetitiveness over time for a vector of common measurements, e) distribution plots, including a plot of each significant digit, f) measures of adverse experience severity and g) Chernoff faces that can combine multiple measurements into pictorial display of site variability. Figures in each domain showed all sites on the same plot for easy comparison. Following the work with the FDA, these methods were applied on a pilot basis to three EMMES projects by a central monitoring team. Each project selected variables of interest and corresponding reports were created. By reviewing the report created by the team, each project identified issues not previously detectable by typical quality-control methods. Efforts are now being made to automate this process so that any project may create on-demand comprehensive reports of side-by-side site-specific clinical data in the domains noted above.

**P14**

A WEIBULL MODEL FOR THE CORRELATION BETWEEN PROGRESSION-FREE SURVIVAL AND OVERALL SURVIVAL

(1) Yimei Li, (2) Qiang Zhang

(1) Division of Oncology, The Children’s Hospital of Philadelphia & 2University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA (2) Radiation Therapy Oncology Group, American College of Radiology, Philadelphia, Pennsylvania, USA

In cancer clinical trials, overall survival (OS), time to progression (TTP) and progression free survival (PFS) are three commonly used endpoints. Empirical correlations between OS and PFS have been investigated and reported for different cancers, but statistical models describing the dependence structures are limited. Recently Fleischer et al. (Statistics in Medicine, 28:2669-2686, 2009) proposed a statistical correlation model based on Exponential distributions. This approach is mathematically tractable and shows some flexibility in describing the dependencies between PFS and OS. However, we usually observe hazard rates that deviate
from the Exponential distribution assumption. In this research, we aim to extend their model to account for the non-constant hazard rates using Weibull distributions. We derived the correlations among different survival outcomes, as well as the distribution of overall survival as induced by the model. Model parameters were estimated using the maximum-likelihood method and the goodness of fit was assessed by comparing the predicted vs. observed overall survival curves. Results are also compared between the two correlation models. Simulations suggested that the Weibull correlation model provides better fit than the Exponential model, for data generated from a range of distributions. We also applied the proposed method to data from two cancer clinical trials. In the non-small-cell lung cancer trial, both the Exponential and Weibull correlation model provide a good fit to the data and the estimated correlations were very similar under each approach. In the prostate cancer trial, the Weibull correlation model showed advantage over the Exponential model with larger estimated correlations. Realistic description of the correlation between PFS and OS from phase II studies is very helpful for planning the phase III trial with OS as the primary endpoint.

P15

AN ASSESSMENT OF AGREEMENT BETWEEN SELF-REPORTED AND CLINIC-REPORTED DATA COLLECTION IN A TYPE 1 DIABETES CLINIC REGISTRY

(1) Kellee M. Miller, (2) Dongyuan Xing, (3) Stephanie N. Dubose, (4) Roy W. Beck for the T1D Exchange Clinic Network

(1, 2, 3, 4) Jaeb Center for Health Research

The T1D Exchange clinic registry consists of 25,833 adults and children with type 1 diabetes (T1D) enrolled at 67 endocrinology clinics through the US. Clinical data are obtained through (1) completion of a questionnaire by the participant/parent of participant (pt) and (2) retrieval of information collected from the medical record (MR). The objective of this analysis is to assess the agreement on the following data elements collected from both the pt and MR: occurrence of diabetic ketoacidosis (DKA) and severe hypoglycemia (SH), frequency of self-monitoring of blood glucose (SMBG), continuous glucose monitor use (CGM), insulin pump use, and total daily insulin (TDI) for insulin pump users. Kappa statistics were reported for categorical variables and Spearman correlation was used for continuous variables. The agreement between self-reported and clinic-reported data ranged from low to moderate for all measures but pump use (Table 1). MRs were much less likely to contain a record of SH and DKA events in the past 12 months. Among 452 pts reporting 3 SH events in the past year, MR review failed to identify any events for 369 pts (82%). Similar trends were observed across all ages. There are substantial discrepancies between clinic- and self-reported data. It is important to consider these discrepancies and possible reasons for poor agreement in data analysis.

P16

EVALUATING POTENTIAL IMPACT OF MODIFIED ELIGIBILITY CRITERIA ON ENROLLMENT AND STUDY RESULTS IN CARDIOVASCULAR CLINICAL TRIALS

(1) Jae Eun Lee, (2) Junghye Sung, (3) Jiyoung Lee

(1) RCMI Translational Research Network Data Coordinating Center @Jackson State University, Mississippi, USA (2) Biostatistics and Epidemiology @ Jackson State University, Jackson, Mississippi, USA (3) University of Pittsburgh, Pittsburgh, PA, USA

Purpose: When recruitment delayed, modifying the eligible criteria may be one of the steps to be considered. However, little is known about guideline to evaluate the potential effect of modified eligibility criteria on enrollment and study goal. This study is to propose an empirical practice to proactively examine how much enrollment can be increased through revising eligibility criteria and how the revision impacts the trial results.

Methods: Key eligibility criteria were obtained from 48 cardiovascular clinical trials registered in ClinicalTrials.gov website as of October 11, 2012: age, low serum vitamin D levels, overweight, well controlled blood pressure, and non-diabetes patients. Potential eligible enrollees were estimated by using the National Health and Nutrition Examination Survey data 2005-6. Its impact on study results was investigated to determine the difference of cardiovascular markers between samples meeting original criteria and those newly-identified by criteria modification by using multiple regression models (SURVEYREG for binary outcome and SURVEYREG for continuous outcome) after controlling age, gender and race.
Results: Our analysis revealed that while decreasing BMI lower limit increased more younger samples, increasing hypertension upper limit increased more female, white, and older samples. Added samples by modifying BMI were better in HDL-Cholesterol, insulin, LDL-cholesterol, Triglyceride, Apolipoprotein, C-reactive protein, parathyroid Hormone, Folate, CVD, and Metabolic syndrome. Meanwhile, added samples by modifying hypertension criterion were more likely to have worse makers in insulin, fasting glucose, triglyceride, Apolipoprotein, Folate, and metabolic syndrome.

Conclusion: Our analyses suggested that modifying the eligible criteria increased potential enrollment, but differentially by select criteria and greater in a specific demographic subgroup, which may cause potential selection bias. This selection bias may misguide the results of the clinical trials. Therefore, change of the eligibility criteria should be carefully done after evaluating its impact on the study goal by using the available data (i.e., open source data).

P17

USE OF PROPENSITY SCORES TO COMBINE DATA FROM TWO SIMILAR CLINICAL TRIALS IN THE URINARY INCONTINENCE TREATMENT NETWORK
Heather Litman, Stephen Kraus, Gary Lemack, Toby Chai, Larry Sirls, Keith Lloyd, Peggy Norton, Elizabeth Mueller, Shawn Menefee, Jonathan Shepherd and John Kusek for the Urinary Incontinence Treatment Network
New England Research Institutes, University of Texas Health Science Center, University of Texas Southwestern, University of Maryland, William Beaumont Hospital, University of Alabama, University of Utah School of Medicine, Loyola Medical Center, University of California San Diego Medical Center/Kaiser Permanente, University of Pittsburgh, National Institute of Diabetes and Digestive and Kidney Diseases

While propensity scores have long been used in observational studies, they are increasingly being utilized to aid analysis of clinical trials. In the Urinary Incontinence Treatment Network (UITN), two randomized clinical trials, the Stress Incontinence Surgical Treatment Efficacy Trial, SISTEr, and the Trial of Midurethral Slings, TOMUS, were conducted on similar yet distinct populations undergoing surgery for stress urinary incontinence. Pre and post-surgery urodynamic (UDS) changes were compared between those randomized to receive an autologous fascia pubovaginal sling in the SISTEr trial to patients who received a midurethral sling in the TOMUS trial. A propensity score analysis was used to help control for bias between the samples by using multiple logistic regression analysis to compute the probability of being enrolled in the SISTEr trial conditional on baseline covariates. The Hosmer-Lemeshow goodness of fit statistic indicated reasonable fit ($p = 0.74$). All propensity-adjusted associations of baseline covariates between studies became statistically non-significant indicating that the propensity score removed bias between the samples for the measures considered. To compare UDS changes between the trials, analysis of covariance was used for the mean difference in continuous parameters and chi-square analysis was used for parameters measured categorically with models including propensity stratum. These models confirmed that after accounting for differing factors between the two study populations using propensity scores, differences between pre and post-surgery UDS measures remained and were similar in both trials. This analysis provides an example of how and why propensity scores can be used in randomized clinical trials.

P18

ASSESSING THE ACTUAL TREATMENT BENEFIT WITH NON-ADHERENCE TO STUDY DRUG IN A LARGE RANDOMIZED TRIAL
(1)Yumi Kubo, (2) Lulu Sterling
(1)Amgen Inc., (2)Amgen Inc.

Intro: The intention-to-treat analysis is the gold standard in assessing the effectiveness of a study drug in randomized clinical trials. However, when non-adherence to study drug is substantial, the actual treatment effect may be underestimated. The EVOLVE trial (Evaluation of Cinacalcet Therapy to Lower Cardiovascular Events), the largest randomized trial ever conducted in the dialysis population, was marked by a large proportion of patients withdrawing from randomized treatment. Various statistical methods were used to assess the impact of non-adherence on the estimated treatment effect.
Methods and Results: The EVOLVE trial (N=3883) was an event-driven trial designed to assess the benefit of cinacalcet compared to placebo on a composite endpoint consisting of all-cause mortality and cardiovascular events. During the 5.5 years of the trial, a large proportion of patients withdrew from treatment but not to follow up in both groups (67% cinacalcet and 70% placebo). A proportion of patients also started taking commercially available cinacalcet (11% cinacalcet and 23% placebo). The median time on treatment (17.5 mo. cinacalcet and 21.2 mo. placebo) was approximately half of the total time patients were in follow-up for endpoints. The hazard ratio (95% CI) for the primary composite endpoint using the ITT analysis was 0.93 (0.85, 1.02). Using lag censoring, where data are censored 6 months after stopping study drug, the HR (95% CI) was 0.85 (0.76, 0.95). Using the inverse probability censoring weight (IPCW) method, the HR (95% CI) was 0.77 (0.66, 0.88). Accelerated failure time models including iterative parameter estimation were also used [HR (95% CI): 0.87 (0.75, 1.01)].

Conclusion: The best method to adjust for non-adherence remains debatable as they each have their limitations. The sensitivity analyses performed, while not providing definitive evidence that cinacalcet is more effective, suggest that the effect size is larger than that estimated by the ITT analysis.

P19
NO COLLABORATION, NO TRIAL: WHY COLLABORATOR OPINION MATTERS.
Claire Cocharan on behalf of the EAGLE Study Group
University of Aberdeen

EAGLE (Effectiveness, in Angle closure Glaucoma, of Lens Extraction) is an international multi-centre, pragmatic, publicly funded randomised controlled trial (RCT). EAGLE addresses whether removal of the lens of the eye for newly diagnosed Primary Angle Closure Glaucoma results in better clinical, economic and patient focussed outcomes compared with standard management. During recruitment 419 participants from 31 centres worldwide were randomised. Clinical and participant-reported outcomes are collected at the recruiting centres for three years following randomisation. Trial data from participant questionnaires and clinical case report forms are uploaded onto the bespoke EAGLE website by collaborators until the end of follow-up (December 2014). Communication and training are key components of managing any trial successfully. Within the UK training of the site collaborators is mainly undertaken via group study days. The aim of this study was to elicit opinions from EAGLE site collaborators to inform future pragmatic multi-centre trials: two focus groups involving eight and nine collaborators respectively were conducted during follow up training; Group training days, rather than individual initiation visits, were preferred as they provide an opportunity to network with peers, take part in practical tutorials and focus on the trial away from the distraction of the everyday environment. Collaborator newsletters remain popular for maintaining trial profile at a group level but at individual level distributing tasks by regular brief emails encourages engagement. Web based data entry gives collaborators a sense of ownership and completion over the data they collect, but thoughtful website design and (telephone) support is key to maintaining data quality. Collaborators like trials that mimic standard clinical care; they are perceived easier to recruit and retain patients within. Barriers included complexity of the inclusion and exclusion criteria, and issues with local financial support. Findings from this study will inform management practice of any future trial requiring collaborator cooperation.

P20
RECOMMENDATIONS FOR FUTURE PATIENT FOCUSED TRIALS ADDRESSING FOOT PAIN IN SYSTEMIC SCLEROSIS
(1*) Joint first authorship, Howard Collier, (1**) Lorraine Loughrey
(1*) Clinical Trials Research Unit, University of Leeds (1**) Division of Rheumatology and Musculoskeletal Disease, University of Leeds

Introduction: A patient initiated Randomised Control Trial was undertaken to evaluate the effectiveness of a simple cushioning and thermal insole in reducing foot pain in patients with systemic sclerosis (SSc). A total of 560 patients were screened across four specialist sites and 141 patients were recruited to target. Three trial contacts were required over 12 weeks, with a total of 11 patient reported assessments completed per participant. Independent patient representatives (with SSc) participated in the trial meetings to inform trial management and design. Findings Compliance and completion of patient reported data was excellent; only 11 participants did not complete follow-up. Both the intervention and a sham device yielded a reduction in...
pain score over 12 weeks but did not meet the pre-specified clinically significant difference. Adjusting for seasonal effect showed only a minimal and insignificant difference in pain between warm and cold months, a factor suggested by the patient representatives. Patient diary comments highlighted severe effects of foot pain on the participant’s psychological status. Much gratitude and support was shown for the trial throughout. Recommendations Patient and public involvement (PPI) is integral for trial success especially in severe systemic diseases and should inform the design from the outset. To yield good recruitment rates, data collection and compliance, numbers of study visits should be minimised with flexibility and support. The nature and source of foot pain in SSc and the interaction with footwear choice requires further evaluation. Future intervention studies should be improved by reporting the nature and source of pain and should reflect the patient need throughout all seasons. An additional non-treatment arm should be incorporated in future trials to explore any potential positive properties of the sham device and any gratitude effect. Consideration of psychological support for participants is recommended when exploring new areas of chronic disease.

**P21**

**BAYESIAN MULTIPLE IMPUTATION FOR MISSING MULTIVARIATE LONGITUDINAL DATA FROM A PARKINSON’S DISEASE CLINICAL TRIAL**

Sheng Luo, Andrew B. Lawson, Bo He, Jordan J. Elm, Barbara C. Tilley

Division of Biostatistics, University of Texas at Houston

In Parkinson’s Disease (PD) clinical trials, PD is studied using multiple outcomes of various types (e.g., binary, ordinal, continuous) collected repeatedly over time. The overall treatment effects across all outcomes can be evaluated based on a global test statistic. However, missing data occur in outcomes for many reasons, e.g., drop-out, death, etc., and need to be imputed in order to conduct an intent-to-treat analysis. We propose a Bayesian method based on item response theory to perform multiple imputation while accounting for multiple sources of correlation. Sensitivity analysis is performed under various scenarios. Our simulation results indicate that the proposed method outperforms standard methods such as last observation carried forward and separate random effects model for each outcome. Our method is motivated by and applied to a PD clinical trial. The proposed method can be broadly applied to longitudinal studies with multiple outcomes subject to missingness.

**P22**

**RANDOMIZATION METRICS: JOINTLY ASSESSING PREDICTABILITY AND EFFICIENCY LOSS IN RANDOMIZATION DESIGNS**

Dennis Sweitzer

Medidata Solutions Worldwide

Randomization methods generally are designed to be both unpredictable and balanced between treatment allocations overall and within strata. However, when planning studies, little consideration is given to measuring these characteristics, nor are they examined jointly, and published comparisons between methods often use incompatible metrics and simulation assumptions. Furthermore, for purposes of real-world planning, such simulations often make unrealistic assumptions (e.g., equal sized strata), and summary statistics give limited information. In order to better reflect real-world study performance, we carried out a series of simulations with 2 treatment arms, and stratification factors that are unequally populated (e.g., 1:2, 1:2:3, or a power law distribution). To measure predictability, we modified the potential selection bias (Efron, Blackwell-Hodges) in which an observer guesses the next treatment to be the one that previously occurred least in the strata containing the subject (i.e., limiting the observer’s knowledge to individual strata, such as site). This reflects a game theory model of randomization pitting observers versus statistician, and is easy to calculate and interpret. To measure imbalances, we calculated efficiency loss using Atkinson’s method because: The main impact of imbalances on the outcomes of a study is a loss of statistical power; Even if treatments are balanced overall, imbalances within small strata can have a disproportionate impact on efficiency; And it is easy to interpret as lost sample size. We applied these methods to evaluate the performance of several popular and novel randomization methods for a variety of parameters, including methods based on permuted blocks, dynamic allocation/minimization, and urn designs. Simulation results were summarized with the median and confidence intervals to estimate best & worst-case scenarios as well as expected performance. Results showed consistent trade-offs between efficiency and unpredictability over methods and parameters, supporting no ‘best’ method to optimize both.
P23
META-ANALYSIS OF ONE OUTCOME FROM GROUP SEQUENTIAL TRIALS WITH COMPOSITE OUTCOMES: ARE STANDARD METHODS APPROPRIATE?
(1) Abigail Shoben
(1) The Ohio State University

Composite outcomes, in which multiple events are combined into one outcome, are the primary outcome in many clinical trials. Composite outcomes are convenient because they provide a way to collapse over competing risks and provide simple interpretation to physicians and others evaluating effectiveness. However, composite outcomes may not provide clarity if treatment effects differ between the outcomes being combined and attempts to separate the effects post-hoc in a single trial or in combined meta-analyses may be problematic. Of additional concern, the composite outcome, or one or more single outcomes may be sequentially monitored for safety or efficacy. Differential monitoring of one or more outcomes further complicates interpretations and post-hoc meta-analyses. This work is motivated by recent meta-analyses of data from randomized trials of magnesium sulfate on neurological deficits and death in preterm infants. This paper illustrates the potential problems resulting from separation of the composite outcome into single outcomes in meta-analyses and provides guidance for future studies.

