Duke-Industry Statistics Symposium

http://sites.duke.edu/biostatsworkshop2015/

Challenges and Innovations in Pharmaceutical Products Development

October 22-23, 2015

Trent Semans Center, Duke University, Durham, NC

abbvie

Boehringer Ingelheim

DUKE BIOSTATISTICS & BIOINFORMATICS

PAREXEL

sas

QUINTILES

DukeMedicine

Executive Committee

Elizabeth DeLong (Duke)
Gregory Campbell (FDA)
Stephen George (Duke)
Kerry Lee (Duke)
Olga Marshenko (Quintiles)
Claude Petit (Boehringer-Ingelheim)
Terry Sosa (Quintiles)
Frank Shen (AbbVie)
Fred Snikeris (PAREXEL)
Maura Stokes (SAS)

Organizing Committee

Joshua Betcher (Quintiles)
Cliburn Chan (Duke)
Shein-Chung Chow (Duke)
Changbin Guo (SAS)
Debbie Medlin (Duke)
Marlina Nasution (PAREXEL)
Michael Pencina (Duke)
Claude Petit (Boehringer-Ingelheim)
Guochen Song (Quintiles)
Yi Tsong (FDA)
Sharon Updike (Duke)
Xiaofei Wang (Duke)
Lanju Zhang (AbbVie)

Duke-Industry Statistics Symposium

Challenges and Innovations in Pharmaceutical Products Development

October 22-23, 2015

Trent Semans Center, Duke University, Durham, NC

The symposium is organized by the Department of Biostatistics and Bioinformatics, Duke University School of Medicine and co-sponsored by AbbVie, Boehringer-Ingelheim, PAREXEL, Quintiles and SAS. It was established to promote research and collaboration among colleagues from industry, academia, and regulatory agencies to discuss challenging issues and recent advances related to the clinical development of drugs and devices. The symposium theme this year is "Challenges and Innovations in Pharmaceutical Products Development."

The keynote speech will be given by Dr Yi Tsong from FDA on "Duality of significance tests and confidence intervals in drug development."

Day 1 (October 22) events will include 4 short (1/2 day) courses on critical topics in clinical research: 1) Adaptive Clinical Trial Design – Case Studies (Shiowjen Lee, FDA; Annie Lin, FDA); 2) Analytical Similarity Assessment in Biosimilar Studies (Yi Tsong, FDA; Shein-Chung Chow, Duke); 3) Phase II Clinical Trial Design and Dose Finding (Naitee Ting, Boehringer-Ingelheim); and 4) Biomarker Utilities in Adaptive Trials (Mark Chang, AMAG Pharmaceuticals).

Day 2 (October 23) will include the keynote speech, nine parallel sessions and a poster session. The topics of the parallel sessions are: Bioequivalence and Biosimilars, Bayesian Non-Inferiority Trials, Data Monitoring Committees, Randomized Concentration-Controlled Trials, Subgrouping Analysis, Biosimilars II, Advanced Survival Analysis, Enrichment Design for Clinical Trials, and Dose Finding and Selection in Clinical Phase.

The Department is pleased to recognize the following sponsors for this event: AbbVie, Boehringer-Ingelheim, PAREXEL, Quintiles and SAS.

Program Schedule

Short Course Schedule for Thursday, October 22, 2015

- Adaptive Clinical Trial Design – Case Studies (Dr. Shiowjen Lee, FDA; Dr. Annie Lin, FDA)
- Analytical Similarity Assessment (Dr. Yi Tsong, FDA; Dr. Shein-Chung Chow, Duke)
- Phase II Clinical Trial Design and Dose Finding (Dr. Naitee Ting, Boehringer-Ingleheim)
- Biomarker Utilities in Adaptive Trials (Dr. Mark Chang, AMAG Pharmaceuticals, Inc)

8:15am-9:00am	Welcome, Registration and Coffee		Atrium
9:00am-10:30am	Short Course 1 : Adaptive Clinical Trial Design – Case Studies	Shiowjen Lee, FDA; Annie Lin, FDA	Great Hall
	Short Course 2: Analytical Similarity Assessment	Yi Tsong, FDA; Shein-Chung Chow, Duke	Great Hall
10:30am-10:45am	Break (Coffee/Tea)		Atrium
10:45am-12:15pm	Continuation of Short Courses		Great Hall
12:30pm-1:30pm	Lunch		Atrium
1:30pm-3:00pm	Short Course 3: Phase II Clinical Trial Design and Dose Finding	Naitee Ting, Boehringer-Ingleheim	Great Hall
	Short Course 4: Biomarker Utilities in Adaptive Trials	Mark Chang, AMAG Pharmaceuticals, Inc	Great Hall
3:00pm-3:15pm	Break (Coffee/Tea)		Atrium
3:15pm-4:45pm	Continuation of Short Courses		Great Hall
5:00pm-6:30pm	Social Mixer		6th floor

Program Schedule

October 22-23, 2015		
Time	**Room**	**Schedule for October 22, 2015**
8:15am – 4:45pm	Great Hall	Short Courses 1-4 (S1 & S3 in Section A S2 & S4 in Section B)
5:00pm-6:30pm	6th floor	Social Mixer
Time	**Room**	**Schedule for October 23, 2015**
8:15am – 9:00am	Great Hall	Registration and Breakfast
9:00am – 9:30am	Great Hall Section A	Opening Remark: Elizabeth DeLong (Duke), René Kubiak (Boehringer-Ingelheim) and Terry Sosa (Quintiles)
9:30am – 10:30am	Great Hall Section A	Keynote Address: Yi Tsong (FDA) *Duality of significance tests and confidence intervals in drug development*
10:30am – 10:45am	Great Hall	Break (Coffee and Tea)
10:45am – 12:00pm	Great Hall Section A	*Session 1: Bioequivalence and Biosimilars* Organizers: Victoria Chang (Boehringer-Ingelheim) and Yi Tsong (FDA) Chair: Yi Tsong (FDA) Victoria Chang (Boehringer-Ingelheim) "Sample size determination for a three-arm equivalence trial of Poisson and negative binomial responses" Meiyu Shen (FDA) "Distributional assumptions for AUC, Cmax and Tmax" Jean Pan (Amgen) "Statistical considerations in biosimilar clinical development" Cassie (Xiaoyu) Dong (FDA) "Statistical approaches to demonstrate analytical similarity of quality attributes"
	Great Hall Section B	*Session 2: Bayesian Non-Inferiority Trials* Organizer: Guochen Song (Quintiles) Chair: Brad Ferguson (Quintiles) Guochen Song (Quintiles) "Controlling Frequentist Type I and Type II Error in Bayesian non-inferiority trials: a case study" Fanni Natanegara (Eli Lilly) "Bayesian considerations for non-inferiority clinical trials with case examples" Sujit Ghosh (NC State University and SAMSI) "Robust Bayesian methods for non-inferiority tests based on dichotomous data"
10:45am – 12:00pm	Great Hall Section C	*Session 3: Data Monitoring Committees* Organizer: Michael Pencina (Duke) Chair: Michael Pencina (Duke)