P24
MULTIPLE PHARMACOKINETIC ASSESSMENTS OF FIVE UNIQUE FORMULATIONS OF EZOGABINE/RETIGABINE COMPARED WITH THE IMMEDIATE RELEASE FORMULATION
Christopher Crean, Mauro Buraglio, Debra Tompson, James Storey, Thangam Arumugham
Valeant Pharmaceuticals

This study was designed to select an ezogabine/retigabine modified release (MR) formulation. The biopharmaceutical comparisons of interest were the bioavailability of the ezogabine MR formulation relative to the immediate release (IR) formulation, and the effect of a high-fat meal versus a standard meal. As previous studies have shown differences in relative bioavailability between single- and repeat-dose administration, this study was conducted at steady state which more closely reflects clinical practice. The study design also addressed the challenges of titration to, and maintenance of, steady state at clinically relevant doses, with minimal dropouts. Study RTG114552 (NCT01332513) was a 35-day study in 36 healthy volunteers (male or female) in which MR formulations were evaluated using steady state twice- (BID) and once-daily (QD) dosing. Subjects received 100 mg ezogabine IR three times daily (TID) on Days 1-3, increasing to 150 mg TID on Days 4-6 (Titration Phase). On Day 7, subjects were randomized into a 6-way crossover (Bioavailability Phase) to receive each of the five ezogabine MR prototypes (300 mg BID) and ezogabine IR (200 mg TID), each dosed for 4 days, with co-administration with a standard meal. On Day 31, subjects were re-randomized to the MR formulations (600 mg QD) for 4 days with co-administration of a high-fat meal on Day 4 (Food Effect Phase). The most common adverse events during the bioavailability phase were CNS-related (dizziness, somnolence, euphoric mood, headache, disturbances in attention). The relative bioavailability based on AUC0-24 for the MR formulations compared with the IR formulation ranged from 0.83 to 0.94. All formulations showed either a small or no impact of a high-fat meal on ezogabine pharmacokinetics. This study design informed MR formulation selection based on key biopharmaceutical criteria.

P25
DEVELOPING APPROACHES FOR RANDOMIZATION WITHIN THE MINI-SENTINEL DISTRIBUTED DATABASE
(1) Roberta Constantine, (2) Pamela Tenaerts, (3) Robert Califf, (4) Richard Platt
(1) Harvard Pilgrim Health Care Institute, (2) Clinical Trials Transformation Initiative, (3) Duke University Medical Center and Duke Translational Medicine Institute, (4) Harvard Medical School and Harvard Pilgrim Health Care Institute

Mini-Sentinel (M-S) is a pilot project sponsored by the U.S. Food and Drug Administration (FDA) to create an active surveillance system to monitor the safety of FDA-regulated medical products using routinely collected electronic health data from over 100 million people. M-S uses pre-existing electronic healthcare data from
multiple sources, employing a distributed data approach in which Data Partners maintain physical and operational control over electronic data in their existing environments. Presently, the M-S only supports observational studies. In concept, the M-S system could identify potential candidates for clinical trials and it could also identify selected clinical trial outcomes. We are assessing the feasibility and develop approaches to using the M-S distributed dataset for these purposes. This activity is a collaboration between M-S and the FDA Clinical Trials Transformation Initiative (CTTI). Investigators are examining the feasibility of using the M-S infrastructure to facilitate patient recruitment including working with the patients’ clinicians to enroll appropriate candidates and follow-up of participants in randomized trials. We are also exploring the ability to ascertain selected outcomes and through a combination of routinely collected electronic health data and medical record review. We are addressing policy and implementation issues concerning patient privacy and confidentiality, and approaches to obtaining informed consent. Investigators will complete a white paper summarizing 1) approaches to randomization, 2) approaches to obtaining informed consent, 3) the implications of use of M-S data from to conduct research (as opposed to public health surveillance), and 4) the appropriate oversight of such research.

P26

ALGORITHMS FOR INFERENCE FOLLOWING BERNOUlli RANDOM WALKS WITH DESIRABLE ABSORBING BOUNDARIES

(1) Neal Oden, (2) Gaurav Sharma

(1 & 2) The EMMES Corporation

Sequential probability ratio tests (SPRTs) and Simon designs are often used to implement phase II clinical trials. At trial’s end, researchers may wish to estimate the parameters underlying the trial design, but this cannot be correctly done without accounting for the design itself. Girshick et al (1946) discussed unbiased estimation of the parameter of a random Bernoulli walk with a simple boundary. Continuing this theme, we review ways to calculate uniformly minimum variance unbiased estimators (UMVUE’s), p-values, expected run lengths, and confidence intervals for Bernoulli walks with finite closed attainable simple boundaries, which we refer to as “desirable”. Single- and multiple-stage Simon trials and truncated binomial sequential probability ratio tests (SPRTs) are desirable walks, as are the k-stage trial of Jennison & Turnbull (1999), curtailed sampling, and a variety other un-named designs. We introduce a simple new method to tell whether a boundary is desirable, and develop a way to name, generate, and count all desirable designs of a given size. We provide several intuitive descriptions of these designs and give explicit computer algorithms for implementing them and determining their characteristics. For example, these algorithms allow truncation of a binomial SPRT at a considerably earlier time than Wald’s (1973) classical method without adversely affecting power and size. Finally, we compare by exact calculation the properties of 5 estimators of the Bernoulli parameter across a range of truncated SPRT’s. The maximum likelihood estimator (MLE), which Chang et al. (1989) found to often have absolute bias less than 0.025 in k-stage Simon designs, does not function as well with truncated SPRTs, where the UMVUE is clearly better. But if root mean squared error is more important than bias, a Bayes estimator, the MLE, and the median unbiased estimator may prove superior to UMVUE and Whitehead estimators.

P27

TRICKS OF THE TRADE: USING SAS® TO ANALYZE IRREGULARLY SCHEDULED CLINICAL DATA

Stacey Slone, John Rinehart

Markey Cancer Center, University of Kentucky Dept. of Internal Medicine, University of Kentucky

Many cancer chemotherapy treatments depend on laboratory values, such as creatinine, to determine dosing levels and schedule. In clinical trials, a patient may have multiple blood draws over a period of days to assess if he/she is capable of receiving chemotherapy and to establish the dosing level. These irregularly scheduled laboratory visits can be tricky to match to a particular chemotherapy cycle since the therapy may not be initiated on the exact day of the blood draw. Although the laboratory and chemotherapy dosing datasets both include the patient identifiers and the respective dates, the dates may not be exact or listed with same cycle code. Hence, merging the datasets becomes a challenge. SAS® PROC SQL can be used to merge datasets in such situations when the matching variables in the datasets of interest do not have exact values but must
be matched within a range. Using a real world example of dosing schedule and laboratory values from a lung cancer trial, I will demonstrate how various SAS functions in conjunction with PROC SQL can be utilized to create a clean analysis dataset to assess dose density. Issues faced include laboratory visits not completed, doses entered as free text and converting doses entered as milligrams back to the AUC dosing required by the protocol.

**P28**

**Evolving best practices to meet the demands of revolutionizing research at coordinating centers**

(1) L. Suzanne Firrell, (2) Haema Nilakanta, (3) Mary A. Foulkes

(1) The Biostatistics Center, The George Washington University (2) The Biostatistics Center, The George Washington University (3) The Biostatistics Center, The George Washington University

The protocol for a research study provides the clinical or scientific framework for recruitment, intervention/observation, assessment, data collection and analysis. A myriad of behind the scenes activities make it possible for a study to be implemented successfully. Most importantly, coordinating centers need to develop a framework to effectively manage these activities, and evaluate and efficiently implement new technology enabling them to deliver the required data and analyses. Technology is rapidly changing the tools for coordinating and managing project infrastructures. Coordinating centers have the experience to evaluate the operational value of these emerging tools, and need to have in place a flexible staff willing and able to develop new skills. These skills and tasks include: establishing effective communications and high-level collaborative relationships with each center; tracking multiple informed consents/renewals; writing scopes of work, requesting bids and selecting vendors and hotels; scheduling and supporting different types of virtual meetings; training a webmaster for a study website; and assuring completion of valid data. To accomplish the above, the recruiting, hiring and training of study staff to have a high level of quality communication skills, project planning and management proficiency is essential. Typically these tasks are not included in a protocol, nor considered a focus in the job description of the staff. In this presentation, we examine these developing demands, how research staff can acquire the skills needed, and look at current best practices and training methods in both assuring that these methods are addressed and ways to ensure continuity and ingenuity for the future.

**P29**

**Common mistakes to avoid in clinical trial case report form design**

Karen, Briggs

Medical University of South Carolina, Statistical and Data Coordination Center

The validity and reliability of the results and interpretations of clinical trials depends heavily on the quality of data collected by Case Report Forms (CRFs). The negative impact of common mistakes in CRF design upon study operation and clinical data quality, have been long recognized. Some problems include: discrepancies between the database workflow and study protocol; data collection schedules lacking coverage for all possible contingency scenarios; CRF items with improper data type specifications; multiple choices without mutually exclusive and exhaustive options; premature coding categorizations with uncontrollable subsequent code additions and edits; data items with arbitrary interpretations; data items difficult to verify with source documentation; data items based upon unrealistic expectations of site personnel; and real-time data validation, rule violation flags introducing the risk of biases. Other mistakes may reduce the efficiency of CRF data collection and procession, such as: CRFs with too many or too few items; CRFs associated with multiple data collection activities across a wide range of time; and an excess of CRF information not directly answering the study questions. These mistakes often go unrecognized until data and operational quality concerns arise in the middle of the trial. This often requires substantial database changes, retrospective data collection, CRF recalls, missing data items, and is likely to adversely affect the data quality and the credibility of trial results. This presentation will cover the lessons learned at the Statistical and Data Coordination Center while working on four multicenter trials in the Neurological Emergencies Treatment Trials (NETT) Network. Strategies and procedures aimed to avoid these mistakes will be shared.
P30

QUALITY CONTROL (QC) OF FIELD-BASED PULMONARY FUNCTION TESTING (PFT) IN THE GULF LONG-TERM FOLLOW-UP STUDY (GULF STUDY)

(1) Steven K. Ramsey,(1) Edward E. Gaunt, (1) Brian D. Blackmon,(1) Matthew D. Curry, (1) Christopher K. Treseder, (1) David A. Johndrow, (2) Richard K. Kwok, (2,3)Larry S. Engel, (2) Dale P. Sandler

(1) SRA International, Inc. (2) National Institute of Environmental Health Sciences; (3) University of North Carolina, Chapel Hill, NC, USA

PFT results are primary outcomes in most studies of pulmonary disease. The literature on PFT QC is largely focused on multi-center, clinic-based studies in which a PFT expert over-reads all tests and provides feedback directly to examiners working under the direction of local investigators. This presentation describes our experience implementing QC activities for field-based PFT measurements in the GuLF STUDY, a large-scale epidemiological study of the potential health effects of exposure to oil and dispersants among clean-up workers who responded to the 2010 Deepwater Horizon oil spill. More than 31,000 participants have enrolled to date, and a subgroup of 8,600 have completed an in-home examination that included PFT. Exams have been conducted by 57 home-based certified medical assistants (CMAs) across five Gulf states. All CMAs received web-based and in-person PFT training and completed practice and certification activities. During training, CMAs are provided with job aides that include a quick reference guide for calibrating and operating the spirometer, standardized scripts and a video to help introduce participants to the maneuver, and a handout demonstrating common PFT problems and solutions. Tests are conducted with standardized spirometers with built-in QC software, which provides real-time quality feedback and recommendations for improving poor quality maneuvers. Field managers periodically attend visits to observe coaching and testing techniques and provide feedback. Finally, CMAs receive weekly reports from field managers about the quality of their PFT scores. Individual and group re-trainings are provided as needed. In large-scale studies with multiple outcomes of interest that are carried out in home-based settings, an approach that relies more heavily on automated software may be an acceptable alternative to expert over-reading, though quality may not be as high as in tightly controlled clinical settings.

P31

A VERSATILE TEST FOR COMPARING SURVIVAL CURVES BASED ON WEIGHTED LOG-RANK STATISTICS

Theodore Karrison
University of Chicago

The log-rank (LR) test is perhaps the most commonly used, nonparametric procedure for comparing two survival curves. The LR test is known to yield maximum power under proportional hazards alternatives. Several authors have described more versatile tests using combinations of weighted log-rank statistics that are more sensitive to non-proportional hazards alternatives. Fleming and Harrington (Wiley, 1991) considered the family of G(?) statistics and their supremum versions, while Lee (Biometrics, 1996) and Lee (Computational Statistics and Data Analysis, 2007) proposed tests based on the more extended G(?, y) family. In this presentation we consider Zm=max(|Z1|, |Z2|, |Z3|), where Z1, Z2, and Z3 are z-statistics obtained from G(0,0), G(1,0), and G(0,1) tests, respectively. G(0,0) corresponds to the LR test while G(1,0) and G(0,1) are sensitive to early and late difference alternatives. Z=(Z1, Z2, Z3) has an asymptotic, trivariate normal distribution with covariance terms that can readily be estimated. The p-value for Zm can therefore be obtained by integrating under the trivariate normal density. An implementation in STATA (College Station, TX) is presented. A simulation study was conducted to compare the performance of the versatile method based on Zm to the LR test and to the “optimally” weighted test under the null hypothesis, proportional hazards, early difference, and late difference alternatives. Results indicate that the method maintains the type I error rate, provides increased power relative to the LR test under early difference and late difference alternatives, and is associated with no more than a 4%-5% power loss relative to the optimally chosen test. The procedure is applied to two real datasets: the Gastrointestinal Tumor Study Group data analyzed in Stablein, Carter, and Novak (Cont Clin Trials, 1981), and the Northern California Oncology Group head-and-neck cancer trial (Efron; JASA, 1988), both of which exhibit a certain degree of non-proportional hazard rates.
P32

EFFECTIVENESS OF A MID-TRIAL TRAINING SESSION FOR MONITORING SERIOUS ADVERSE EVENTS: EXPERIENCE IN ACTION FOR HEALTH IN DIABETES (LOOK AHEAD) STUDY
Sarah A. Gaussoin, Judy Bahnson, Alain Bertoni, Fred Brancati, Lawton Cooper, Jeff Curtis, Siran Ghazarian, Leslie Gregg, Steven Kahn, BJ Maschak-Carey, Anne Murillo, Anne Peters, F. Xavier Pi-Sunyer
Wake Forest Health Sciences

Monitoring participant safety in clinical trials is often done via collecting serious adverse events (SAEs). While reporting guidelines exist for drug intervention studies through the FDA, there is limited literature related to behavioral interventions. Data regarding the training of clinical staff on SAEs reporting are also rare. Look AHEAD, a multi-center, randomized clinical trial of 5,145 overweight or obese participants with Type 2 diabetes evaluating the long-term health outcomes of an intensive lifestyle intervention program compared to diabetes support and education, attempted to bridge these gaps. At the start of the study, central training was provided in all aspects of the study protocol, include SAE reporting. During the trial, each SAE report was centrally reviewed, and it was noted that there appeared to be an increasing lack of consistency over time in interpretations of adverse events among sites. To address this, staff responsible for reporting adverse events were given a written test designed to detect areas regarding identifying and reporting SAEs where additional training may be needed. A one-time training was conducted five years into the study to reinforce common SAE terms and their definitions and review SAE reports and narratives that contained errors. After the training, the test was repeated on average 12 days after the first test. The percentage of significant errors decreased 16% after the training to a level of 0.5% and the scores on the test increased 10%. In longer studies, mid-trial SAE testing and training can be important to maintain high levels of quality control.

P33

THE IMPORTANCE OF INNOVATIVE DATA MANAGEMENT PRACTICES FOR CLINICAL REGISTRIES
(1) Jason Stock, (2) Daniel Jeffers, (3) Jennifer Nobbe, (4) Eileen King
Cincinnati Children’s Hospital Medical Center

The importance of registries for clinical research has been increasing in recent years. Registries allow a look at real world situations that cannot be recreated in randomized clinical trials. Data generated from registries can be used for comparative effectiveness research, safety monitoring and/or quality improvement efforts. Many quality improvement registries lack the data management infrastructure common to clinical trials. Without a monitor on site to assist users with data entry training and data quality issues, the role of the data manager for these projects is greatly increased. With this realization, innovative data management techniques are necessary to successfully ensure the level of data quality for these registries meets the needs of its customers. Each registry presents different challenges to incorporating data quality procedures. Since these registries have a large scope with fewer resources than traditional RCTs, we have found it is not always feasible to institute a data querying process that requires sites to address all potential data entry errors. One registry attempted a full data querying process after 2 years of data collection had already been completed. When queries were issued, sites were overwhelmed and the response rate was much lower than expected. As a result, a key variable monitoring program is being developed instead. Other projects use front end logic checks on data collection forms as a means of ensuring quality. One project has instituted a process monitoring program for data quality similar to that used in the engineering industry to detect changes in data quality processes over time. With the help of innovative data management practices, we have been able to increase the level of data quality in registries without the level of resources typically employed for clinical trials.

P34

DEVELOPMENT OF AN ALGORITHM AND A WEB-BASED SOFTWARE MODULE FOR SINGLE AND MULTI-STAGE PATIENT ENROLLMENT PROCESSES FOR WEB RANDOMIZATION
John E. Podoba
Dacima Software Inc.

Randomizing patients through a web Electronic Data Capture (EDC) platform allows patients to be randomly allocated to treatment groups in real-time using a web browser. The web randomization process can
include an enrollment web form which includes questions (data entry fields) to assess the eligibility or inclusion/exclusion criteria of the patients. Incorporation of validation syntax in the enrollment form allows the web randomization system to verify the eligibility criteria of the patients, classifying them as either eligible or ineligible and then randomize those patients who are eligible for the trial. Some trial designs require patients to be randomized at the time of patient enrollment, while in other trial designs the patient is not randomized at the time of initial enrollment but rather at a later time point, such as after completion of laboratory test or other patient assessments. In order to facilitate different trial enrollment designs, Dacima Software Inc. developed an enrollment/randomization algorithm and module for its web EDC software platform (Dacima Clinical) that allows the trial enrollment and randomization processes to be configured and implemented without the need for computer programming. Through an easy to use graphical user interface the system allows single-stage enrollment or multi-stage enrollment processes to be setup and configured. The multi-stage enrollment algorithm allows for the configuration of hierarchical enrollment form structures which verify eligibility criteria and track patient enrollment status at each stage of the enrollment process. Patients who meet the eligibility criteria at each stage in the sequential enrollment process are randomized after completion of the last enrollment form. The configuration of status levels at each stage in the enrollment process facilitates the management, tracking and monitoring of patients as they precede through the different enrollment stages. The algorithm and software provides a powerful and flexible tool for patient enrollment and randomization.

P35
KEY DATA EDITING FEATURES OF THE BSC DATA MANAGEMENT SYSTEM FOR COORDINATING CENTERS IN MULTI-CENTER CLINICAL TRIALS
Sam Bergman, Zenobia Liendo, Pam Mangat, Susan Reamer, Alla Sapozhnikova
The George Washington University Biostatistics Center

Timely cleaning of data is a critical process to ensure internal validity of a clinical trial. The George Washington University Biostatistics Center (BSC) has implemented a complex data management system that integrates validation during data entry with complex cross-form discrepancy checking and online resolution in real time. The application includes a robust scripting environment where Research Assistants without previous programming experience can independently add validation criteria. Furthermore, the scripting system is multi-purpose, and can be employed in participant eligibility confirmation, data validation, medical monitoring and notification, and adjudication. Simpler data management systems have limited editing functionality and edit evaluation is a time-consuming process. The BSC system automates much of the edit review process and provides online interactive resolution. Among the key features enhancing edit review efficiency are the ability to automatically update edit status to “fixed” when data are corrected by the keyer, and sorting and categorizing of edits to enable previously reviewed edits to be viewed separately from newly generated ones. Our reviewer investigation features include links for direct access to each form referenced in an edit, thus avoiding a multi-step process of accessing participant data. Discrepancy investigation is further improved by direct sharing of edits between central units and clinical centers for resolution. A valuable blog-type communication tool facilitates resolutions between the coordinating center and all related centers, providing context when an edit is redirected. A compatible test environment is a fundamental component that uncovers problems with edit logic and verifies the expected output. This presentation will discuss the complex data editing features of the BSC data management system that offer time savings, consistency, and ensure a high degree of data validity.

P36
RISK OF USING INSTRUMENTS IN INTERNATIONAL CLINICAL TRIALS: THE SCORES MAY NOT BE COMPARABLE ACROSS DIFFERENT COUNTRIES
(1) Weiquan Wei, (1) Chengwu Yang, Yang Xiao, Yuehua Peng, Dingyan Chen, Jincong Yu, Dongming Wang, Fang Ding, Zengzhen Wang 1These two authors contributed equally to this work
Pennsylvania State University, College of Medicine, Hershey, PA, USA (CY) Huazhong University of Technology and Science, Tongji Medical College, Wuhan, China (WW)

Patient-Reported Outcomes (PRO) such as Quality of Life and their measurement tools (i.e., instruments) such as SF-36 have been increasingly used in clinical trials. On the other hand, international clinical trials are becoming more popular in recent years. In order to make the scores from these instruments comparable
across different countries, the instruments’ factor structure must stay the same. However, due to many reasons including culture differences, this key assumption can be violated. As an illustration of this issue, in current study the widely used Obsessive Compulsive Drug Use Scale (OCDUS) was translated from English into Chinese, and then implemented onto 298 heroin addicts in China. Confirmatory factor analyses (CFA) failed to confirm the original 3-factor structure of the OCDUS, which implied that the 13 items cannot be summarized into the 3 sub-scale scores as specified in the original OCDUS to measure the 3 specific aspects of heroin addiction. Exploratory factor analyses (EFA) and another CFA explored and confirmed a new 3-factor structure, which is totally different from the original one. Therefore, if a drug abuse clinical trial needs to recruit participants from both of China and the western countries, and the OCDUS is chosen as the primary outcome measure given its high popularity, then, due to the totally different factor structures in populations from different countries, the sub-scale scores of the OCDUS cannot be used. In international clinical trials that used instruments, if factor structure and other psychometric properties of instruments were not investigated across patient populations in different countries, and a single factor structure was used to summarize the item scores into sub-scale scores and then comparisons across countries were made based on these sub-scale scores, then, the results can be misleading and the conclusions can be wrong.

P37
THE EFFECT OF COMPLIANCE CUT-POINTS ON CACE TREATMENT EFFECT ESTIMATES
James Peugh
Behavioral Medicine & Clinical Psychology, Cincinnati Children’s Hospital Medical Center

Intent-to-treat (ITT) analyses of randomized control trial (RCT) data are commonplace because ITT assumes full compliance with the treatment regimen. Although unlikely, the ITT full treatment compliance assumption results in a conservative “lower-bound” estimate for the effect of treatment. Recent advances in finite longitudinal mixture modeling techniques have resulted in compliance average causal effect (CACE) statistical analyses of RCT data that allow treatment regimen compliance to be not only predicted by covariates, but also incorporated into treatment effect estimates to quantify the additive effects of both treatment condition assignment plus treatment regimen compliance upon the dependent variable of interest. However, CACE estimation techniques currently allow only a binary indicator of compliance to be used in statistical analyses. Forcing treatment compliance to be considered as a binary variable only has at least two undesirable results: patients with very different treatment regimen compliance rates could be assigned the same binary compliance value, and where the compliant/noncompliant decision cut-point is set can have a notable effect on all analysis results, including treatment effect estimates. To illustrate, longitudinal CACE analyses were conducted on data collected from juvenile fibromyalgia (N = 114) patients randomly assigned to a cognitive-behavioral therapeutic treatment condition designed to reduce functional disability, or a fibromyalgia educational control condition. Results showed treatment effect estimates can differ by as much as 350% across several (i.e., 30% - 95%) cut-points used to define a binary indication of treatment regimen compliance. Results also showed that the ability of covariates such as age, depression, and pain severity to significantly predict which patients would be more likely to comply with the treatment regimen can also differ substantially depending upon the cut-point location used in statistical analyses. Additional suggestions for the use of CACE statistical analysis techniques with RCT data, and future CACE analytic possibilities are also discussed.

P38
SUCCEEDING AS A DATA EXPERT AT AN FDA ADVISORY COMMITTEE MEETING
Mark P. McIlduff
Boston Biomedical Associates (BBA)

Upon completion of a pivotal clinical trial, the final step in the FDA approval process often comes down to the recommendation made at an FDA advisory committee panel meeting. Preparing for panel well in advance is critical and will increase the chance of a successful presentation. Data experts may not be in the forefront on panel day, but their work in the lead up to the meeting is a key factor in how well the product is presented. Effective presentation and communication of data are one of the most important components of high-stakes presentations. If managed well, a data expert can support the message of the product and increase the chances of a favorable outcome. The data experts’ key responsibilities include creating succinct slides and assisting the panel team in training for the advisory committee question and answer period. The data
An expert should be familiar with and understand crucial questions at a panel. They should be relaxed in calling up critical data for the presenters on the day of the meeting. A successful data expert is familiar with all areas of the product and presentation. Having a well-prepared data expert on the panel team is critical to a company’s achievement at an FDA advisory committee meeting.

P39
DEVELOPING AN ONLINE TRAINING MODULE FOR MULTICENTRE TRIALS
Johanna Sanchez, Dalah Mason, Elizabeth Asztalos
The Centre for Mother, Infant, and Child Research, Sunnybrook Research Institute

The Centre for Mother, Infant, and Child Research (CMICR) is the central coordinating centre for several large, multi-centre, international, academic randomized controlled trials that aim to improve clinical practice and the health outcomes of women and their children. In order to comply with Good Clinical Practice principles (GCP), it is essential that all participating sites receive protocol training, and that it is documented. To accomplish this, individual trial teams organize pre-initiation investigator meetings for all sites investigators, coordinators, and relevant study staff. In addition, a pre-initiation visit may be conducted to provide additional and individual training. However, it has become evident that since trials are conducted for several years, revisions to the study protocol or procedures must be reviewed with participating sites. In addition, staff changes at the participating sites result in additional people requiring training. To address this, trial teams are in constant communication with the sites by email and teleconference. Due to budgetary limitations, it is not possible to conduct interim investigator meetings, and it is not feasible to conduct site visits each time there are revisions to the study protocol or there is new personnel. To address this, CMICR has been developing an online training module for its research trials. The module will include the following trial-specific sections: study protocol, study procedures, and data entry. In addition, general sections on the module, which will be applicable to all trials, will include: ICH-GCP principles, the randomization process, and electronic data capture. After reading each section, site personnel will be prompted to complete a quiz. Upon successful completion of the quiz, the site and central coordinating trial team will receive an email of completion. The goal is to have an effective tool that will support training and also serve as a reference throughout the conduct of the trial.

P40
IMPROVED RECRUITMENT RATES IN A RANDOMIZED CONTROLLED TRIAL OF ACUTE MYOCARDIAL INFARCTION THROUGH USE OF A WEB-BASED INTERFACE FOR CENTRAL RANDOMIZATION
Brandi Meeks
Population Health Research Institute, Hamilton Health Sciences/McMaster University

Randomized controlled trials in the setting of acute myocardial infarction (AMI) pose many challenges in recruitment. AMI is an incident disease which can occur at any day or time and, most importantly, AMI is a medical emergency and the primary focus is on the patient’s care. It is imperative that any clinical trial-related procedures not cause undue delays since both safety and efficacy depend heavily on the time to treatment. Given the challenges in recruitment, selecting an appropriate randomization process is a fundamental decision in operationalizing a successful trial. Central randomization systems have been traditionally accessed by a telephone connection to an interactive voice response system (IVRS). Using IVRS for randomizations can be cumbersome for the user given the limitations of data entry using a telephone keypad. In a clinical trial of AMI, an investigator may be reluctant to use a tedious system themselves and it is not practical have a study nurse available 24-7. The internet now provides an alternative for accessing central randomization. Web-based interfaces involve the user accessing the central system using a web browser and entering information using a computer keypad onto an online form on a secure website. In an ongoing clinical trial in the setting of AMI patients undergoing primary PCI, feedback from participating investigators is that the web interface for central randomization is more convenient, easy to use and, most importantly in the emergency setting, quick. Observations from the trial suggest that the web-based interface increases off-hour recruitment and improves the rate of accrual. Achieving improved recruitment rates with a simple operational decision will ensure that study completion targets are met and that study results will be available sooner and while still relevant.
P41

NONPARAMETRIC COVARIATE-ADJUSTED HYPOTHESIS TESTS USING R ESTIMATION

(1) John Kloke, (2) Tom Cook
University of Wisconsin – Madison

Standard nonparametric tests are fairly common in clinical practice, for example the Wilcoxon rank-sum. These approaches, however, do not allow for the adjustment of covariates. Aligned rank tests as well as Wald tests and reduction in dispersion tests based on a rank-based fit allow the analyst to adjust for baseline demographics or other patient characteristics while maintaining the robustness and nonparametric properties of classical nonparametric tests. In this presentation we provide a brief overview of rank-based tests that allow for the adjustment of covariates. We then illustrated the use of rank tests on a large clinical dataset. We use this dataset as a basis for a Monte Carlo simulation which demonstrates the increase in power over standard parametric approaches when the assumption of normality is not met.

P42

FAMILY FORUMS: A MODEL FOR PARTICIPANT EDUCATION AND ENGAGEMENT

Charles Heisey, Dejan B. Budimirovic, Tanjala T. Gipson, Rebecca Hinton, Victor Talisa, Michael V. Johnston
Kennedy Krieger Institute

Discussions with families affected by Fragile X syndrome and their clinicians, led us to recognize a communication gap. Families have advocacy and LINKS groups. Clinicians have conventions and conferences. However, there were no local settings for small group discussions and expert presentations on topics of the families’ interest. We decided to target this great need by creating family forums. Using this approach, families have learned from clinicians, but our research team has learned a great deal from the families. Currently, the forums center around Fragile X Syndrome. We not only cover important subject matter, but also create an open atmosphere where individuals feel comfortable enough to discuss their real day to day issues. This enables them to receive guidance from professionals and reassurance from other families that they are not alone. Simultaneous webcasting is provided for those interested, but unable to physically attend. These webcasts are also stored on our website for later viewing. Since the response has been positive, and the audience has grown with each installment, the program will be expanded to include autism and tuberous sclerosis complex. Many clinical trials are hampered by poor recruitment. Our interactions with families led us to hypothesize that decreased opportunities for educational interactions with clinicians/investigators may contribute to poor recruitment. Family forums represent an attempt to target this barrier. This approach may become an important aspect of clinical trial operations that not only increases recruitment, but more importantly gives families an opportunity to be active participants in the process and receive thorough information to help them make an informed choice about clinical trial participation.

P43

BEWARE OF TOO MANY COOKS IN THE KITCHEN: CONCOCTING THE PERFECT QUERY RECIPE

Christina Tittsworth, Alexandra Stout
SGS Life Science Services, KAI Research, Inc.

During the course of a study, there are two individuals who act to ensure that the data being collected are valid and reliable: Clinical Research Associates (CRAs)/Monitors and Data Managers (DMs). The CRA conducts monitoring visits at the site, meets with the Investigative team and ensures that data are appropriately transcribed from source documentation. The DM remotely reviews the data entered in the database for consistency and accuracy by looking at trends in the data. Although they are working towards the same end-point, they approach their roles from different perspectives, and their relationship during a study is not typically an extensive collaboration. This lack of communication can lead to the following issues: Sites issued duplicate queries, Monitor writes inappropriate queries as database specifications are not known, Monitor spends additional time writing and resolving basic queries. These issues are not only burdensome to sites but also wastes resources. If the roles of the CRA and DM are clearly outlined and implemented at study start-up, efforts can be appropriately handled across the study management team. Instead, try the CoOK
Approach: Coordinate query efforts between CRA and DM, Open the lines of communication by collaborating on the resolution of outstanding queries, Keep to the division of responsibility for generating data queries as outlined at study start-up Communication between the CRA and DM will ensure maximum efficiency of query writing and streamlined resolution by the site. Adhering to the agreed-upon division of responsibilities is a key factor in collecting high quality data, which in turn sets a good example to sites. A collaborative CRA/DM relationship will produce the least amount of queries, the highest quality data and likely a more successful study.