		Karim Calis (FDA) "Challenges and opportunities in data monitoring and trial oversight" **Frank Rockhold (GSK) "Risk versus benefit considerations in data monitoring"** **Bob Bigelow (Duke) "Interim data analysis: Distinguishing signal from noise"** **Susan Halabi (Duke) "Group sequential design: Uses and abuses"**
12:00pm – 1:30pm	**6th floor**	**Lunch, Poster Session and Job Fair (AbbVie, Boehringer-Ingelheim and FDA)**
1:30pm – 2:45pm	**Great Hall Section C**	***Session 4: Randomized Concentration-Controlled Trials*** **Organizer: Russell Reeve (Quintiles)** **Chair: Shein-Chung Chow (Duke)** **Seth Berry (Quintiles) "Pharmacokinetic/pharmacodynamic modeling and simulation in the design and analysis of RCCTs"** **Russell Reeve (Quintiles) "Efficiency of randomized concentration-controlled trials relative to randomized dose-controlled trials, and application to personalized dosing trials"** **Michael Hale (Baxalta) "Practical reasons your randomized concentration controlled trial might flop"**
1:30pm – 2:45pm	**Great Hall Section A**	***Session 5: Subgrouping Analysis*** **Organizer: Xuan Liu (AbbVie)** **Chair: Xuan Liu (AbbVie)** **Martin King (AbbVie) "Identifying subgroups in product labeling: Two recent case studies"** **Michael Rosenblum (Johns Hopkins University) "Optimal, two stage, adaptive enrichment designs for randomized trials, using sparse linear programming"** **Shuai Chen (University of Wisconsin) "A flexible framework for treatment scoring in clinical studies"**
1:30pm – 2:45pm	**Great Hall Section B**	***Session 6: Biosimilars II*** **Organizer: Lanju Zhang (AbbVie) and Guochen Song (Quintiles)** **Chair: Guochen Song (Quintiles)** **Lanju Zhang (AbbVie) "Statistical considerations for biosimilarity assessment"** **Thomas Gwise (FDA) "Points to consider for biosimilar clinical studies"** **Sujit Ghosh (NC State University and SAMSI) "Dynamic model based methods to test for biosimilarity"**
2:45pm –	**Great Hall**	**Break (Coffee and Tea)**

3:00pm		
3:00pm – 4:30pm	**Great Hall Section B**	***Session 7: Advanced Survival Analysis*** **Organizers: Marlina Nasution (PAREXEL) and Changbin Guo (SAS)** **Chair: Marlina Nasution (PAREXEL)** **Peter Jakobs (PAREXEL) "Analysis of recurrent adverse events of special Interest: an application for hazard-based models"** **Audrey Boruvka (University of Michigan) "Understanding the effect of treatment on progression-free survival and overall survival"** **Changbin Guo (SAS) "Current methods in survival analysis using SAS/STAT® software"**
3:00pm – 4:30pm	**Great Hall Section A**	***Session 8: Enrichment Design for Clinical Trials*** **Organizer: Jane Qian (AbbVie)** **Chair: Jane Qian (AbbVie)** **Yijie Zhou (AbbVie) "Enrichment design with patient population augmentation"** **Shu-Chih Su (Merck) "A population-enrichment adaptive design strategy for vaccine efficacy trial"** **Hui Quan (Sanofi) "Adaptive patient population selection design in clinical trials"**
3:00pm – 4:30pm	**Great Hall Section C**	***Session 9: Dose Finding and Selection in Clinical Phase*** **Organizers: Qiqi Deng (Boehringer-Ingelheim) and Joshua Betcher (Quintiles)** **Chair: Susan Wang (Boehringer-Ingelheim)** **Rebhi Bsharat (Quintiles) "Using utility index to evaluate risk-benefit of several doses to help in dose selection"** **Li Wang (AbbVie) "Enhanced understanding of MCPMod in dose-ranging studies"** **Qiqi Deng (Boehringer-Ingelheim) "A robust method to design dose ranging study, followed by modeling for dose selection"** **Yaning Wang (FDA) "Regulatory application of exposure-response analyses in dose selection"**
4:30pm – 5:00pm	**Great Hall Section A**	**Panel Discussion and Closing Remarks**

Keynote Address

Dr. Yi Tsong

Division Director for Biometrics VI
Office of Biostatistics
Office of Translational Sciences
Center for Drug Evaluation and Research, FDA

Title: Duality of significance tests and confidence intervals in drug development

Abstract: At FDA as well as industry, we promote powerful test with type I error rate controlled. It leads to various tests using different standard error or confidence interval. We often ignored that the confidence interval should be consistent with the test. It happens also often in equivalence testing when 90% confidence interval is used to decide if the two treatments are equivalent. I examined cases with normal and binary data. We will extend it to cover Poisson and survival data.

Yi Tsong received his Ph.D. in Statistics from the University of North Carolina at Chapel Hill in 1979. He did his post-doctoral training in cardiovascular prevention and biostatistics at Northwestern Medical School (1978-1980). He worked as senior statistician in pattern recognition at Lockheed Engineering and Management Company (1981-1983) and biostatistical consultant at the University of Texas Medical Branch at Galveston (1984-1987) before joining FDA. He served as team leader of postmarketing risk assessment and statistical reviewer of NDA submission of critical care and pain relief products.

He is currently the Division Director and Acting Team Leader for statistical team of Chemistry and Manufacturing Control. He specializes in postmarketing risk assessment, drug manufacturing process control and quality assurance, active control noninferiority/equivalence tests, adaptive designs and QTc trials. He received 8 CDER and 12 FDA level awards for contributions in postmarketing drug risk assessment, for advisory on CDER postmarketing risk assessment external contracts, medication errors, quality control evaluation, drug compliance, in vitro bioequivalence, drug compliance, drug abuse potential studies, setting quota of scheduled substances, adaptive design and non-inferiority tests, et al. He publishes frequently in numerous professional journals. He served as Treasurer, Board Director and President of International Chinese Statistical Association. He serves also as the Associate Editor of Statistics in Medicine and Journal of Biopharmaceutical Statistics.

Parallel Session

Parallel Session 1: Bioequivalence and Biosimilars

Organizers: Victoria Chang (Boehringer-Ingelheim) and Yi Tsong (FDA)
Chair: Yi Tsong (FDA)

Victoria Chang (Boehringer-Ingelheim) "Sample Size Determination for a Three-Arm Equivalence Trial of Poisson and Negative Binomial Responses"

Assessing equivalence or similarity has drawn much attention recently as the US pharmaceutical industry is under threat from biologics patent cliff. To claim equivalence between the test treatment and the reference treatment when assay sensitivity is well-established from historical data, one has to demonstrate both superiority of the test treatment over placebo and equivalence between the test treatment and the reference treatment. Thus, there is urgency for practitioners to derive a practical way to calculate sample size for a three-arm equivalence trial. In this paper, we derive power function and discuss sample size requirement for a three-arm equivalence trial with Poisson and negative binomial clinical endpoints as an extension to the prior research on continuous endpoints. In addition, we examine the effect of the dispersion parameter on the power and the sample size by varying its coefficient from small to large. In extensive numerical studies, we demonstrate that required sample size heavily depends on the dispersion parameter. Therefore, misusing a Poisson model for negative binomial data may easily lose power up to 20%, depending on the value of the dispersion parameter.

Meiyu Shen (FDA) "Distributional Assumptions for AUC, Cmax and Tmax"

In a typical pharmacokinetic bioequivalence study with a single dose administration, one of the drug products is a reference formulation and the other a test formulation. Each subject is administered both formulations in a randomized two-period crossover design. A concentration-time profile is determined for each subject given each formulation. Each single concentration-time profile can be modeled by a pharmacokinetic compartmental model. Many software programs exist for estimating the pharmacokinetic parameters such as the absorption rate, the volume of distribution, etc. Then, AUC, Cmax, and Tmax can be obtained from the fitted pharmacokinetic model. In spite of these elaborate pharmacokinetic models, the AUC, Cmax, and Tmax are obtained from the nonparametric method for bioequivalence assessment. In practice, the univariate response variables such as log(AUC) and log(Cmax) are often assumed to follow a normal distribution without much experimental data support. For instance, an investigation of observed pharmacokinetic studies was based on numbers of subjects

from 29 to 69 and so the power of the Shapiro-Wilk test to detect departures from either distribution (lognormal or normal) may have been limited. In this presentation, we investigate the normality assumption of log(AUC) or log(Cmax) using pharmacokinetic compartmental models typically used to describe concentration profiles over time. In particular, if data is generated using the simplest pharmacokinetic models (namely one and two compartment models), will it ultimately lead to deciding which distribution of log(AUC), log(Cmax), or log(Tmax) is most plausible?

Jean Pan (Amgen) "Statistical Considerations in Biosimilar Clinical Development"

Clinical development for a biosimilar product is aimed at demonstrating similarity to a reference biologic product. It is not intended to prove clinical safety and efficacy all over again. With this understanding, there are specific challenges to the design and analysis of biosimilar clinical studies. In this talk we will discuss several statistical strategies and challenges for biosimilar clinical studies, including selection of endpoints, determination of margins, and evaluation of the totality-of-evidence. Experiences from working with regulatory agencies on clinical development of some biosimilar molecules will be shared from a statistical perspective.

Cassie (Xiaoyu) Dong (FDA) "Statistical Approaches to Demonstrate Analytical Similarity of Quality Attributes"

In conventional equivalence testing, the equivalence margin is usually fixed, e.g. (80%, 125%) in PK studies. However, such a fixed margin may not be suitable for highly variable medicines or for testing quality attributes of biologics. Considering those practical issues, we proposed to establish the equivalence margin as a constant times the variability of the reference product. This constant is obtained by achieving a given power with a pre-specified samples sizes and the true mean difference. With this equivalence margin, test statistics of the equivalence testing on the mean values need to be carefully derived and examined. When the variability of the reference product is a known constant, we developed an exact t-statistics. When the reference variability is unknown, we need to consider the variability of the sample variance when we conduct the hypothesis testing. We developed approximate approaches, confidence interval approaches, and exact statistics. We investigated type I error rate, power function of our proposed statistical methods for each scenario.

Parallel Session 2: Bayesian Non-Inferiority Trials

Organizer: Guochen Song (Quintiles)
Chair: Brad Ferguson (Quintiles)

Guochen Song (Quintiles) "Controlling Frequentist Type I and Type II Error in Bayesian Non-inferiority Trials: a Case Study"

In phase III biosimilar studies, utilizing Bayesian method and borrowing information from historical data for the control arm can effectively reduce the sample size. From the Frequentist point of view, however, the type I error from such studies can be inflated if not properly controlled. This case study demonstrates how to control the type I and type II error together in a setting where the endpoint variable is binary and the conjugate beta prior is assumed.

Fanni Natanegara (Eli Lilly) "Bayesian Considerations for Non-Inferiority Clinical Trials with Case Examples"

The gold standard for evaluating treatment efficacy of a pharmaceutical product is a placebo controlled study. However, when a placebo controlled study is considered to be unethical or impractical to conduct, a viable alternative is a non-inferiority (NI) study in which an experimental treatment is compared to an active control treatment. The objective of such study is to determine whether the experimental treatment is not inferior to the active control by a pre-specified NI margin. The availability of historical studies in designing and analyzing NI study makes these types of studies conducive to the use of the Bayesian approach. In this presentation, we will highlight case examples for utilizing Bayesian methods in NI study and provide recommendations.