P44

A PRACTICAL APPROACH TO ESTIMATING EFFECT OF CHANGE IN RECRUITMENT DURATION, CLUSTER SIZE, AND NUMBER OF CLUSTER UNITS ON STATISTICAL POWER OF A CLUSTER RANDOMIZED TRIAL

Karisse M. Torres and Bambi Arnold-Bush for the Diabetic Retinopathy Clinical Research Network (DRCR.net)

Jaeb Center for Health Research, Diabetic Retinopathy Clinical Research Network (DRCR.net)

Background: The DRCR.net is performing a cluster randomized trial to evaluate the effect of an educational intervention administered in the ophthalmologist’s office on blood glucose control in persons with diabetes. Initial sample size calculations were performed assuming an equal cluster size across 50 clusters. During the recruitment period, a number of clusters were not achieving the targeted cluster size. Therefore, it was necessary to assess the potential impact on statistical power if there was a modification to the number of clusters or cluster size.

Methods: A conventional power computation was performed in lieu of computer simulations to obtain minimum estimates of power in order to assess how long recruitment should be extended for clusters not meeting the recruitment goal, whether additional clusters should be added, or whether the cluster size should be increased where possible. Letting m=number of clusters and n=number of participants per cluster, a power computation formula for a parallel cluster-randomized trial with a continuous outcome assuming equal cluster size was used with varying n and m and only including clusters where sample size was n or projected to be n after an additional 3, 6, or 9 months of recruitment.

Conclusions: The power estimates were used to choose the minimum number of months that recruitment needed to continue while maintaining acceptable statistical power. This is a conservative approach to power calculations since clusters with less than n participants are not included in the power calculations and only n participants are assumed for clusters with more than n participants, both of which will be included in the final analyses. The approach demonstrated may be useful to other cluster randomized trials experiencing similar recruitment difficulties.

P45

USING A GLOBAL OUTCOME (SUM OF PRE-POST RATINGS) IN ASCERTAINING EFFECT OF AN INTERVENTION

Geraldine E. Baggs

Abbott Nutrition Research and Development

Intolerance symptoms were collected at entry and 3 days post treatment. Subjects may have had different symptoms at the two timepoints. A global outcome combining the symptoms was defined to show effect of the treatment (i.e. infant formula) in resolving intolerance. A pre minus post score was obtained for each symptom and the scores added together. Positive (>0) total pre - post score indicates that more symptoms were present at entry than at 3 days post treatment. Let Xi = entry rating = 0 (absent) or 1 (present), i=1,2,...,w= total symptoms. Same for Yi = post treatment rating. Define Di = Xi - Yi. Then D = Sum(Di) + w has a Binomial (n=2w, p= 0.5) distribution, under H0. Suppose there are two groups under study, Control (C) and Experimental (E). Let mC, mE denote the number of subjects in C and E, respectively. Define DE = Sum(Di j) + mE w, the sum over all mE subjects in group E. Same for DC. In this clinical trial, p-values were calculated using the asymptotic distribution of DE – DC. We compared and contrasted results when using the exact and bootstrap distributions, and ascertained for this dataset how well asymptotic theory works, the basis of readily available computing methods.
STANDARDIZED THE WAY TO MANAGEMENT QUERIES WITHOUT RELYING DATA CAPTURING OR CLEANING SYSTEMS
Keiko Ohta, Akiko Kada, Megumi Sakakibara, Yoko Kai, Kanae Takahashi, Haruko Yamamoto
National Cerebral and Cardiovascular Center

Background: Clinical researches need effective data capturing systems (DMS), which should have the function of consistency checks, range checks and logical checks, so that possible errors can be caught and corrected in real time. Because of their high costs and low research budgets, these systems are not always available, especially to young researchers. Some previous researches our data center supported used DMS with limited function or paper case report form without real time data checking, so we have set the query-generating program to run the defined checks automatically against the entire database, and to generate output lists of queries. The queries are then sent to the site. However, before sending, we should manually compare the new lists with previously sent queries each time, because a system cannot ignore queries which had been already responded.

Objective: To standardize a way to manage queries without relying on DMS using a method we developed recently.

Methods: Using Microsoft Access®, we developed a query management system (QMS) which has four steps; 1) store queries with records of their date and type, 2) compare new queries to stored queries, 3) make status (necessary to send/ unnecessary to send), 4) create a reporting form. Our QMS was used in two different researches to test its applicability and usefulness.

Results: In the two researches, our QMS ran without difficulties. The query management process became simple and smooth. [Conclusion] Our standardized way of managing queries alleviated the difficulties and simplified the process of data constancy checks, and thus can be considered a good alternative for DMS.

P47
THE USE OF A WEIGHTED COMBINATION OF THE LIFE-TABLE SURVIVAL ESTIMATES TO EVALUATE THE TREATMENT EFFECT IN A RANDOMIZED CLINICAL TRIAL WHEN THE PROPORTIONAL HAZARDS ASSUMPTION IS VIOLATED
Haijing Qin, Michele Melia
Jaeb Center for Health Research

The proportional hazard model is widely used for the analysis of treatment effect with censored survival data assuming constant hazard ratio. When this assumption is violated, different methods should be used to deal with non-proportionality of hazards such as using time-covariate interactions or stratification when the covariate that interacts with time is not of direct interest. In this presentation, an alternative approach is illustrated by an analysis of data in a Diabetic Retinopathy Clinical Research Network randomized clinical trial to evaluate the effect of intravitreal ranibizumab or triamcinolone acetonide on worsening of diabetic retinopathy for eyes without proliferative diabetic retinopathy at baseline that were followed for up to 3 years. The proportional hazards model could not be used to compute P-values for treatment effects due to significant violation of the proportional hazards assumption with the treatment effect changing with time. As treatment effect was of primary interest, the approach of using covariate-time interaction or stratification could not be applied. Instead, a weighted combination of the life-table survival estimates, stratifying by baseline covariates, was used to obtain an overall estimate of survival and standard error for each treatment group at each annual time point that was adjusted for the covariates. Weights were equal to the proportion of participants in each of the covariate strata with all treatment groups combined. The adjusted estimates were compared between treatment groups using the Z test.

P48
CHALLENGES IN QUANTIFYING PHYSICAL ACTIVITY WITH TRI-AXIAL ACCELEROMETERS IN A RANDOMIZED CONTROLLED CLINICAL TRIAL OF A LIFESTYLE INTERVENTION IN SEDENTARY OLDER WOMEN
(1) Shelly Y. Lensing, (2) Leanne L. Lefler, (3) Jean C. McSweeney, (4) Kimberly K. Garner
(1) Department of Biostatistics, College of Medicine; University of Arkansas for Medical Sciences; (2 & 3) College of Nursing, University of Arkansas for Medical Sciences; (4) Geriatric Research, Education and Clinical Center, Central Arkansas Veterans Healthcare System

Purpose: Little is published on quantifying physical activity in older women using tri-axial capable accelerometers. Processing of this objective data may be needed that recognizes age-related differences in motion and intensity.

http://ctj.sagepub.com Clinical Trials 2013; 10: S1–S88
Methods: Twenty-two women > 60 years of age, able to perform activities of daily living, and who did not engage in regular physical activity were recruited to participate in a randomized study of an intervention to increase lifestyle physical activity. Metabolic Equivalent Tasks (METs), activity minutes, activity in bouts, and intensity levels were measured by the GT3X+ Tri-Axial accelerometer. As different methods for handling data can result in dramatically different values for the same outcome variables, we compared standard accelerometer settings to those that were modified to be sensitive to typical movement in older adults.

Results: Average wear-time was significantly higher for modified 52% (SD 12) vs. 21% (SD 19) than for standard settings. Nine women (41%) did not meet criteria for usable data (>3 days of 10-hours) using standard settings, reducing the analytic sample to 13; whereas, all women were evaluable using our modified settings. Findings of our modified settings were confirmed with individual graph analyses of data. Comparing activity measures, we found significantly lower METS [Mean (Standard Deviation)] 1.09 (0.06) vs. 1.12 (0.10) and higher sedentary minutes per day, 674 (114) vs. 473 (66) for the modified vs. standard settings. There were trends for lower average kilocalories per day, higher number of bouts in moderate activity, and minutes/day in light activity for modified settings vs. standard.

Conclusions: Findings indicate standard accelerometers settings may not be sensitive enough to accurately detect activity of sedentary older women engaging in activity that is well below the standard intensity thresholds. Further research must be conducted to determine accurate measurements in older women.

P49
THE POTENTIAL OF THE POINCARE PLOT IN THE SEARCH FOR MEANINGFUL MEASURES OF HEART RATE VARIABILITY
(1) Steffanie M. Halberstadt, (2) Lei Gao, (3) Anne S. Lindblad, (4) Zorayr Manukyan, (5) Susan Mirow, (6) Lindell K. Weaver
(1) The EMMES Corporation (2) Department of Statistics, George Mason University (3) The EMMES Corporation (4) Biogen Idec Inc. (5) Department of Psychiatry, University of Utah School of Medicine (6) Department of Hyperbaric Medicine, LDS Hospital

In many clinical trials, identifying outcomes with desirable statistical properties that can distinguish between healthy and unhealthy populations and detect treatment effects is challenging. An example is the search for an outcome that accurately describes heart rate variability (HRV), a measure of fluctuation around the mean heart rate used to evaluate cardiovascular activity. HRV is most commonly measured using time domain and frequency domain methods, such as the standard deviation of inter-beat (RR) intervals or the root of squared successive RR intervals. Another approach that has gained popularity is the geometric analysis of the Poincare plot, a dynamic visual display of each RR interval against the subsequent RR interval across time. Most analyses of Poincare plots focus qualitatively or quantitatively on the shape of the plot and how to interpret this shape in terms of short term or long term HRV. Recent attention focuses on the asymmetry of the Poincare plot and its potential for distinguishing between populations. We reviewed current analysis techniques for assessing the shape of the Poincare plot and propose a new approach to capture changes in the shape of the plot across time by quantifying the asymmetry in the plot. The new approach summarizes the asymmetry along the minor axis of the plot for each 5-minute window of the RR intervals to determine the differences over time between accelerations and decelerations of the heart rate for subjects in the same physical state (e.g. sleep or exercise). We will present data to show that this approach is a useful and robust means of describing HRV. We plan to implement this method in evaluating data from a mild traumatic brain injury population and to compare our approach to standard measures of HRV as well as other clinical outcome measures.

P50
DEVELOPMENT OF A PARENT INFORMATIONAL VIDEO TO AID RECRUITMENT IN A PEDIATRIC SURGICAL RANDOMIZED TRIAL
Danielle L. Chandler, Michele Melia, David A. Leske, Jonathan M. Holmes, and Raymond T. Kraker, for the Pediatric Eye Disease Investigator Group
Jaeb Center for Health Research Mayo Clinic

Recruiting children for randomized trials comparing different surgical procedures can be particularly challenging. After experiencing slower than expected recruitment in a randomized trial comparing two types of
eye muscle surgery for treatment of intermittent exotropia (eye misalignment), we surveyed participating investigators and clinic coordinators to determine what specific recruitment barriers they were encountering. We identified the major recruitment barrier to be parental refusal to have their child’s surgical procedure assigned by randomization, a factor which was often related to the investigators’ difficulty in effectively communicating their clinical equipoise when discussing the study with parents. Although participating surgeons had genuine uncertainty about which surgical procedure is more effective, communicating this clinical equipoise to parents can be particularly difficult because in routine practice most surgeons tend to predominantly use one procedure or the other for treating this condition, a choice that is based more on where individual surgeons were trained than on medical evidence. To help overcome these barriers, we developed an informational video for parents who are considering enrolling their child into the randomized trial. The five-minute video describes the purpose of the study, illustrates the difference between the two surgical procedures, explains that each procedure is a very reasonable treatment option with a high reported success rate, and clarifies why randomization is necessary to determine whether one procedure is more effective over the long term. The video was posted on a public website so that it could be viewed by parents in the clinic or at home. IRB approval was required for each clinical site that elected to distribute the video to parents. Since introducing the video in February 2012, the average recruitment per month per site has increased from 0.101 patients to 0.127 patients (26% increase). We believe the video was one factor which contributed to the increased recruitment.

P51
EVALUATION OF RETENTION RATES BY DEMOGRAPHIC CHARACTERISTICS IN THE ENVIRONMENTAL POLYMORPHISM REGISTRY (EPR)
(1) Rebecca Elmore, (2) Kathryn Rose, (3) Matthew Curry, (4) Dale Moffett, (5) Andrea Zombeck, (6) Beverly Warden, (7) Shepherd Schurman, (8) Stavros Garantziotis
(1) SRA International, Inc. (2) SRA International, Inc. (3) SRA International, Inc. (4) SRA International, Inc. (5) SRA International, Inc. (6) SRA International, Inc. (7) National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH) (8) National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH)

Registries are potentially rich resources for clinical and epidemiologic studies. Retention of a registry cohort is critical to study designs that require additional data collection to address key aims. Periodic re-contacting and tracing strategies are often employed to minimize loss to follow-up and associated biases that could impact follow-up data collection, though retention of certain demographic subgroups remains a challenge for many studies. Thus, a comprehensive, multi-modal approach to retention is being implemented in the Environmental Polymorphism Registry (EPR), a large-scale registry established by the National Institute of Environmental Health Sciences (NIEHS) to facilitate genotype-driven translational research of complex disease. To date, the EPR has enrolled and collected blood samples from more than 15,000 participants, including sizeable demographic segments that are often difficult to retain (e.g., 38% male, 27% African American, and 36% less than 36 years old at the time of enrollment). In order to maximize retention, our annual re-contact plan includes efforts to reach participants by email, mail, and phone. Participants who cannot be reached by any of these methods are traced, followed by attempts to contact them by phone, if new information is obtained. This poster compares response rates for various contact methods and tracing outcomes by age, gender, race/ethnicity, and highlights challenges and potential solutions for retaining hard-to-reach registry subgroups for future study.

P52
A COMPARISON OF THE 3+3 DESIGN, THE ROLLING SIX DESIGN, THE CRM, AND THE MODIFIED CRM IN PHASE I CLINICAL TRIALS USING DISCRETE EVENT SIMULATION
(1) Tao Wang, (2) Susan Hilsenbeck, (3) Sarah Baraniuk, (4) Dejian Lai, (5) Xianglin Du, (6) Lem MoyÈ
(1) Baylor College of Medicine (2) School of Public Health, University of Texas

Phase I clinical trials are the “first in human” studies in medical research to examine the toxicities of new agents. Many phase I clinical trial designs proposed in the past 30 years aimed to determine the maximum
tolerable dose (MTD) accurately with least trial duration. The purpose of the present research was to simulate phase I clinical trials and compare 4 commonly used designs. Patients' population was created by the discrete event simulation (DES) method. The patients' characteristics were generated by several distributions with the parameters derived from a historical phase I clinical trial data review. Patients were then selected and enrolled in clinical trials, each of which used the 3+3 design, the rolling six, the continual reassessment method (CRM), and the modified CRM. Five scenarios of dose-toxicity relationship were used to compare the performance of the designs. One thousand trials were simulated for each phase I clinical trial design under each dose-toxicity scenario. The results showed the rolling six design was not superior to the 3+3 design in terms of trial duration. The time to trial completion was comparable between the rolling six and the 3+3 design. However, they both shorten the duration as compared to the two CRM designs. Both CRMs were superior to the 3+3 design and the rolling six in accuracy of MTD estimation. The 3+3 design and rolling six tended to assign more patients to undesired lower dose levels. The toxicities were slightly greater in the CRMs.

P53
MEASURING MENTAL HEALTH, ATTACHMENT, AND BEHAVIOURAL DEVELOPMENT IN 0-1 YEAR OLD INFANTS AND THEIR PARENTS
Maiken Pontoppidan
SFI – The Danish National Center for Social Research

In recent years interest in early intervention for disadvantaged families has increased and preventive interventions have been implemented in Denmark, e.g. The Incredible Years Baby programme (IY Baby). Early intervention programmes aim to improve parent-infant relationship, parental self-worth, and infant development. It is however difficult to measure if improvements occur because both infants and parents go through rapid physical and psychological developments in the first year of a child's life. This presentation will outline the criteria that have been developed to select endpoint assessment instruments for an RCT on the effects of IY Baby in Denmark. The challenges that have occurred in the process will be discussed, including issues such as: Should the primary endpoint be a parent or infant measure? Can attachment between mother and infant be measured without observation? Can a short version of the instrument be used instead of a long version? Will the test work in a Danish context (a country with a strong welfare state, high taxation, free access to school, hospitals etc., and less traditional family values than e.g. the US)? An additional challenge is that the majority of assessment instruments developed for this age group are not translated into Danish and have to be translated and validated as a part of the project.