Sujit Ghosh (NC State University and SAMSI) "Robust Bayesian Methods for Non-Inferiority Tests Based on Dichotomous Data"

In a non-inferiority trial, the experimental treatment is compared against an active control instead of placebo. The goal of this study is often to show that the experimental treatment is non-inferior to the control by some pre-specified margin. The standard approach for these problems, which relies on asymptotic normality, usually requires large sample size to achieve some desired power level. This talk presents robust Bayesian approaches based on Bayes factor and posterior probability for testing non-inferiority in the context of two-sample dichotomous data. A novel aspect of the proposed Bayesian methods is that the cut-off value for Bayes factors and posterior probabilities are determined from the data that approximately controls the overall

errors. Results based on simulated data indicate that both of the proposed Bayesian approaches provide significant improvement in terms of statistical power as well as the total error rate over the popularly used frequentist procedures. This in turn indicates that the required sample size to achieve certain power level could be substantially lowered by using the proposed Bayesian approaches. [This is a joint work with Muhtar Osman and major part of the talk is based on two published papers: 1 and 2]

Parallel Session 3: Data Monitoring Committees

Organizer: Michael Pencina (Duke)
Chair: Michael Pencina (Duke)

Karim Calis (FDA) "Challenges and Opportunities in Data Monitoring and Trial Oversight"

Clinical trial oversight requires coordination and review by a number of groups and committees whose diverse focus includes safety, quality, ethics, adjudication, operations, and logistics. Although some of these evolving roles and responsibilities invariably overlap, independent data monitoring committees (IDMCs) hold a unique place in trial oversight. IDMCs periodically review the accumulating safety and efficacy data by treatment group and advise the sponsor on whether to continue, modify, or terminate a trial based on risk-benefit assessment. They also play a critical role in assessing the validity and integrity of the trial to enhance its potential to generate reliable findings. IDMCs typically oversee a single trial but occasionally review multiple related trials. Emerging functions of IDMCs include the monitoring of pragmatic clinical trials and perhaps even the entire portfolio of research related to an investigational product throughout its clinical development life cycle. IDMCs should be composed of qualified individuals with knowledge of ethical principles and expertise in biostatistics, research methodology, and relevant areas of science and clinical medicine. IDMC members must be independent of the sponsor and afforded adequate resources and flexibility to perform their duties. Roles and responsibilities of the IDMC—including contingency and communication plans—should be clearly delineated in a succinct, well-organized charter that empowers IDMC members. Although IDMCs have been established for decades, the transformation of the clinical trial landscape has created new opportunities as well as scientific and regulatory challenges in the oversight of clinical trials.

Frank Rockhold (GSK) "Benefit to Risk Considerations and Methods Applied to the Ongoing Monitoring of Clinical Trials"

The overall goal of the clinical trial is to assess a primary objective and endpoint (usually a benefit) over the background of secondary endpoints including patient safety. The objective of the IDMC is to integrate the information efficacy and safety information in some fashion to make ongoing decisions about whether to continue the trial as is, have the design altered, or prematurely discontinue based on the benefits and harms they are observing in the trial. Thus in some fashion the IDMC is tasked with creating a "benefit to risk" picture for the trial patients and future patients. The science of Benefit to risk for quantitatively summarizing completed trials (one or many) has evolved over the past decade. The purpose of this talk to is to explore how one might apply these

techniques in amore more structured and systematic way in an ongoing IDMC setting. Some things to be discussed are a review basic IDMC and B-R practices and processes, examples of how to integrate data in an evolving manner as part of data monitoring, use graphical and other methods usually used to display BR data at the end of the trial adapted to interim looks, and an example reworked in a BR framework over the life of a trial using methods outlined. Some thought questions will also be proposed around what, if any, impact of type I error adjustment (if any) on the data review and presentation and the impact of regulatory guidelines, if any, on these recommendations. The intent of this talk is to start the discussion of combining classical IDMC process with the more recent advances in Benefit to Risk methodology.

Bob Bigelow (Duke) "Interim Data Analysis: Distinguishing Signal from Noise"

The goal of many clinical trials is to reach a decision on comparative treatment effects based entirely on information from the trial. Pre-specification of endpoints, sample size, acceptable type I and II errors, and statistical analyses increase the chances that hypothetical treatment differences (or similarities) can be demonstrated in the presence of background noise. However, patient safety information from an ongoing trial is often not sufficient for an IDMC to make conclusive recommendations, and reliance on expert clinical judgment and familiarity with external data are necessary. In this presentation we will discuss challenges of interim safety assessment and consider methods to improve statistical rigor in IDMC analyses.

Susan Halabi (Duke) "Group Sequential Design: Uses and Abuses"

Group sequential design (GSD) is considered part of standard statistical practice has been developed for interim monitoring (and potential termination) of clinical trials to minimize the role of subjective judgment. Most randomized clinical trials include strategies for terminating the trial early if a treatment arm is found to be either effective or harmful to the patients. Although GSDs serve as an aid in monitoring throughout the trial, the decision to stop a trial early is complex. In this talk, the consequences of terminating a trial early will be discussed with an emphasis on statistical issues related to the estimation of the treatment effect and the analysis and interpretation of the primary and secondary endpoints. Several examples of oncology trials that were stopped early for superiority will be considered.

Parallel Session 4: Randomized Concentration-Controlled Trials

Organizer: Russell Reeve (Quintiles)
Chair: Shein-Chung Chow (Duke)

Seth Berry (Quintiles) "Pharmacokinetic/Pharmacodynamic Modeling and Simulation in the Design and Analysis of RCCTs"

Pharmacokinetic (PK) and pharmacodynamic (PD) modeling is needed to design and understand the properties of a randomized concentration-controlled trial (RCCT). We will cover the exposure-response causal chain principles underlying these designs; will explore how PK variability in concentration-time profiles impacts dose-response analyses; and will discuss how implementing RCCTs can help control for the variability in PK, improving the signal while reducing noise in the PD properties of the biological system, and consequently enhancing the trial design.

Russell Reeve (Quintiles) "Efficiency of Randomized Concentration-Controlled Trials Relative to Randomized Dose-Controlled Trials, and Application to Personalized Dosing Trials"

The literature on randomized concentration-controlled trials (RCCTs) is surveyed, comparing this trial design to the more traditional randomized dose-controlled trial (RDCT). It is shown that RCCTs require smaller sample sizes than RDCTs for the same power, and that they elicit more informative information on the exposure-response relationship. RCCTs are similar in spirit to personalized titration designs, and the relationship is explored, where it is shown that personalized titration designs have similar power, even in the face of categorical responses, such as a rheumatoid arthritis trial using the binary ACR20 as the primary endpoint.

Michael Hale (Baxter) "Practical Reasons Your Randomized Concentration Controlled Trial Might Flop"

The Randomized Concentration Controlled Trial (RCCT) is based on the idea that the clinical response to a dose of drug is mediated through exposure, and so randomizing people to different exposure targets should reduce "experimental noise", compared with randomizing to different doses. Implementing an RCCT can be very challenging, however, and few have actually been performed. This talk will consider some of the practical hurdles which must be overcome, based on the speaker's experience in designing and implementing a well-known RCCT for mycophenolate mofetil in renal transplantation.

Parallel Session 5: Subgrouping Analysis

Organizer: Xuan Liu (AbbVie)
Chair: Xuan Liu (AbbVie)

Martin King (AbbVie) "Identifying Subgroups in Product Labeling: Two Recent Case Studies"

We evaluate 2 recently approved new drugs for which subgroups were identified in product labeling for potentially different treatment. For each case, trial results and labeling decisions are reviewed in light of the EMA draft guideline on subgroups in confirmatory clinical trials. We discuss the relative contributions of various factors, including evidence of heterogeneity, biological plausibility, pre-specification, and risk of misclassification.

Michael Rosenblum (JHU) "Optimal, Two Stage, Adaptive Enrichment Designs for Randomized Trials, using Sparse Linear Programming"

Adaptive enrichment designs involve preplanned rules for modifying enrollment criteria based on accruing data in a randomized trial. Such designs have been proposed, for example, when the population of interest consists of biomarker positive and biomarker negative individuals. The goal is to learn which populations benefit from an experimental treatment. Two critical components of adaptive enrichment designs are the decision rule for modifying enrollment, and the multiple testing procedure. We provide the first general method for simultaneously optimizing both of these components for two stage, adaptive enrichment designs. We minimize expected sample size under constraints on power and the familywise Type I error rate. It is computationally infeasible to directly solve this optimization problem since it is not convex. The key to our approach is a novel representation of a discretized version of this optimization problem as a sparse linear program. We apply advanced optimization methods to solve this problem to high accuracy, revealing new, approximately optimal designs.

Shuai Chen (University of Wisconsin) "A Flexible Framework for Treatment Scoring in Clinical Studies"

To identify subgroups of patients who have different responses to different treatments, one essentially needs to investigate interactions between the treatments and covariates. Instead of using the traditional outcome-modeling approach, we propose two

alternative frameworks for treatment scoring in both observational studies and clinical trials. In particular, we construct personalized scores ranking the patients according to their potential treatment effects. In contrast to outcome-modeling, under our framework, there is no need to model the main effects of covariates. The proposed methods are quite flexible and we show that several recently proposed estimators can be represented as special cases within our frameworks. As a result, some estimators which were originally proposed for randomized clinical trials can be extended to observational studies. Moreover, our approaches allow regularization in presence of a large number of covariates. Many powerful M-estimation technologies can be used in estimation.

Parallel Session 6: Biosimilars II

Organizer: Lanju Zhang (AbbVie) and Guochen Song (Quintiles)
Chair: Guochen Song (Quintiles)

Lanju Zhang (AbbVie) "How to set up biosimilarity bounds in biosimilar product development"

FDA published three biosimilar guidances in 2012 and one guidance in 2014. With the approval of the first biosimilar product in March 2015, FDA cleared a regulatory pathway for biosimilar product development in US. This calls for stepwise development approach, including analytical biosimilarity, pharmacological biosimilarity and clinical biosimilarity. A tiered approach has been proposed for demonstrating analytical biosimilarity, requiring more statistical rigor with increasing criticality of quality attributes. Specifically, critical quality attributes in tier 1 call for a head to head comparison between reference product and biosimilar product, using a statistical equivalence test to show biosimilarity based on appropriate equivalence bounds. The key is therefore to set up these goalposts. In this talk, we will review some methods to set up equivalence bounds, with a focus on 1.5 times standard deviation of reference product, which was used in the briefing document for the FDA's first biosimilar product approval. This is a joint work with Sutan Wu from SutanStats.

Thomas Gwise (FDA) "Points to Consider for Biosimilar Clinical Studies"

The Biologics Price Competition and Innovation Act of 2009 (BPCI Act) amends the PHS Act and other statutes to create an abbreviated licensure pathway for biological products shown to be biosimilar to an FDA-licensed biological reference product. This presentation will discuss the objectives of the BPCI Act to place into context the role of clinical studies in establishing a product as biosimilar to a reference product. The recent biosimilar application reviewed by FDA's Oncologic Drug Advisory Committee will serve as an example to explore clinical study design issues particular to biosimilars, including margin determination.

Sujit Ghosh (NC State University and SAMSI) "Dynamic Model Based Methods to Test for Biosimilarity"

In recent years there has been a lot of interest to test for similarity between biological drug products, commonly known as biologics. Biologics are large and complex molecule drugs that are produced by living cells and hence these are sensitive to the

environmental changes. In addition, biologics usually induce antibodies which raise the safety and efficacy issues. The manufacturing process is also much more complicated and costly than the small-molecule generic drugs. Because of these complexities and inherent variability of the biologics, the testing paradigm of the traditional generic drugs cannot be directly used to test for biosimilarity. Taking into account some of these concerns we propose a dynamic model based methodology that takes into consideration of the entire time course of the study based a class of flexible models. The empirical results show that the proposed approach is more sensitive than the classical equivalence test approach, and require much less sample size for detecting biosimilarity. [This is a joint work with Dr. Yifang Li, Novartis Inc.]

Parallel Session 7: Advanced Survival Analysis

Organizers: Marlina Nasution (PAREXEL) and Changbin Guo (SAS)
Chair: Marlina Nasution (PAREXEL)

Peter Jakobs (PAREXEL) "Analysis of Recurrent Adverse Events of Special Interest: an Application for Hazard-Based Models"

For decades, safety risks on study or product level have been summarized by incidence estimates: typically, with N_j denoting the number of subjects who received at least one dose of study treatment j and $n_{j,x}$ denoting the number of subjects in treatment group j who experienced at least once an adverse event x (e.g., categorized as MedDRA Preferred Term), such an incidence is estimated by $n_{j,x} /N_j$ times 100%. Treatment groups have been compared by related estimates like risk difference, risk ratio -and odds ratio. Timing, duration, recurrence as well as duration of adverse events has been ignored frequently. A trend for utilizing time-to-first-event methodology (like cumulative incidence estimates and Cox proportional hazard regression models) in safety assessments has been observed over the last years, but this approach is still limited. My presentation will outline some statistical methodology for evaluating risks for recurrent (or otherwise complex) safety events of special interest, focusing on hazard-based models for counting processes and multi-state models. For example, states in a multi-state model for adverse events of special interest may be defined by administration of certain concomitant medication(s) over the course of the study (that either change the risk for such adverse events or are used to treat such adverse events). If time allows, a fictitious case study analysis will be presented as well.

Audrey Boruvka (University of Michigan) "Understanding the Effect of Treatment on Progression-Free Survival and Overall Survival"

Cancer clinical trials are routinely designed on the basis of event-free survival time where the event of interest may represent a complication, metastasis, relapse, or progression. This talk is concerned with a number of statistical issues arising with use of such endpoints including interpretation and dual censoring schemes. We consider methods to evaluate this endpoint based on the Cox model. However, even when treatment is randomized, the resulting hazard ratios have limited interpretation as causal effects. We point to some ways in which one can draw causal inferences in this particular setting. This talk is based on joint work with Richard J. Cook and Leilei Zeng at the University of Waterloo.