P54
THE CASE FOR BAYESIAN PREDICTIVE PROBABILITIES FOR INTERIM MONITORING OF CLINICAL TRIALS
(1) Benjamin Saville, (2) Jason Connor, (3) Dan Ayers, (4) JoAnna Alvarez
(1) Vanderbilt University (2) Berry Consultants (3) Vanderbilt University (4) Vanderbilt University

We explore the advantages of a fully Bayesian predictive probability approach versus a traditional Bayesian approach using posterior probabilities for the interim monitoring of clinical trials. Compared to posterior probabilities, predictive probabilities dramatically increase computing time required to design trials with known operating characteristics via simulations. However, some argue that the cost of the computational burden is overwhelmed by the benefits obtained from using predictive probabilities. In this manuscript, we provide evidence that the predictive probability approach is more closely aligned with the clinical decision process than the posterior probability approach. It is more straightforward to derive stopping rules based on predictive probabilities and we show examples of poorly chosen rules based on posterior probabilities. We explore the relationship between predictive probabilities and posterior probabilities as a function of the amount of remaining data to collect in the trial, and argue that the predictive probability approach is a superior strategy for clinical trial design.
ENHANCED PROTOCOL COMPLIANCE IN A COMPLEX CLINICAL TRIAL USING WEB-BASED RESPONSE ADAPTIVE TREATMENT INSTRUCTIONS

(1) Martina Mueller, (2) Wenle Zhao, (3) Charles H. Kellner, (4) Sarah H. Lisanby, (5) Abeba Teklehaimanot, (6) Rebecca G. Knapp

(1), (2), (5), (6) Medical University of South Carolina (3) Mount Sinai School of Medicine (4) Duke University

Maintaining patients with severe depression symptom-free after remission due to an initial course of treatment remains challenging. Even mild residual symptoms contribute to functional impairment and increase the likelihood of relapse. Prolonging Remission In Depressed Elderly (PRIDE) is a multi-center randomized controlled clinical trial funded by NIMH. The study design entails a variable-length acute phase for treatment of severe depression (1 to 4 weeks depending on patient response) and a 24 week long randomized phase to maintain remission. The protocol requires that the schedule for study treatments and depression severity assessments must be determined based on the longitudinal trajectory of the clinical status as reflected in the subject’s current, previous, and baseline clinical depression rating score. To enhance protocol adherence and minimize the burden on study coordinators, a data-driven rule-based treatment instruction algorithm was developed and implemented as part of the web-based clinical trial management system (CTMS). The algorithm rules were developed based on the consensus understanding of the study protocol among clinicians, statisticians and computer programmers. Study coordinators obtain instructions from the study website related to the timing of the subject’s next visit and the number of electroconvulsive treatments that should be administered based on current, baseline and previous total scores on the patient’s depression rating. Use of this system results in three advantageous outcomes: 1) data entry is timely, because it is required to obtain treatment instructions, 2) decrease in study coordinator burden through removal of a substantial amount of data the study coordinator would otherwise be required to remember, therefore allowing the SC to focus on patient-centered activities, such as regular follow-up phone contact, and 3) removal of the possibility of human error in implementation of the complex algorithm, thereby improving the fidelity of the study intervention.

CHARACTERIZATION OF BIAS BETWEEN VALIDATED ASSAYS USING REGRESSION METHODS

(1) Uma Kher, (2) Aditi Sapre, (3) Michael Stepanavage

(1) Merck &Co (2) Merck &Co (3) Merck &Co

The efficacy of a new compound under development may be assessed based on a biomarker that is measured using several validated assays. Difficulties in the interpretation of the assay results may arise when treatment with the experimental compound impacts the biomarker results based on varying assays. Based on the assay utilized, biased biomarker assessments may result which can lead to an inaccurate treatment effect of the experimental compound. In recognition of this problem, the development team developed a plan to fully characterize the bias exhibited by varying assays across varying patient populations, subgroups, and biomarker ranges using statistical techniques commonly used in analysis of assay comparison studies. Inherent in the bias characterization is the classification of the bias components as constant or proportional. If the bias is similar up and down the range of expected values, the bias can be defined as constant. If the difference between assay methods is dependent on the true value of the assay, the bias can be characterized as proportional. We propose to compare the properties of two different regression methods in characterizing assay bias, specifically Deming regression analyses and Magari regression analyses (1998). Finally, we will utilize the methodology by Linnet (1999) to estimate the adequate sample size to detect a significant difference between assay methods in the presence of constant and proportional bias, using real life examples.

HOW DO TRIAL STEERING COMMITTEES AND MANAGEMENT GROUPS CONTRIBUTE TO TRIAL CONDUCT?

(1) J. Athene Lane, (1) Rhiannon Macefield, (2) Michael Clarke, (3) Carol Gamble, (4) Sheila McCann, (5) Matthew Sydes, (1) Alison Haewood

(1) School of Social and Community Medicine, University of Bristol, (2) All-Ireland HTMR, Queen’s University of Belfast (3) University of Liverpool (4) Aberdeen University SMRC CTU

Aims and Introduction: Good clinical practice guidelines (GCP) emphasise trial oversight, usually through a Data Monitoring and Ethics Committee (DMEC) reporting to a Trial Steering Committee (TSC). The Trial
Management Group (TMG) is responsible for the delivery of the trial. Empirical evidence underpins the role and functions of DMECs (DAMOCLES project). We aimed to investigate the roles of TSC and TMGs in decision making regarding trial conduct.

**Methods:** Qualitative and semi-quantitative research to identify how TSCs and TMGs make decisions, who are the key members and how these committees interact. A purposeful sample of clinical trials from UK registered trials units selected across a range of trial designs, interventions and oversight structures. TSC and TMG meetings will be observed by a trained researcher using ethnographic methods. Follow-up interviews will also be held with key committee members (e.g. the Chair, the trial CI) and compared with the ethnographic data as well as to gain individual’s understandings of the committees roles and decision making. A parallel process evaluation will examine trial documents (e.g. protocol), databases, TSC and TMG minutes and reports to funders to establish the impact of these meetings on trial conduct.

**Results:** Review of the existing published evidence base revealed no equivalent detailed guidance to that for DMECs for TSC or TMGs. A cohort of trials for ethnographic observation and analysis have been identified and are being contacted currently. Piloting the research process within one trial was successful and modifications have been made to accommodate teleconference meetings which substantially altered the meeting dynamics.

**Conclusion:** Mixed methods research is becoming more widely utilised for enhancing trial design and delivery and should also be valuable for identifying the role and functions of TSC and TMG in efficient trial conduct. This research should contribute to the evidence base and guidance available to trialists.

---

**P58**

**ON THE ANALYSIS OF TUBERCULOSIS STUDIES WITH MISSING SPUTUM DATA**

(1) Daniel O. Scharfstein, (2) Andrea Rotnitzky, (3) Maria Abraham, (4) Richard Chaisson, (5) Jonghyeon Kim, (6) Lawrence Geiter

(1) Department of Biostatistics, Johns Hopkins School of Public Health

In the design of randomized studies evaluating competing treatments for patients with tuberculosis, it is common to culture sputum for the presence of tuberculosis at regularly scheduled clinic visits over a specified time horizon. A primary goal in such studies is to estimate the treatment-specific distribution of the visit of first culture conversion. Culture conversion is said to have occurred for a patient at a given visit, if all culture sputums for that visit and all subsequent visits are negative. A key complication in the analysis is that culture sputum results are often missing intermittently at some visits and thus culture conversion status is interval censored in some study participants. As a result, the treatment-specific distribution of time of culture conversion is not identified without the imposition of untestable assumptions about the distribution of the culture conversion status within the censoring intervals. When positing assumptions, it is important to condition on as much of the relevant observed data as possible including observed sputum culture results both inside and outside of the censoring interval and auxiliary baseline and time-varying factors that are associated with the unknown sputum cultures inside the censoring interval. In this poster, we develop a methodology that estimates the treatment-specific distribution of time of culture conversion under a class of assumptions on the distribution of the censored culture conversion status that conditions on all of the available data. Each assumption in the class is indexed by a sensitivity analysis parameter which quantifies the magnitude of discrepancy from a specific benchmark assumption. We use a randomized tuberculosis study previously analyzed by Conde et al. (2009) as a case study.

---

**P59**

**CLINICAL MONITORING: A COLLABORATIVE APPROACH. PROCESS AND OUTCOMES OF IMPLEMENTING A CLINICAL MONITORING PROGRAM IN THE NATIONAL INSTITUTE OF ENVIRONMENTAL HEALTH SCIENCES CLINICAL RESEARCH UNIT**

(1) Dawn Hunt, (2) Stavros Garantziotis, (3) Chloe Katz, (4) Lisa Murphy, (5) Silver Wevill

(1), (3), (4), & (5) SRA International, Inc., Durham, North Carolina (2) National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC

In 2011, SRA International, Inc. (SRA) developed a clinical monitoring program for studies implemented through the National Institute of Environmental Health Sciences (NIEHS) intramural Clinical Research Program. The NIEHS Clinical Research Unit (CRU), a new ambulatory care facility, began conducting research
in 2009 and is currently implementing eleven active protocols, several of which were initiated prior to CRU management. Each active CRU study is monitored soon after the initial participant’s enrollment, and at least annually prior to study closure. SRA utilized a collaborative approach for monitoring these studies. The Study Monitor (SM) drafts a study-specific Monitoring Plan and obtains input from the CRU team prior to finalization. This identifies concerns, generates information on study operations, and facilitates relationship development and role identification among CRU staff. The SM subsequently schedules and conducts the monitoring visit and holds a de-briefing with the study team to discuss findings. This allows for recommendations from both the study team and SM, resulting in uniform practices, equal representation in study reports, and minimized ambiguities. Finally, the SM updates the Monitoring Plan using information from the monitoring visit to enhance the monitoring process at future visits. Early monitoring of new studies facilitates timely detection and resolution of issues that arise soon after study initiation and sets expectations from a regulatory and operational perspective. CRU study team members have reported that as a result of monitoring, the study teams have addressed consenting and quality assurance processes and refined research and documentation practices. Close collaboration between the SM and study team facilitated increased monitor understanding of the study, additional study team preparation, mutual consensus-building, expeditied on-site monitoring, positive working relationships, and open communication.

P60

CHALLENGES IN ANALYZING SAFETY DATA IN A PHASE III TRIAL TO EVALUATE THE SAFETY OF EXTENDED REGIMEN OF NEVIRAPINE (NVP) IN INFANTS BORN TO HIV-INFECTED MOTHERS TO PREVENT VERTICAL HIV TRANSMISSION DURING BREASTFEEDING (HPTN 046)

(1) Jenny Tseng, (1) Casey Herron, (1) Elizabeth Brown, (2) Philip Andrew, (1) Lynda Emel
(1) Statistical Center for HIV/AIDS Research and Prevention (SCHARP), Fred Hutchinson Cancer Research Center, Seattle, WA (2) Science Facilitation, FHI 360, Durham, NC

Background: As a result of the systematic assessment of safety in large phase III clinical trials, many unexpected adverse events (AEs) are encountered that have no known association with the intervention. When comparing these AEs, p-values are often reported unadjusted, which can result in a high false discovery rate (FDR). In the analysis of HPTN 046 phase III trial, using the Mehrotra-Heyse-Tukey “Double FDR” adjustment provides an objective method to address issues of multiplicity[1].

Methods: Randomized infants who initiated study drug were included in the safety assessment. AEs, reported from 6 weeks through 8 months of life, were coded using the Medical Dictionary for Regulatory Activities (MedDRA) and grouped by Preferred Term (PT) and System Organ Class (SOC). Number and proportion of infants with specific PTs were tabulated by study arm (nevirapine vs. placebo) and compared using Fisher’s exact test, adjusting the resulting p-values using the Mehrotra-Heyse-Tukey “Double FDR” approach.

Results: 1519 randomized infants initiated study drug (758 in nevirapine and 761 in placebo). In the first 8 months, 1250 (82%) infants reported a total of 3,696 AEs (1,876 in nevirapine and 1,820 in placebo). The non-rare AEs (those with overall incidence greater than 5) were grouped into 50 MedDRA PTs and 10 SOCs. Six PTs had unadjusted p-values less than 0.05; conjunctivitis, pneumonia, respiratory tract infection, Tinea faciei, upper respiratory tract infection and atopic dermatitis. After applying the “double FDR” adjustment only one PT achieved statistical significance (upper respiratory tract infection).

Conclusion: Safety data analysis using the “Double FDR” method to detect differences between study arms is a useful analytical tool when AE data are grouped into numerous categories. [1] Mehrotra DV and JF Heyse. Use of the false discovery rate for evaluating clinical safety data. Statistical Methods in Medical Research 2004; 13:227-238

P61

WHEN SURVIVAL CURVES CROSS...ADD A LANDMARK ANALYSIS: AN EXAMPLE FROM CARDIOLOGY

(1) Anne Glass, (2) Christoph Nienaber, (3) Guenther Kundt
(1) (3) Institute for Biostatistics and Informatics in Medicine and Ageing Research, Rostock University Hospital, Germany (2) Medical Center I, Cardiology Department, Rostock University Hospital, Germany

Objectives: When comparing survival probabilities of different treatment studies it may happen that Kaplan-Meier curves cross, and thereby indicate non-proportional hazards (HR). This violates assumptions of survival analysis and reduces the power of Log-Rank test. One alternative is dividing survival time into two parts by

http://ctj.sagepub.com

Clinical Trials 2013; 10: S1–S88
setting a landmark, and analyzing both parts separately. Results will give more background, but must be discussed cautiously, if concerning the second phase, due to limitations of Landmark analysis [1].

**Methods:** 140 Patients were randomly assigned to elective standard (medical, n=68), or interventional therapy (standard and repair by stent-graft, n=72). The impact of intervention was assessed after 5ys as patients were followed for two mortality parameters and evidence of progression of disease. All analyses were based on intention to treat design: Cox regression (HR), survival curves (Log-Rank-test). As survival curves cross, Landmark analysis at 2ys after randomization was carried out, with HRs calculated separately for events 0-24 months and, 24-60 months. Additional landmarks at 1 month and 1y after randomization were investigated to diffuse the risk of dependence on the 2ys-landmark.

**Results:** A benefit of endovascular repair at 5ys follow-up could be shown in the trial descriptively (decreased risks of mortality and progression), and graphically by Kaplan-Meier curves. HRs were significant for two parameters. More insight gave Landmark analysis: all parameters showed significant p-values after landmark, but not before. Additional set landmarks showed consistent results. Interactions between treatment and time were significant. The long-term benefit of repair versus medication starts beyond 2ys, if surgery-related risks are over.

**Conclusions:** Landmark analyses appear to be of special interest and need in cardiovascular research. Though this method is, in case of crossing survival curves, still better than a “naïve” approach, results can’t be generalized and have to be interpreted with regard to its conditional nature. [1] Dafni (2011) Circulation

---

**P62**

THE SENSITIVITY AND POSITIVE PREDICTIVE VALUE OF STROKE PATTERN INDICATORS IN TRADITIONAL KOREAN MEDICINE: THE KOREAN STANDARD PATTERN IDENTIFICATION QUESTIONNAIRE FOR STROKE

Tae-Yong Park, Tae-Woong Moon, Byoung-Kab Kang, Ju Ah Lee, Myeong Soo Lee, Mi Mi Ko

**Background:** Individual clinicians differ in the clinical importance they place on each clinical indicator used for pattern identification (PI). We aimed to identify the key indicators for stroke in traditional Korean medicine (TKM).

**Methods:** The Korean Standard Pattern Identification for Stroke III (K-SPI-Stroke-III), a stroke questionnaire for PI that includes 44 clinical indicators associated with Dampness-Phlegm, Fire-Heat, Qi-deficiency, or Yin-deficiency, has been developed. TKM doctors from 11 hospitals evaluated 1,286 subjects, indicating the patients’ symptoms and signs on the K-SPI-Stroke-III. They were asked to leave a special mark beside the indicators that they considered important for PI, regardless of the number of indicators. Using a frequency analysis, the pattern sensitivity (PS) and the pattern positive predictive value (PPPV) for each pattern indicator were obtained.

**Results:** Four indicators (reddened complexion, slippery pulse, looks weak and lethargic, and feels weak and lethargic) were chosen as the key indicators exceeding 50% for PS (differentiating between patterns) and PPPV (frequency of appearing within a pattern). In contrast, seven other indicators were determined to be the least useful indicators.

**Conclusion:** A new, more time-efficient version of the K-SPI-Stroke-III that maintains the current diagnostic accuracy should be published, reflecting the insignificant indicators and the key indicators.

---

**P63**

LITERATURE REVIEW TO COMPARE INTERIM AND FINAL RESULTS OF PUBLISHED CLINICAL TRIALS

(1) Nicholas Counsell, (2) Despina Biri, (3) Allan Hackhaw

(1) Cancer Research Uk & Ucl Cancer Trial Centre, University College London (2) Cancer Research Uk & UCL Cancer Trial Centre, University College London (3) Cancer Research UK & UCL Cancer Trial Centre, University College London

**Background:** Interim analyses of randomised controlled trials (RCTs) are sometimes published before the final analysis. Because we occasionally noticed a marked difference in the reported treatment effects between the interim and final results, we aimed to determine how common this was, and the extent of the difference.