Changbin Guo (SAS) “Current Methods in Survival Analysis Using SAS/STAT® Software”

Interval censoring occurs in clinical trials and medical studies when patients are assessed only periodically. As a result, an event is known to have occurred only within two assessment times. Traditional survival analysis methods for right-censored data are not applicable, and so specialized methods are needed for interval-censored data. The goal of this presentation is to give an overview of these techniques and their recent implementation in SAS software, both for estimation and comparison of survival functions as well as for proportional hazards regression. Competing risks arise in studies when individuals are subject to a number of potential failure events and the occurrence of one event may impede the occurrence of other events. A useful quantity in competing-risks analysis is the cumulative incidence function, which is the probability sub-distribution function of failure from a specific cause. This presentation describes how to use the LIFETEST procedure to compute the nonparametric estimate of the cumulative incidence function and test for group differences. In addition, this presentation will describe two approaches that are available with the PHREG procedure for evaluating the relationship of covariates to the cause-specific failure. The first approach models the cause-specific hazard, and the second approach models the cumulative incidence (Fine and Gray 1999).

Parallel Session 8: Enrichment Design for Clinical Trials

Organizer: Jane Qian (AbbVie)
Chair: Jane Qian (AbbVie)

Yijie Zhou (AbbVie) "Enrichment Design with Patient Population Augmentation"

Clinical trials can be enriched on subpopulations that may be more responsive to treatments to improve the chance of trial success. In 2012 FDA issued a draft guidance to facilitate enrichment design, where it pointed out the uncertainty on the subpopulation classification and on the treatment effect outside of the identified subpopulation. We consider a novel design strategy where the identified subpopulation (biomarker-positive) is augmented by some biomarker-negative patients. Specifically, after sufficiently powering biomarker-positive subpopulation we propose to enroll biomarker-negative patients, enough to assess the overall treatment benefit. We derive a weighted statistic for this assessment, correcting for the disproportionality of biomarker-positive and biomarker-negative subpopulations under enriched trial setting. Screening information is utilized for weight determination. This statistic is an unbiased estimate of the overall treatment effect as that in all-comer trials, and is the basis to power for the overall treatment effect. For analysis, testing will be first performed on biomarker-positive subpopulation; only if treatment benefit is established in this subpopulation will overall treatment effect be tested using the weighted statistic. [Joint with Bo Yang from AbbVie]

Shu-Chih Su (Merck) "A Population-Enrichment Adaptive Design Strategy for Vaccine Efficacy Trial"

Adaptive design has the flexibility allowing pre-specified modifications to an ongoing trial to mitigate the potential risk associated with the assumptions made at the design stage. It allows studies to include broader target patient population and to evaluate the performance of vaccine/drug across subpopulations simultaneously. Our work is motivated by a Phase III event-driven vaccine efficacy trial. Two target patient populations are being enrolled with the assumption that vaccine efficacy can be demonstrated based on the two patient subpopulations combined. It is recognized due to the heterogeneity of the patient characteristics, the two subpopulations might respond to the vaccine differently. i.e., the vaccine efficacy (VE) in one population could be lower than that in the other. To maximize the probability of demonstrating vaccine efficacy in at least one patient population while taking advantage of combining two populations in one single trial, an adaptive design strategy with potential population enrichment is developed. Specifically, if the observed vaccine efficacy at interim for one

subpopulation is not promising to warrant carrying forward, the enrollment in the other population can be enriched. Simulations were conducted to evaluate the operational characteristics of different timing and futility boundaries for interim analysis. This population enrichment design provides a more efficient way as compared to the conventional approaches with several target subpopulations. If executed and planned with caution, it can improve the probability of having a successful trial. [Joint with Ivan S.F. Chan from Merck]

Hui Quan (Sanofi) "Adaptive Patient Population Selection Design in Clinical Trials"

For the success of a new drug development, it is crucial to select the sensitive patient populations. To potentially reduce timeline and cost, we may apply a two-stage adaptive patient population selection design to a therapeutic trial. In such a design, based on early results of the trial, patient population(s) will be selected/determined for the final stage and analysis. Because of this adaptive nature and the multiple between-treatment comparisons for multiple populations, an alpha adjustment is necessary. In this paper, we propose a closed step down testing procedure to assess treatment effects on multiple populations and a weighted combination test to combine data from the two stages after sample size adaptation. Computation/simulation is used to compare the performances of the proposed procedure and the other multiplicity adjustment procedures. A trial simulation is presented to illustrate the application of the methods. [Joint with Dongli Zhou, Pierre Mancini, Yi He and Gary Koch from Sanofi]

Parallel Session 9: Dose Finding and Selection in Clinical Phase

Organizers: Qiqi Deng (Boehringer-Ingelheim) and Joshua Betcher (Quintiles)
Chair: Susan Wang (Boehringer-Ingelheim)

Rebhi Bsharat (Quintiles) "Using Utility Index to Evaluate Risk-Benefit of Several Doses to Help in Dose Selection"

In early phase studies where the sample size is small and several endpoints are used for evaluation to choose candidate doses for latter stage of drug development, the challenge is choosing the best doses that have maximum efficacy and the best safety profile. Treatment arms including active control and/or placebo could give conflicting messages when evaluated based on different endpoints. A technique is presented to summarize overall utility of each dose and compare different doses to placebo or active control using a clinical utility index which is a multivariate utility function that summarizes the utility for each subject across all endpoints. Active doses are compared to placebo or active control using bootstrap confidence intervals. The technique supports informed-decision making based on evaluation of different scenarios using simulation.

Li Wang (AbbVie) "Enhanced Understanding of MCPMod in Dose-Ranging Studies"

MCPMod (Bretz et al, 2005 and Pinheiro et al. 2014) is the approach to provide additional insights for the selection of the "best" underlying dose-response model and controls FWER at the POC stage for model selection. The target dose is selected from the final model and is not bounded in the candidate set of the doses evaluated in the trial. They enable the more informative Phase II trial study design to provide a more solid basis for all subsequent dose selection strategies and decisions. Specifically, MCPMod approach receives regulatory supportive opinion, e.g. EMA – CHMP qualification opinion on 10/01/2013. In this research, we further evaluated the performance of MCPMod on MED estimation using weighted and unweighted AIC criteria and the impact of number of doses and different prior assumptions on model selection and restricted MED estimation.

Qiqi Deng (Boehringer-Ingelheim) “A Robust Method Using Ordinal Linear Contrast Test to Design Dose Ranging Study”

Nowadays, many sponsors are working to speed up the clinical development process. A commonly used strategy is to combine the Proof of Concept (PoC) and the dose-ranging clinical studies into a single trial at the early Phase II development. In such trials, the primary objective is to establish the POC and make go-no go decision. And the important secondary objective is to identify a range of doses to move into phase III. We propose to use ordinal linear contrast test (also referred to as trend test) to design such a trial, which is easy to communicate to non-statisticians, simple to implement, and provides robust performance for different dose response curves under monotonic assumption. We will also discusses the implication of different ways of allocating patients to each treatment group, under a given total sample size – which is often limited by budget and ethical concerns.

Yaning Wang (FDA) “Regulatory Application of Exposure-Response Analyses in Dose Selection”

Exposure-response analyses are routinely conducted by Pharmacometric reviewers at FDA to address the key question: is the dose/dosing regimen selected consistent with the exposure-response relationships for both efficacy and safety? Such analyses are used to support the approved dose/dosing regimen and justify additional studies such as post-marketing requirement (PMR) or post-marketing commitment (PMC) studies to further optimize the dose/dosing regimen. Case studies will be shared to demonstrate the application of exposure-response analyses in regulatory decision making process.

Poster Session

Who will benefit from antidepressants in the acute treatment of bipolar depression? A follow up observational data analysis of STEP-BD

Fan Wu (NCSU)

There is substantial uncertainty regarding the efficacy of antidepressants in the acute treatment of bipolar depression. In our recent paper (Wu et al., 2015, in press, International Journal of Bipolar Disorders), by using data from the acute depression randomized care (RAD) pathway of the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) study (Sachs et al., 2007, NEJM), we estimate an optimal dynamic treatment regime via Q-learning. The estimated optimal dynamic treatment regime presents some evidence that patients in RAD pathway of STEP-BD with a (hypo)manic episode prior to onset of the current depressive episode should not be given an antidepressant in addition to a mood stabilizer, while all the other patients would benefit from an additional antidepressant. The goal of our current analyses is to validate this finding, using an independent sample, but a similar methodology regarding outcome criteria and rating instruments. In STEP-BD study, there is another pathway named standardized care pathway (SCP), which is an observational study. We construct a dataset named SAD (Standard Acute Depression) Pathway, which contains patients in SCP pathway that satisfy RAD pathway entering criteria. By using this dataset, we construct three different models to validate the performance of the findings in RAD pathway via Inverse Probability Weighting (IPW) as well as Augmented Inverse Probability Weighting (AIPW), which are two popular methods used to estimate the mean outcome of a treatment regime. The analyses are still ongoing, but preliminary results indicate that there is some uncertainty in reproducing the findings from RAD, and that the results are dependent on the estimation method used. There are some differences between RAD dataset and SAD dataset, which may, besides other variables, affect the results from SAD data analysis. In the future, we will estimate optimal treatment regime based on SAD dataset.

TumorENVO a standardized representation of the tumor microenvironment

Anna Maria Masci (Duke)

As all other tissues, tumor is made by a mixture of parenchymal and stromal cells, that include a variety of immune and non-immune cell types. The complex network of physical and functional interactions between these cell populations regulates the development of the blood and lymphatic vasculature, as well as the chemical-physical properties of the extracellular matrix. All these cellular and extracellular factors can be included in the broad definition of Tumor Microenvironment (TME). The tumor microenvironment (TME) is being increasingly recognized as a key factor in multiple stages of disease progression, particularly local resistance, immune-escaping, and distant metastasis, thereby substantially impacting the future development of frontline interventions in clinical oncology. An appropriate understanding of the TME promotes evaluation and selection of candidate agents to control malignancies at both the primary sites as well as the metastatic settings. In the omics era, a massive production of data on tumors composition are generated in relatively short time, and this is offering the unique opportunity to shed light on the role of TME in mechanisms regulating growth, invasiveness and spreading of human tumor. To maximize their usability, these large data sets require to be integrated, and this is not a trivial process. The absence of standardized, computable representations of TME and its influence in progression of cancer poses a significant barrier to progress in research, diagnosis and treatment of cancer. Standardized vocabularies are needed to ensure the consistent use of terminology, thereby facilitating the sharing of research results between subspecialties and the translation of research results into clinical practice. Machine processable representations are of further value and are becoming critical as biomedical research becomes ever more reliant on computational assistance in the management and analysis of data and information. Standardized, computable representations, here referred to as ontologies, are at the foundation of methods for sharing data and making heterogeneous data resources interoperable. To address this gap in the information and computational resources critical to progress in cancer research, we are developing the Tumor micro Environmental Ontology The potential impact of this ontology is significant, broadly impacting both the research and clinical care communities.

Separating Variability in Practice Patterns from Statistical Error; an Opportunity for Quality Improvement

Laine Thomas (Duke)

Quality improvement studies seek to establish the degree of variability in practice patterns and outcomes across different providers. Wide variation suggests that institutional factors play a role in affecting outcomes, and high performing institutions should be studied and emulated. Therefore, the magnitude of variation is a key parameter of interest. Despite the extensive literature on methods for hospital monitoring and profiling, variability across providers is usually displayed in figures and histograms using techniques that either over-estimate or under-estimate the actual degree of variation. As a result, conclusions regarding the extent of variation based on these figures may be wrong. Instead, the distribution across providers can be estimated directly from a hierarchical model, where provider is included as a random effect. Hierarchical models are commonplace for other objectives, such as hospital-specific prediction, and may substantially improve the estimation and illustration of inter-hospital variability. However, the predominant Gaussian hierarchical model imposes an assumption of normally distributed variation across providers. This and other parametric models will not be adequate when unknown features of the distribution, such as bi-modality or skewness, are of particular interest. Semi-parametric Bayesian methods for density estimation offer a flexible alternative. In addition, we cast this as a measurement error problem and apply a recently developed method for density estimation in the presence of measurement error. Alternative approaches are compared by simulation and results are interpreted in the context of a motivating example.

Non-mixture cure model with left truncation and complex censoring and its application for a spontaneous abortion data set

Yuan Wu (Duke)

We propose a semi-parametric spline-based estimation for non-mixture cure model when left truncation, interval censoring and observed events all appearing. To our

knowledge this spline-based approach is the first attempt to handle this type of complex survival data structure in non-mixture cure model in only one step. The semi-parametric estimation is obtained by a constrained maximization algorithm. The observed efficient Fisher Information for the parametric part is also easily estimated for the spline-based estimation. The proposed approach is applied to a spontaneous abortion data set.