**Methods:** We conducted a systematic review to compare the interim and final results of published RCTs. We used Medline (1990-2012) and some manual searches of several journals to identify pairs of articles reporting
the interim and final results of the same RCT. 2712 articles were identified from the searches, in which there were 51 pairs of papers with relevant data.

**Findings:** Our preliminary findings here are based on an initial 15 pairs, all of which refer to the first published article as “interim analyses”. Examination of these trials suggest that the final results were generally of a smaller magnitude than the interim ones, with smaller effect sizes in 11 (73%) trials and a marked reduction (the effect size reduced by 20% from interim to final report) in 6 (40%) of these. For example, in a trial examining trastuzumab after adjuvant chemotherapy in patients with HER2-positive early breast cancer, the reported interim hazard ratio was 0.54, but the subsequent reported hazard ratio was 0.76.

**Conclusions:** Our preliminary analysis suggests that published interim results may need to be considered with caution, because the final treatment effect could be noticeably smaller. Although our findings above are based on 15 of the 51 pairs, full analyses will be available for a presentation/poster. However, we can already say that at least 12% of such pairs (6/51) show a marked difference in effect sizes. It is essential to have reliable estimates of treatment effects, which are also important for regulatory authority reviews (marketing license) and health economic analyses.

---

**P64**

**SIGNAL DETECTION IN SMALL SAFETY DATABASES**

(1) Katharine Poole, (2) Victoria Williams

(1) Rho, Inc (2) Rho, Inc.

On September 29, 2010, the FDA published a new final rule amending the IND safety reporting requirements under 21 CFR Part 312 to improve safety monitoring in clinical trials. Under this new regulation, effective March 28, 2011, all trials involving an investigational new drug (IND) application are held to more stringent analysis and reporting guidelines for adverse events. One specific requirement of the regulation is that sponsors should have a systematic approach in place for safety surveillance of their entire safety database. This signal detection process extends throughout the investigational lifetime of a drug and is designed to determine if the incidence any adverse events associated with a study drug is higher than their incidence associated with other drugs or placebos. It is intended to aid in detection of safety signals present at low frequencies that may escape detection by looking solely at individual trials. Several methods for analyzing large safety datasets aggregated across multiple studies have been published, but publications regarding signal detection in smaller safety databases are scarce. The NIAID, NIDDK, and JDRH -fund the Immune Tolerance Network (ITN) which includes a portfolio of several smaller clinical trials, many of which are under INDs held by DAIT, NIAID or by the investigators. To comply with the new regulations, attempts were made to extend the methods for analyzing large databases for use on much smaller scale analyses of single studies or groups of studies. This poster will focus on the feasibility of extrapolating a variety of methods used on large databases to smaller studies. We will also display graphical tools developed to enhance the evaluation of possible adverse event signals and discuss additional ways to group safety data for further analysis. Funded by NIAID #HHSN272200800029C.

---

**P65**

**CHALLENGES OF PROVIDING MEDICAL RESULTS TO PARTICIPANTS IN FIELD-BASED STUDIES: EXPERIENCES IN THE GULF LONG-TERM FOLLOW-UP STUDY (GULF STUDY)**

(1) Jessica M. Stucker, (1) Steven K. Ramsey, (1) Edward E. Gaunt, (1) Kathryn M. Rose, (1) Matthew D. Curry, (2) Aubrey Miller, (2) Richard K. Kwok, (2) Dale P. Sandler

(1) SRA International, Inc; (2) National Institute of Environmental Health Sciences

Clinical and epidemiological studies often collect health data that are relevant to the medical care of study participants. Clinically relevant findings are typically shared with participants, many of whom view receipt of findings as a benefit of participation. This presentation focuses on our approach to sharing findings with participants in the GuLF STUDY, a prospective cohort study examining potential health effects of exposure to oil and dispersants among clean-up workers who responded to the 2010 Deepwater Horizon oil spill. Specifically, we will describe our approaches to collecting and sharing clinical measurements in the home setting, our resources and procedures for providing referrals to participants without health insurance or a primary care provider, and our procedures for responding to urgent physical and mental health problems.
We will also discuss some of the challenges we have encountered in this large scale, field-based study. Some challenges are due to characteristics of the cohort, many of whom are economically disadvantaged, unhealthy, and medically underserved. Other challenges relate to the research setting where paraprofessionals interact with participants in home environments that are sometimes difficult to control. Our experiences sharing results in the home setting may be of interest to other field-based epidemiological studies.

P66

STRATEGIES IMPLEMENTED TO ENSURE KNOWLEDGE TRANSFER
Kathryn Mangoff, Dalah Mason, Johanna Sanchez, Elizabeth Asztalos, and Jon Barrett
The Centre for Mother, Infant, and Child Research, Sunnybrook Research Institute

The Twin Birth Study (TBS) is an international multicentre randomized controlled trial that recruited 2804 patients from 106 centres in 25 countries. TBS seeks to determine, in women expecting twins, whether a policy of planned Caesarean section decreases the likelihood of perinatal or neonatal mortality or serious neonatal morbidity, compared to a policy of planned vaginal birth. Recruitment for the study was completed in April 2011, and the data analysis was completed in early 2012. It was therefore important to develop a dissemination plan to share the results with the collaborating sites and the international perinatal community. In order to inform the TBS collaborators of the results, an international meeting was held in Toronto in May 2012. The Principal Investigator (PI) of the TBS study presented the results, which were followed by an extensive discussion with the collaborators, where everyone had the opportunity to ask questions. This meeting allowed everyone in attendance to gain a collective understanding of the results. Once the results are published or presented at a scientific meeting by the PI, a plan will then be devised on how best to disseminate the findings globally and into everyday practice. One strategy is to make a comprehensive PowerPoint slide presentation on the findings and distribute it to all of the investigators involved in TBS so they can share the information with their colleagues in their respective countries. Another strategy is to create a sub-committee made of collaborators representing different countries and regions to bring in an international perspective about the findings. Yet another strategy is to have the PI or members of the Steering Committee travel to different conferences to disseminate the findings reaching as many groups as possible. This presentation will share the strategies used to ensure successful dissemination of the TBS results worldwide.

P67

DELEGATION LOGS IN A REGULATED CLINICAL TRIAL INVOLVING MULTIDISCIPLINARY TEAMS
Jessica Stortz, Melissa Brown, Dalah Mason, Johanna Sanchez, Elizabeth Asztalos
The Centre for Mother, Infant, and Child Research, Sunnybrook Research Institute

The Enhancing Breast Milk Production with Domperidone in Mothers of Preterm Neonates study (EMPOWER) is an international, multi-centre, randomized controlled trial investigating safety of domperidone for mothers of preterm infants as well as timing of initiation and duration of treatment. EMPOWER is a Health Canada regulated trial with approximately 20 participating sites. Study teams are comprised of individuals with various professional backgrounds including neonatologists, nurses, lactation consultants, pharmacists, research coordinators and assistants. The implementation and use of the Site Signature and Task Delegation Log is an ICH-GCP (International Conference on Harmonisation Good Clinical Practice) requirement and is in compliance with Health Canada regulations. This log assists in organizing delegated responsibilities at each site. The site investigator is required to authorize tasks delegated to individuals on the log, via signature, indicating each individual listed has been trained and is qualified to conduct the assigned task(s). EMPOWER sites initially faced challenges in correctly completing the Site Signature and Task Delegation Log. Some tasks on the original version of the log should only be performed by registered pharmacists or individuals with a medical license, but had been assigned to individuals without these qualifications. There were also inconsistencies with respect to the format of dates recorded in the log. Additionally, the date section was incorrectly interpreted as the start date as opposed to the date of training. Due to challenges faced at multiple sites, particularly with task assignment, the Site Signature and Task Delegation Log was revised to increase comprehensiveness of instructions for completion. This ensured compliance with both ICH-GCP and Health Canada regulations and resulted in significantly fewer errors in completion of the log. This presentation will
discuss strategies utilized in the revision process and share the revised version of the EMPOWER Site Signature and Task Delegation Log.

P68

LOCATING STUDY PARTICIPANTS 5 YEARS AFTER CLOSEOUT
(1) Mae Gordon, (2) Patricia Morris, (3) Deborah Dunn, (4) Leonard Haertter, (5) The Ocular Hypertension Treatment Study
(1) Washington University School of Medicine (2) Washington University School of Medicine (3) Washington University School of Medicine (4) Washington University School of Medicine (5) Washington University School of Medicine

Coordinating centers and clinics typically do not have adequate tracing resources to locate participants 3-5 years after study close-out. We describe how The Ocular Hypertension Treatment Study (OHTS) triaged the task of tracing study participants from 30 clinics to a locator service. Our goal was to re-examine 80% of the surviving participants (median age of 67 years at study close-out, range 58-94). Adjusting for deaths on the National Death Index, 1,025 study participants were classified as survivors. Clinics had PHI (birth date and place, SS, participant’s father’s name, last known address/telephone number, spouse’s name, and closest friend/relative not living with participant) collected at baseline and updated annually during the 15 year study. Clinics were able to have voice contact with 65% (669 of 1,025 survivors). The Coordinating Center referred these 356 non-contactable participants to a locator service. The locator service identified 41 additional deaths and retrieved updated contact information for the 315 survivors. Using updated contact information, Clinics were able to contact 82% (258 of the 315 surviving participants) of the “non-contactable” participants to date. Information from the locator service enabled us to increase the contact rate of surviving study participants from 65% (669 of 1,025 survivors) to 94% (927 of the 984 survivors). We will document how the locator service was useful in locating participants “lost to follow-up” and our cost benefit analysis.

P69

COMPARABILITY OF CTCAE GRADING AND CLINICAL SIGNIFICANCE IN ABNORMAL CLINICAL LABORATORY RESULTS
Ashley Pinckney
Rho, Inc.

In clinical trials, laboratory data is a key element of the safety profile for the treatment being studied. When local labs are used, there may be a large amount of variance in determining clinically significant abnormal lab values in the reporting of adverse events (AEs). Many clinical trials are moving towards the use of grading scales such as the National Cancer Institute common terminology criteria for adverse events (NCI-CTCAE) to provide standard ranges and implement suggested grading across clinical sites. Rho, Inc. has developed a macro to incorporate the CTCAE grading criteria into laboratory analysis datasets. The CTCAE grading is complex and encompasses a wide variety of laboratory tests making this macro extremely useful. The purpose of this poster is to explore the differences in AEs generated by CTCAE and those generated by clinical significance alone across local sites. Discussion will be presented on various reporting techniques and the distribution and comparability of abnormal clinical laboratory results across sites. This paper will also present the above macro that requires minimal user input to efficiently grade abnormal lab values based on the CTCAE.

P70

USE OF JOINT MODELING FOR LONGITUDINAL AND SURVIVAL DATA IN LONG-TERM RENAL CLINICAL TRIALS
Antonietta Chianca, Giovanni A Giuliano, Annalisa Perna
‘Aldo e Cele Daccò’ Clinical Research Center, Mario Negri Institute for Pharmacological Research

Many long-term clinical trials generate both longitudinal (repeated measurement) and survival (time to event) data, which are usually analyzed separately, using linear mixed effects models for repeated measures
data, and semiparametric (Cox) proportional hazards models or parametric models for survival data. Their separate use however may be suboptimal when longitudinal variables are correlated with patient health status and survival endpoints. An unified, flexible approach by means of joint models has been proposed (Henderson 2000) in order to obtain less biased and more efficient inferences (Guo & Carlin, 2004). The ‘Aldo e Cele Daccò’ Clinical Research Center coordinates several multicenter, multi-national long-term randomized clinical trials in progressive non-diabetic chronic kidney diseases. In these studies the rate of change in measured glomerular filtration rate (mGFR) or time to end stage renal disease (ESRD) or both are selected as primary efficacy variables. We explored the performance of joint modeling through a real-data application using mGFR longitudinal data and time to ESRD measured in participants enrolled in such trials. The inclusion of time-varying covariates in survival analyses and the longitudinal response data affected by informative dropout are evaluated.

**P71**

**THE EFFECT OF THE TRANSITION FROM PAPER-BASED TO ELECTRONIC DATA CAPTURE ON CLINICAL TRIAL METADATA**

(1) Sunny Chan, (2) Michael Shi, (3) Cathy Yang, (4) David Lau, (5) Elizabeth Asztalos  
(1,2,3,4,5) The Centre for Mother, Infant, and Child Research, Sunnybrook Research Institute

Clinical trial metadata define the characteristics of data and documents related to trials. Common metadata are attributes of datasets from data forms and supporting documents. The Centre for Mother, Infant, and Child Research (CMICR), is the data and clinical coordinating centre for several multi-centre randomized controlled trials. In earlier trials, paper-based case report forms were found to be a reliable but costly method of data collection. Since 2011, CMICR coordinated trials use electronic data capture (EDC) as the primary method to collect data, with the exception of patient administrative questionnaires that are collected on paper-based case report forms. Clinical trial data are commonly stored electronically using a clinical database management system. With the adoption of an electronic data capture system, there exists a need to organize, standardize and manage clinical research data and metadata. A series of SAS programs will be developed to facilitate EDC data integration with existing systems, and ensure accurate and consistent data for analysis. This process will automate repeatable clinical data integration tasks and also allow central trial coordinators at CMICR to perform further data quality checks implemented in SAS logic programs. Efforts are underway to standardize the way the clinical data are organized and to ensure metadata transparency between paper-based and EDC data collection methods.

**P72**

**WEB-BASED ADMINISTRATION AND TRACKING OF COGNITIVE DATA IN THE ACCORDION-MIND STUDY**

(1) Patricia Davis, (2) John Hepler  
(1-2) Wake Forest University School of Medicine

In recent years more and more clinical trials have used web-based technology to aid in the management of portions of their studies. Whether used for aspects of recruitment, data collection, reporting or study analysis, the web is a useful means of communication between sites and a coordinating center within the clinical trial setting. In the ACCORDION Study, we use the web not only for data collection and analysis but also for many administrative tasks such as scheduling, sub-study management, outcomes tracking and event adjudication. By building administrative functionality into the website, data become more accessible by study staff, thus enabling the streamlining of tasks that were previously accomplished via much slower methods (phone, fax, email). In this poster, we will describe how the ACCORDION-Mind Study uses a web-based cognitive booklet tracking tool to monitor and track the mailing, completion, return and data entry of cognitive booklets for the purpose of determining status for payment. Prior to implementation of this system, booklets were managed manually with spreadsheets on multiple computers, resulting in difficulty accounting for all booklets. The implementation of a web-based process to administer and monitor this activity has greatly improved the process by which booklets are monitored and tracked, while providing a more cost- and time-efficient method for submitting requests for reimbursement to the coordinating center.
P73

POWER ANALYSIS AND SAMPLE SIZE CALCULATION IN DESIGNING A SMART TRIAL FOR BUILDING DYNAMIC TREATMENT REGIMENS

(1) Huitian Lei, (2) Kevin Lynch, (3) David Oslin, (4) Susan Murphy

(1) Department of Statistics, University of Michigan (2) Treatment Research Center and Center for Studies of Addictions, Department of Psychiatry, University of Pennsylvania (3) Philadelphia Veterans Administration Medical Center, Philadelphia, Pennsylvania 19104, and Treatment Research Center and Center for Studies of Addictions, Department of Psychiatry, University of Pennsylvania (4) Department of Statistics, Institute for Social Research, and Department of Psychiatry, University of Michigan

Dynamic treatment regimen is a sequence of decision rules that specifies which treatment to provide first and how to adjust and adapt treatment according to patients’ ongoing performances and changing needs. The sequential multiple assignment randomized trial (SMART) is an experimental design used to inform the development of high-quality dynamic treatment regimens. In a two-stage SMART design, the first stage usually involves randomizing patients to two initial treatments/options. Patients’ responses to their initial treatments are measured, based on which they are classified as early responders or non-responders. The second stage of a SMART design usually involves randomizing (responders) non-responders to two (maintenance) rescue treatments/options. When designing a SMART study various research questions may be considered, some of which concern the comparison of two treatment components while others concern the comparison of embedded dynamic treatment regimens. This review focuses on the power analysis and sample size calculation in designing SMART trials. When the primary hypothesis of a SMART study is to compare two treatment components, standard sample size formulae for the two-arm randomized clinical trials are extended where a good guess of the early non-response rate may be needed. When the primary hypothesis concerns the comparison of dynamic treatment regimens as a whole, conservative sample size formulae/algorithm are developed. We conduct simulation studies to demonstrate the power performance of the formulae/algorithm and their robustness under the violation of (some) working assumptions. Throughout the review, we use the Extending Treatment Effectiveness of Naltrexone (ExTENd) SMART trial as an illustrative example.

P74

A MULTI-SYSTEM SETUP FOR DEVELOPING ELECTRONIC CASE REPORT FORMS (E-CRFS)

Cathy Yang, Michael Shi, Sunny Chan, Elizabeth Asztalos

The Centre for Mother, Infant, and Child Research Sunnybrook Research Institute

The Centre for Mother, Infant, and Child Research (CMICR) is the data and clinical coordinating centre for several large, national, and international multi-centre randomized controlled trials. Since 2011, data collection for newly funded trials are implemented with a third party Electronic Data Capture (EDC) system, which provides cleaner data, a more efficient data entry and management process, and faster access to the data. To ensure that all patient data defined by the protocol are captured, data form design plays a key role in EDC implementation. EDC allows fast delivery and easy sharing of e-CRFs across trials, however, the challenges include: developing the e-CRFs in a short time frame, efficiently maintaining multi-version e-CRFs, and minimizing the downtime of the EDC system. Three EDC systems are set up to accommodate CMICR’s needs: a development site hosted on an internal virtual server; a production site hosted by a third party hosting provider; and a test site which is identical to the production site and hosted at the same location as the production site. This system setup allows for developing, testing, and managing live-running e-CRFs in parallel, which greatly shortens the development life cycle of e-CRFs and reduces the downtime of EDC when e-CRF updates are required at a minimal cost. A version control method was also developed to maintain the consistency of the version numbers across systems. This presentation will discuss the advantages and disadvantages of this three EDC system setup and the details of a version control method.
P75

COMPARING THE IMPLEMENTATION OF A MODIFIED CONTINUAL REASSESSMENT
METHOD TO A TRADITIONAL 3+3 DESIGN IN A PHASE I ACUTE MYELOID LEUKAEMIA TRIAL
(1) Christina Yap, (2) Charlie Craddock, (3) John ‘O’ Quigley, (1,4) Lucinda Billingham
(1) MRC Midland Hub for Trials Methodology Research, University of Birmingham (2) Centre
for Clinical Haematology, Queen Elizabeth Hospital (3) Université Pierre et Marie Curie (4) Cancer
Research UK Clinical Trials Unit, University of Birmingham

The majority of Phase I dose-finding trials in oncology have been dominated by the traditional up-and-down
designs, such as the 3+3 designs. However, in recent years, there is an emerging interest in implementing
more innovative model-based dose finding methods such as the Continual Reassessment Method (CRM).
Such approaches, despite being of higher statistical complexities, have demonstrated much more superior
operating characteristics in correctly identifying the right dose to take forward to a Phase II trial, ability to
expose fewer patients to potentially toxic doses and allocating more patients to the maximum tolerated
dose. In this study, we provide further support for the use of a model based design in an Acute Myeloid
Leukaemia trial. This trial studied the use of combined 5-Azacitidine and Lenalidomide as a salvage therapy
in patients with acute myeloid leukaemia who relapsed after allogenic stem cell transplantation. The dose-
limiting toxicity was defined as severe grade 3/4 Graft versus Host Disease and severe grade 3/4 non-haema-
tological toxicities considered to be related to Lenalidomide or Azacitidine. We examine the operating
characteristics of different variants of the continual reassessment method (CRM), with modifications to the
CRM that are largely guided by the challenges in this trial. We then compare these operating characteristics
with those from a traditional 3+3 design via a simulation study. Examining the operating characteristics
of the 3+3 design is seldom done in practice when designing a Phase I trial. Using this trial as an example, we
demonstrate how useful it is to guide the clinicians to the most appropriate design, when the operating
characteristics of both the model based designs and 3+3 are presented and compared.

P76

THE USE OF LOGS AND FORMS FOR THE TRACKING OF RT-CGM DEVICES IN THE CONCEPTT TRIAL
Aquila Farrell, Sonya Mergler, Dalah Mason, Johanna Sanchez, Denice S. Feig, Elizabeth Asztalos
(1) Sunnybrook Health Sciences Centre, The Centre for Mother, Infant, and Child Research, Toronto, Ontario,
Canada (2) Mount Sinai Hospital, Toronto, Canada

The Centre for Mother, Infant, and Child Research (CMICR) is the data and coordinating centre for
Continuous Glucose Monitoring in Women with Type 1 Diabetes in Pregnancy Trial (CONCEPTT), a multi-
centre, parallel randomized controlled trial (RCT) with a sample size of 324 women. In CONCEPTT, preg-
nant women or women planning pregnancy will be randomized to either receive the Real-Time Continuous
Glucose Monitor (RT-CGM) sensor added to standard therapy or standard therapy of home glucose monitor-
ning (HGM) to determine whether continuous monitoring will improve glycemic control. Since half of the
women will be randomized to receive the RT-CGM device, it is paramount to create an efficient tracking
system for the management of the devices. To document the serial number of each device assigned to the
participants, forms and logs were created to track devices assigned and in use. Three documents have been
developed to track the RT-CGM devices: 1) Device Order Forms: used by the coordinating centre to document
the devices that were sent to the site. 2) Device Inventory Form: used by the site for ongoing logging of
the inventory of the devices. 3) Device Serial Number Log: used by sites to document the serial number,
date the device was given and type of RT-CGM. In the event of device exchange, the reason for the switch is
recorded. The coordinating centre and the sites, will be able to use these tracking documents to assign
devices to a patient, order devices, and track the return of defective devices to the manufacturer. This pres-
tentation will illustrate the forms implemented to develop the monitoring systems.

P77

ENSURING ACCURATE OUTCOMES REVIEW FOR DATA ANALYSIS
Ainy Zahid, Sunny Chan, Trinh Hoac, Johanna Sanchez, Dalah Mason, Elizabeth Asztalos
The Centre for Mother, Infant, and Child Research, Sunnybrook Research Institute

The Centre for Mother, Infant, and Child Research (CMICR) is the central coordinating centre for several
multi-centre, international RCTs that aim to improve clinical practice and the health outcomes of women
and their children. Outcomes are an integral part of a clinical study; therefore, it is necessary that reported outcomes are accurately reviewed to confirm they meet the definitions outlined in the study protocol. Each trial at CMICR has a process for the review of outcomes by an outcomes review committee. A tracking system is used to record cases with reported outcomes and the decisions of the outcomes review committee. For the purposes of the data analysis, a report generated from the tracking system was used as a cross referencing tool against data extracted from a database, to ensure that all cases with reported outcomes are accounted for and correctly designated as outcomes. In this process, two programs are used to ensure that true outcomes, as defined by the study protocol, are included in the data analysis. First, a SAS program is used to extract cases where outcomes have been reported on the case report forms. Then, the reported outcomes are compared against the outcomes review committee’s final decision as recorded in the tracking system. The outcomes review committee maintains the authority on the final decision regarding the designation of the reported outcomes; therefore, it is imperative that the adjudication tracking is used in conjunction with the SAS program to confirm all outcomes are accurately accounted for. The recent tracking system has proved to be useful in ensuring accurate outcomes data for data analysis, as discrepancies were revealed upon comparison of the two lists. This presentation will illustrate the process implemented to ensure accuracy for data analysis.

**P78**

**ELECTRONIC CONVERSION AND ARCHIVING OF CLINICAL TRIAL PAPER RECORDS**
Dong Vo

*Ottawa Hospital Research Institute Clinical Epidemiology Program, Methods Centre-DMS*

Electronic Conversion and Archiving of clinical trial paper records Moving to Electronic Archiving or Backfile digital conversion can provide remarkable value to any organization. However, it is a tremendous undertaking from an administrative perspective. Too often, completed trial paper records end up taking a lot of valuable office space and it is extremely challenging to re-access the old trial records once they are locked and stored in a secure, remote location. The OHRI-CEP, Methods Centre has developed an electronic archiving system that is comprehensive in scope, yet general enough that it can be adapted for use in different clinical centres. Here are some of the benefits to which the electronic archiving system offers: Save space by replacing money-eating space with revenue-producing office space, Save time with better accessibility: old study data are readily available, Save money with efficiency. This presentation will document the many challenges inherent to designing and implementing an effective electronic archiving system. Below are some key points in planning and developing an e-Archiving system: Roles and Responsibilities Human Resources Hardware and Software Requirements Document Assessment, Type of Documents, Retention Periods, Privacy and Confidentiality, Security, Scanning, Accessibility, Validation and Attestation Process, Document Standard Naming Convention, Document Electronic Record Format, Purging of duplicate and redundant records Document Tracking End of Life Purging Process

**P79**

**A BAYESIAN ANALYSIS OF A CLINICAL TRIAL FOR TREATMENT OF UVEITIS**
(1) Natalie Nardone, (2) S. R. Rathinam, (3) Manohar Babu, (4) Thomas Lietman, (5) Travis Porco, (6) Nisha Acharya

(1) University of California, San Francisco (2) Aravind Eye Hospitals (3) Aravind Eye Hospitals (4) University of California, San Francisco (5) University of California, San Francisco (6) University of California, San Francisco

**Purpose:** The purpose of this study was to perform a Bayesian analysis of a recent randomized clinical trial comparing the relative effectiveness of two antimetabolites (oral methotrexate and oral mycophenolate mofetil), using the pre-trial opinion of a panel of uveitis experts as a prior.

**Methods:** A block-randomized observer-masked clinical trial comparing the effectiveness of the antimetabolites oral methotrexate to oral mycophenolate mofetil for non-infectious uveitis was previously conducted. The primary outcome was corticosteroid-sparing control of inflammation that was sustained for at least 28 days. Expert opinion on the most effective antimetabolite for treatment of non-infectious uveitis was queried through an online questionnaire sent to the Executive Board of the American Uveitis Society. A subjective Bayesian analysis was performed using expert opinion as the prior and the results of the clinical trial as the likelihood.
Results: Twelve uveitis experts completed the online questionnaire out of 12 surveyed. The majority (67%) thought that mycophenolate mofetil was the more effective antimetabolite. A frequentist analysis of the trial found that 71% of methotrexate treated patients achieved treatment success, compared to 53% success with mycophenolate (OR 2.2, p=0.14). The Bayesian analysis posterior suggested that methotrexate offered 1.5-fold better odds for success (95% credible interval of 0.6 to 4.9-fold), with an estimated 71% belief that methotrexate was more effective.

Conclusions: While the frequentist analysis of the trial indicated that methotrexate was more effective, though not significantly so, the expert opinions indicated that mycophenolate mofetil would be most effective. The Bayesian analysis suggested that there was little evidence that one agent was better than the other. The clinical equipoise demonstrated in the Bayesian analysis allows justification for a larger clinical trial (NEI: U10 EY021125-01).

P80
DYNAMIC WEB-BASED MANAGEMENT OF A LIFESTYLE INTERVENTION: THE STUDY OF NOVEL APPROACHES TO PREVENTION CLINICAL TRIAL
(1) Leah Griffin, (2) Karen Erickson, (3) Erica Ferguson, (4) Letitia Perdue, (5) Mark Espeland
Wake Forest Baptist Health

The Study of Novel Approaches to Prevention (SNAP) is a two-center randomized, controlled clinical trial of 599 young adults designed to compare two approaches to weight gain prevention. Participants were randomly assigned to one of three conditions: interventions targeting large or small behavioral changes plus self-regulation or a control group. Both intervention groups participated in an initial four month in-person program which included eight weekly group sessions followed by two monthly meetings. Following the initial intervention period, continued contact with study participants is driven by regular reporting on a group-specific intervention website. The main purpose of the website is for regular reporting of body weight. Participants also have access to treatment lessons for their group, quarterly newsletters, weekly tips, and a diary to report diet and exercise behaviors. Participants are also encouraged to use the website to submit their current weights at least weekly. They also have the ability to submit additional information such as exercise minutes, caloric intake, and number of diet changes made. In order to track overall and individual progress and monitor adherence, we developed an extensive series of interactive intervention reporting tools that take the information entered by the participant and report that data back to the clinics. The reports provide the clinics with up-to-date information including details such as website usage, weight tracking, diary submissions, and diary content. Both summaries and detailed information on each individual and their intervention history are provided. With these reporting tools, clinics monitor participants and continuously provide them with the feedback they need to maintain intervention participation such as personalized correspondence based on weight and website usage, tip sheets, and offers for individual counseling. Web-based monitoring provides an efficient means to track adherence to behavioral interventions in multi-center trials.

P81
THE TRANSFORMING OF MULTICENTER CLINICAL TRIALS TO OBSERVATIONAL FOLLOW-UP STUDIES
(1) Lea Drye, (2) Anne Casper, (3) Janet Holbrook, (4) Gabrielle Jenkins, (5) Curtis Meinert
(1) Johns Hopkins Bloomberg School of Public Health, Epidemiology Department and Center for Clinical Trials
(2) Johns Hopkins Bloomberg School of Public Health, Epidemiology Department and Center for Clinical Trials
(3) Johns Hopkins Bloomberg School of Public Health, Epidemiology Department and Center for Clinical Trials
(4) University of Alabama at Birmingham School of Public Health, Epidemiology Department (5) Johns Hopkins Bloomberg School of Public Health, Epidemiology Department and Center for Clinical Trials

Investigators may elect to continue to follow participants in an observational study after the trial portion of a clinical trial comes to an end. The additional follow-up may be initiated because of a need for more data on adverse events, efficacy, or costs related to treatment, or for reasons unrelated to treatment such as to observe the natural history of the disease taking advantage of the established cohort from the trial. However, the transition from trial to follow-up study can be complicated by issues in funding, maintaining contact with and
cooperation of participants and the relevance of the scientific question. At the Johns Hopkins Coordinating Centers, several trials have transitioned to observational follow-up. The Alzheimer's Disease Anti-inflammatory Prevention Trial (ADAPT) follow-up study was carried out to examine whether a late emerging trend toward decreased risk of Alzheimer's Disease in the naproxen treatment group was sustained in the long term and to compare mortality across treatment groups. The Childhood Asthma Management Program (CAMP) continuation studies were designed to assess long term treatment effects of budesonide and nedocromil on physical growth and development of children as well as to investigate the natural history of asthma and repercussions of persistent asthma. The objective of the Multicenter Uveitis Steroid Treatment (MUST) trial follow-up study is to extend the comparison of the effects of fluocinolone acetonide intraocular implant versus systemic steroid therapy for chronic uveitis as well as comparison of the treatments costs and to collect data on need for re-implantation. Issues encountered and lessons learned from these transitions and other trials will be reviewed. The authors will discuss issues such as obtaining funding and funding cuts, IRB and consent requirements, recruitment and enrollment and difficulty in combining trial and follow-up data.

P82
INTEGRATING STUDY DATA FROM MULTIPLE SOURCES USING SQL SERVER - THE ACCORDION STUDY
Jason Griffin, John Hepler, Gregory Evans, Robert P. Byington
Wake Forest Health Sciences

The Action to Control Cardiovascular Risk in Diabetes (ACCORD) Trial was a randomized clinical trial in which 10,251 people with type 2 diabetes were treated and followed for an average of approximately 5 years from 2001 through mid-2009 to determine whether intensive risk factor management could reduce cardiovascular disease risk. A prospective observational follow-up study of ACCORD participants called ACCORDION is now underway and is designed to clarify the long-term effects of the ACCORD treatment strategies and provide additional data on the long-term relationships among various cardiovascular and diabetic risk factors. Data collection for ACCORDION includes clinical data from participant visits entered remotely via a web interface, as well as data from central agencies, including a central laboratory and fundus photograph reading center, which transfer data to the coordinating center on a periodic basis in csv, excel and xml formats. Merging data from multiple sources can be a time consuming and tedious process. Incoming data must be gathered, cleaned and organized in a way that will ensure an efficient structure for study management, reporting and analysis. Once data are received and stored in a secure location, automated processes are put in place to import and document the insertions of new data, along with corrections to existing data. Sql Server 2008 provides a set of robust tools and technologies that can simplify and improve the processes needed to this end. This presentation will describe the methods and tools within Sql Server 2008 which are used to import and integrate data within the ACCORDION study.

P83
MISSING DATA IN CLINICAL TRIALS
Vicki Lu
Boston Biomedical Associates

Missing data is a prevailing problem in clinical studies and the credibility of clinical trial results could be substantially undermined by missing data. Missing data can arise during any stage of a clinical study due to a variety of reasons including patient dropout, incomplete data collection, and invalid or missed lab measurements. In practice, frequently used approaches to address missing data primarily focus on handling missing data during the data analysis phase of a study. However, the prevention and avoidance of missing data are more important in addressing the problem in clinical trials. Domestic and international guidelines and recommendations on missing data in clinical trials have been published. The current presentation will review strategies and steps that could and should be taken to reduce the amount of missing data and the impact of missing data in clinical trials. The focus is on three critical elements: (1) careful trial design; (2) good trial conduct; (3) appropriate data analysis in handling missing data.
CREATING A NEW SCORING ENDPOINT OF CARDIAC EVENTS AND
QUESTIONNAIRE-BASED HEART FAILURE SYMPTOM

(1) Akiko Kada, (2) Masanori Asakura, (3) Hiroyuki Uesaka, (4) Kaori Doi, (5) Haruko Yamamoto,
(6) Masafumi Kitakaze

(1) National Cerebral and Cardiovascular Center (2) National Cerebral and Cardiovascular Center
(3) Osaka University, National Cerebral and Cardiovascular Center (4) National Cerebral and Cardiovascular
Center (5) National Cerebral and Cardiovascular Center (6) National Cerebral and Cardiovascular Center

Background: It is difficult to evaluate the efficacy of treatments for acute heart failure. According to “The
guidelines for methods of clinical evaluation of anti-cardiac failure drugs”, essential efficacy variables should
be clinical signs/symptoms, hemodynamics, prognosis, and improvement of quality of life (QOL). But there
is no single endpoint to measure these variables simultaneously. So we create a new endpoint judging the
effectiveness of treatment for acute heart failure in short-term outcome.

Methods: We searched clinical trials of acute heart failure, and selected several measurements of cardiac
events and QOL. We determined how to combine them, and the characteristics of the combinations. Sample
size calculation and power analysis were also examined.

Results: Composite scoring endpoints of all-cause mortality, hospitalization due to heart failure and ques-
tionnaire-based heart failure symptom (Kansas City Cardiomyopathy Questionnaire: KCCQ) at 3 months
after the start of drug administration were evaluated. The summary score by KCCQ (scale from 0 to 100) at
3 months after the start of drug administration is divided into 10 points to give a score of -2.5, -2.0, -1.5, -1.0,
-0.5, 0, 0.5, 1.0, 1.5, or 2.0. If hospitalization due to worsening of heart failure is observed during 3 months
after randomization, one point is deducted from the above score. And if all-cause death within 3 months
after randomization, the score is given with -4.

Conclusion: We created the new composite scoring endpoint of all-cause mortality, hospitalization due to
heart failure and KCCQ. The endpoint will be confirmed in a current trial of aldosterone antagonist in acute
heart failure exploratory.

DIGITAL DATA COLLECTION – THE LIFE EXPERIENCE

Scott Rushing, Delilah Cook, Kate Youngman, Erica McDavitt

Wake Forest University School of Medicine, Wake Forest University School of Medicine,
Stanford University, Tufts University

Digital data collection technologies afford clinical researchers the ability to capture data that would oth-
erwise be impractical to obtain. Consider collection of physical activity levels of subjects in an interven-
tion. Subjects could record their activity manually at set time-points during the day for estimated levels
of activity, or they could wear an accelerometer and gather precise activity levels for the entire day. This
new era of digital data collection allows us to more easily obtain detailed analytic measures, but comes
with challenges of its own. In this presentation, we will explore the challenges faced during a current
trial, and describe procedures and tools that were put in place to mediate those challenges. The Lifestyle
Interventions and Independence for Elders (LIFE) Study is a phase 3, randomized controlled trial being
conducted to compare a moderate-intensity physical activity program to a successful aging health educa-
tion program in 1,635 sedentary older adults who are followed for an average of 2.7 years. LIFE Study
subjects interact with computerized data capture tools including accelerometers for physical activity and
two interactive programs for cognitive function and self-perception of mobility. Each of these proprietary
systems collect and store data to a local computer which is systematically transferred to the Coordinating
Center for analysis. While data transfers sound like a trivial task in today’s environment, in LIFE we expe-
rienced a number of challenges in transmitting and reconciling these data. In this presentation we will
discuss systematic ways to ensure accurate and consistent electronic data transmission. As the medical
research field continues to adopt new data collection technologies, our ability to maximize the quality of
electronic datasets will become exceedingly important. Our discussion of the procedures we implemented
in LIFE will guide a conversation about how we can fully benefit from the rich data sets these technolo-
gies afford us.
P86
UTILIZING R FOR THE GRAPHICAL REPORTING OF ADVERSE EVENT CLINICAL TRIAL DATA TO FACILITATE DSMB REVIEW
Miguel Villarreal, James Rochon, Jeremy Wildfire, Agustin Calatroni, Tee Bahnson, Rebecca Zabel, Katy Jaffee
Rho, Inc.

Rho serves as the Statistical and Data Coordinating Center (SDCC) for the Immune Tolerance Network (ITN) and the Inner City Asthma Consortium (ICAC) and is responsible for generating materials for Data Safety Monitoring Board (DSMB) review. Depending on the size of the clinical trial, reviewers can sometimes be presented with an expansive amount of Adverse Event (AE) incidence summary data in tabular format. Tables and supporting listings can be overwhelming and ineffective in identifying safety signals and can distract the reviewer from absorbing meaningful patterns in the data. In order to facilitate and expedite safety data review, it is necessary to condense AE incidence into discernable chunks using graphical techniques. Utilizing R, an open-source statistical and computing language and environment, we expand on the graphical techniques prompted by Dr. Frank Harrell and summarize graphically the hierarchical classification of coded verbatim AE terms into System Organ Class (SOC) and Preferred Term (PT) using annotated dotplots. This technique combines the structural familiarity of a summary table with the visual economy of a graphic and conveys a considerable amount of information in a compact, intuitive, and readable manner. This makes it easier for the reviewer to quickly absorb meaningful patterns in the data and discern safety signals that could potentially be obscured by conventional tabular displays. Through the use and flexibility of R graphics, this method of reporting Adverse Events can be expanded and applied to a broad range of clinical trials and reproduced for other DSMB reports. Several examples of AE plots will be presented, and the pros and cons associated with using R will be discussed. This project is funded by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, under contracts HHSN272200800029C and HHSN2722010000521.

P87
THE EVALUATION OF MEDICAL DIAGNOSTIC TESTS IN CLINICAL STUDIES USING LOCAL LIKELIHOOD DENSITY ESTIMATION AND ROC CURVES
(1) Norberto Pantoja-Galicia, (2) Mary Thompson
(1) Food and Drug Administration (2) University of Waterloo

An important component in the evaluation of diagnostic tests in clinical studies is the estimation of the distribution of diagnostic test results for the diseased and non-diseased subjects in a study population. The intuitive appeal of kernel density estimation makes it an attractive candidate to obtain nonparametric estimates of such probability density functions. However, in spite of the advantages presented by standard kernel density estimation techniques, a problem of increased bias at and near a known boundary is present in certain scenarios. We develop a local likelihood estimation method that is suitable for a density with a discontinuity at a boundary. We propose an estimator that corrects boundary bias. Applications in the context of diagnostic test evaluation in clinical studies using ROC curve analysis are discussed and displayed.

P88
ENSURING COMPLIANCE TO GOOD CLINICAL PRACTICE REQUIREMENTS IN MULTI-CENTRE STUDIES
Asma Qureshi, Melissa Brown, Johanna Sanchez, Dalah Mason, Elizabeth Asztalos
The Centre for Mother, Infant, and Child Research Sunnybrook Research Institute

The Centre for Mother, Infant, and Child Research (CMICR) is the central coordinating centre for several large, multi-centre, international randomized controlled trials that aim to improve clinical practice and the health outcomes of women and their children. As the sponsor for multiple regulated trials, it is important to ensure compliance to the International Conference of Harmonization-Good Clinical Practice (ICH-GCP) requirements and Health Canada Division 5 Food and Drug Regulations, at all sites. Efforts to maintain the collection of GCP essential documents from each person involved with the trial included constant
communication with the sites via email, phone calls, site visits, and monthly newsletters. Attempts to retrieve essential documents and keep certificates and licenses up to date were intensified as protocol amendments occurred, along with the inclusion of international sites that have their own national regulatory requirements. As collection and management of these numerous documents became increasingly difficult, a new tracking system was required to increase efficiency and effectiveness of this process. A solution to this problem was to develop a process to automate site reminders of outstanding essential documents. The process involved creating a database to track and manage all of the required documents listed for each phase of the trial. Once created, the received documents are inputted into the database and the date is recorded. A program was then created to remind the sites to submit their renewals annually. This system would generate an email for each site, which would notify them of any outstanding documents. The introduction of automated tracking sheets has allowed for documents to be recorded and tracked in one place efficiently and reduced the potential for oversight.

P89

NO-FAULT COMPENSATION TO PARTICIPANTS IN UN-NOTIFIED CLINICAL TRIALS INCLUDING STEM CELL STUDY IN JAPAN

Toshinori Murayama, Eriko Sumi, Manabu Minami, Toshiko Ihara, Masayuki Yokode
Clinical Innovative Medicine, Translational Research Center, Kyoto University Hospital

CIOMS International Ethical Guidelines for Biomedical Research Involving Human Subjects (2002) state that investigators should ensure that research subjects who suffer injury are entitled to free medical treatment and financial assistance as compensates them equitably for any resultant impairment. While Helsinki Declaration of World Medical Association (2008) describes that the protocol should include information regarding provisions for treating and/or compensating subjects who are harmed as a consequence of participation in the research study. This no-fault compensation is a different notion from legal liability/indemnity/reparation due to malpractice or negligence in clinical trials. Un-notified clinical trials to the authorities are still allowed in Japan, other than IND/IDE trials. The Ethical Guidelines for Clinical Studies, the only regulation for those, were fundamentally revised and enacted in April 2009, which obligate researchers to take measures on compensation such as insurance in clinical trials to assess pharmaceuticals or medical devices. The Guidelines for Clinical Studies utilizing Human Stem Cells were also revised and enacted in November 2010, which require researchers to conduct compensation. Since casualty insurance companies have not accumulated know-how to estimate the risk of un-notified trials besides IND/IDE trials, compensation insurance remains inadequate in quality; 1) Medical expense or medical allowance cannot be paid. 2) There are a considerable range of exceptions for such clinical trials as employ anticancer agents, immunosuppressants, cells/tissue and implantable devices in this insurance. 3) Malpractice or Product Liability is exempted from the insurance responsibility. In order to overcome the situation, we have investigated legal restriction for insurance, medical expense reduction system in the academic hospital, and an academic guideline for compensation in researcher-initiated un-notified clinical trials with a casualty company, so that we could contribute to improving participants protection.

P90

REPRODUCIBLE RESEARCH METHODS WITH THE SAS SYSTEM AND MICROSOFT WORD

Paul A. Thompson
Sanford Research/USD

Reproducible research methods are defined as performing the analysis and data management components of a project in a way which enables the process to be repeated strictly and easily. With the use of R, sweave, and LaTeX, this can be accomplished due to the entirely scripted nature of all components of the analysis and documentation systems. With SAS, Word .rtf file, and Adobe .pdf files, this has been previously less do-able. The system defined within Sanford Research/USD for DSMB reports is described. This system uses SAS to prepare both the report and perform the analysis. As SAS can be considered a scripting engine as well as an analysis engine, the entire process can be “canned-up” within SAS. While more attention has been paid to R, sweave, and LaTeX, these tools are often confined to the more sophisticated end of the analysis spectrum. Using the methods described here, all parts of the process can be done within SAS itself.
P91

COMPARISON OF A LINEAR MIXED EFFECTS MODEL AND MIXED MODEL REPEATED MEASURES FOR ANALYZING INCOMPLETE LONGITUDINAL DATA IN CLINICAL TRIALS

(1) Amit K. Chowdhry, (2) Michael P. McDermott
(1) (2) Department of Biostatistics and Computational Biology, University of Rochester School of Medicine and Dentistry

Consider a randomized clinical trial with longitudinal measurements having a multivariate normal distribution obtained from participants at a number of fixed time points, where the measurement at the last time point is the primary outcome variable. A common analytic strategy to accommodate missing data is mixed model repeated measures (MMRM), a method based on direct likelihood that treats time as a categorical variable. An alternative strategy that might make more efficient use of information from the earlier time points is a linear mixed effects model (LMEM) that specifies a functional form for the relationship between response and time. We performed a simulation study to examine the consequences on power, variance, and bias of using the LMEM strategy for analysis. This strategy would be expected to be superior to MMRM in cases where the functional form of the relationship between response and time is correctly specified, but the sensitivity to model misspecification is important to investigate. The study considered a variety of assumptions concerning the number of time points, pattern of mean responses over time, correlation structure for the repeated measurements, and rates of subject withdrawal (under the missing at random assumption). Estimates of power, variance, and bias were based on 100,000 replications of the simulation. When the response evolved linearly over time, a correctly specified LMEM yielded noticeable gains in power and precision that depended on the correlation structure; a smaller benefit was seen for the setting with a quadratic relationship between time and response. When the LMEM was misspecified, substantial bias was possible depending on the degree of misspecification. LMEM can provide a substantial increase in power over MMRM when the functional form of the relationship between response and time is simple and correctly specified, but there is a risk of substantial bias if the model is incorrectly specified.

P92

TITLES VERSUS TITLES-AND-ABSTRACTS FOR INITIAL SCREENING OF ARTICLES FOR SYSTEMATIC REVIEWS

(1,2) Farrah J. Mateen, (1,2) Jiwon Oh, (1,3) Ana I. Tergas, (1,4) Neil H. Bhayani, (1,5) Biren B. Kamdar
(1) Bloomberg School of Public Health, Johns Hopkins University (2) Department of Neurology, Johns Hopkins Hospital (3) Division of Gynecologic Oncology, Johns Hopkins Hospital (4) Department of Surgery, Howard University Hospital (5) Division of Pulmonary and Critical Care Medicine

Background: There is no consensus on whether screening titles or titles and abstracts simultaneously is the preferable screening strategy for inclusion of articles into a systematic review.

Methods: Two methods of screening articles for inclusion in a systematic review were compared: titles-first versus titles-abstracts simultaneously. Each citation found in Medline or Embase was reviewed by two physician reviewers for pre-specified criteria: the citation included (1) primary data; (2) the exposure of interest; and (3) the outcome of interest.

Results: There were 2,965 unique citations. The titles-first strategy resulted in an immediate rejection of 2,558 (86%) of the records after reading the title alone, requiring review of 239 titles and abstracts, and subsequently 176 full text articles. The simultaneous titles-and-abstracts review led to rejection of 2,782 citations (94%) and review of 183 full text articles. Inter-reviewer agreement to include an article for full text review using the titles-first screening strategy was 89-94% (kappa=0.54) and 96-97% (kappa=0.56) for titles-and-abstracts combined. The final systematic review included 13 articles, all of which were identified by both screening strategies.

Conclusions: Screening via a titles-first approach may be more efficient than screening titles and abstracts together.
COMPARISON OF BENEFITS AND HARMS OF ROFLUMILAST IN PATIENTS WITH MODERATE TO SEVERE CHRONIC OBSTRUCTIVE PULMONARY DISEASE AND A HISTORY OF EXACERBATIONS

(1) Tsung Yu, (2) Kevin M. Fain, (3) Sonal Singh, (4) Carlos O. Weiss, (5) Tianjing Li, (6) Ravi Varadhan, (7) Cynthia M. Boyd, (8) Milo A. Puhan

(1) Epidemiology, Johns Hopkins Bloomberg School of Public Health (2) Epidemiology, Johns Hopkins Bloomberg School of Public Health (3) Division of General Internal Medicine, Johns Hopkins School of Medicine (4) Geriatric Medicine and Gerontology, Michigan State University (5) Epidemiology, Johns Hopkins Bloomberg School of Public Health (6) Division of Geriatrics, Johns Hopkins School of Medicine (7) Division of Geriatrics, Johns Hopkins School of Medicine (8) Epidemiology, Johns Hopkins Bloomberg School of Public Health

Background: Roflumilast is a phosphodiesterase 4 inhibitor that reduces risk of exacerbations in patients with chronic obstructive pulmonary disease (COPD). Roflumilast was approved by the FDA and EMA but has a complicated regulatory history because of its gastrointestinal and psychiatric harms. The aim of this study was to use a multidimensional approach that considers multiple outcomes, patient characteristics and preferences to quantitatively compare the benefit and harm of roflumilast (500 mcg per day) as compared to placebo.

Methods: The outcomes, evaluated for a time horizon of 1 year, were exacerbations prevented and gastrointestinal and psychiatric harms associated with roflumilast. We used publicly available documents from the FDA and calculated, based on an approach developed by the National Cancer Institute, the net benefit/harm index per 10,000 COPD patients treated over 1 year. Each outcome was weighted by their relative importance and we considered death as a competing risk. In sensitivity analyses we explored the impact of using different weights for the importance of outcomes.

Results: The net benefit/harm indexes for each category of men and women with different age and baseline risk of exacerbations (20%, 40%, 60%, and 80% of patients having at least 1 exacerbation per year) suggest that roflumilast causes more harms than benefits overall. For example, in men age 65 and older with a baseline risk where 40% of patients having at least 1 exacerbation per year, the index is -346, which suggests more harms caused by roflumilast. Only in sensitivity analyses ignoring minor outcomes the index was positive for patients at higher risk of exacerbations.

Conclusions: Our results suggest that roflumilast causes more harms than benefits overall, even for patients at higher risk of exacerbations. The results of this multidimensional and transparent approach for benefit harm assessment of roflumilast challenge its approval in the US and Europe.