An evaluation of constrained randomization for the design and analysis of group-randomized trials

Fan Li, Yuliya Lokhnygina, David M. Murray, Patrick J. Heagerty, Elizabeth R. DeLong (Duke)

In group-randomized trials, a frequent practical limitation to adopting rigorous research designs is that only a small number of groups may be available, and therefore simple randomization cannot be relied upon to balance key group-level prognostic factors across the comparison arms. Constrained randomization is an allocation technique proposed for ensuring balance, and can be used together with a permutation test for randomization-based inference. However, several statistical issues have not been thoroughly studied when constrained randomization is considered. Therefore, we used simulations to evaluate key issues including: the impact of the choice of the candidate set size and the balance metric used to guide randomization; the choice of adjusted versus unadjusted analysis; and the use of model-based versus randomization-based tests. We conducted a simulation study to compare the type I error and power of the F-test and the permutation test in the presence of group-level potential confounders. Our results indicate that the adjusted F-test and the permutation test perform similarly and slightly better for constrained randomization relative to simple randomization in terms of power, and the candidate set size does not substantially affect their power. Under constrained randomization, however, the unadjusted F-test is conservative while the unadjusted permutation test carries the desired type I error rate as long as the candidate set size is not too small; the unadjusted permutation test is consistently more powerful than the unadjusted F-test, and gains power as candidate set size changes. Finally, we caution against the inappropriate specification of permutation distribution under constrained randomization.

General Statistical Methodologies for Adaptive Design

Liddy Chen (PAREXEL International)

The Pharmaceutical Research and Manufacturers of America (PhRMA) working group on adaptive design defined adaptive designs as "a clinical study design that uses accumulating data to decide how to modify aspects of the study as it continues without undermining the validity and integrity of the trial (Dragalin 2006)." Statistical methods are mainly designed to safeguard validity, which means to provide the correct statistical inferences. E.g., the overall pre-specified Type I error rate must be maintained. In general, varies statistical methods to control Type I Error in adaptive clinical trials can be classified into the following categories: 1) repeat testing such as group sequential trials, 2) multiple hypothesis testing or multiple comparison, 3) combination of pre- and post- adaptation data. In more complex adaptive trials where statistical methodologies are not available to ensure the Type I Error rate is controlled at its nominal level, simulations are usually used to demonstrate the operating characteristic. Methods for repeat testing includes Pocock (1977), O'Brien and Fleming (1979), and the alpha spending function developed by Lan and DeMets (1983) etc. The stepwise multiple comparison procedures based on the closure principle (Marcus et al. 1976), also known as closed testing methods, can also be applied to adaptive clinical trials. Another idea is to combine stage-wise p-values to allow for adaptations. These combination methods can be applied to different adaptive designs to provide combined critical values or to combine p-values from pre- and post-adaptation. Methods include the Fisher's combination function by Bauer and Kieser (1999), inverse normal function by Lehmacher and Wassmer (1999), conditional error function by Proschan and Hunsberger (1995), down weighting combination strategy by Cui et al (1999). In complex designs, more than one statistical methods can be used together to ensure the Type I Error rate is not inflated. The Bayes theorem in which the posterior distribution depends on the prior knowledge fits naturally in the adaptive design framework. The Bayesian methods, such as the Continual Reassessment Method (CRM) Design (Sweeting et al. 2013) and Normal Dynamic Linear Model (NDLM) approach (Smith et al. 2006) had been widely used in adaptive designs.

Detection of immunogenic reactions in clinical trials of biosimilars

Marek Ancukiewicz (PAREXEL International)

Immunogenicity assessment is required for the approval of biosimilar products both by FDA an EMA. Both innovator drug and biosimilar drug have potential for inducing anti-therapeutic antigens. We consider the identification of individual patients participating in clinical comparison of biosimilar and innovator drugs, based on 1) pre-clinical and clinical evaluations of immugenocity and 2) covariate information such as patient characteristics affecting immunogenicity, and possibly on other data such as lab measurements or T-cell counts. We consider here a problem of non-parametric, Bayes classification rule for detection of immunogenic reactions in individual patients taking part in clinical comparative trials of biosimilars, using a vector of features. Assume two treatment groups and (patients randomized to comparator and biosimilar products). We model, in each group, the density of features as a mixture from individuals with and without immunogenic reactions, assuming mixing proportions and for comparator product and for the biosimilar product. We assume that mixing proportions are known. The proposed solution is based inversion of the mixture equations: and then a classification rule such that the density is maximized. We provide methods to estimate the error rote, based on procedure described in Ancukiewicz, Marek. “An Unsupervised and Nonparametric Classification Procedure Based on Mixtures with Known Weights.” Journal of Classification 15, no. 1 (January 1, 1998): 129–41.

Evaluating methods of estimating common risk difference for stratified binomial clinical trials for less common events

Kate Fisher (PAREXEL International)

Many clinical trials use a stratified randomization approach to ensure balanced treatment assignment within subgroups, helping to more accurately estimate treatment effect. For example, a trial may stratify the randomization on gender, age groupings, or some other covariate that investigators believe may influence treatment effect on a binomial outcome. For stratified binomial trials, a risk difference can be calculated per stratum. These estimates can be combined to produce a common risk difference to describe the overall treatment effect. However, there are many ways to estimate the

common risk difference and confidence interval. Popular weighting schemes include using Cochrane-Mantel-Haenzel, inverse of variance, and minimum risk. Confidence intervals can be constructed using a variety of methods including Wald-Type(Sato), stratified Newcombe, Miettinen-Nurminen(Score), and a new approach presented by Klingenberg. We simulate data using 4 strata and an overall event probability similar to an actual trial. Across all simulations we keep the underlying overall event probability and the sample sizes per stratum and treatment group stable. We simulate multiple scenarios exploring how 1) the magnitude of risk difference and 2) the magnitude of heterogeneity of risk difference affect the performance of various common risk difference estimation methods. We vary the magnitude of the difference between the treatment and control arm probability of events to explore setting 1. For setting 2, we introduce heterogeneity by sampling from a uniform distribution centered at 0 to add noise to the treatment arm probability of event for each stratum. We capture how often 0, indicating an insignificant result, and how often the true underlying risk difference is captured in the confidence interval to help determine if certain methods outperform others across simulations.

Impact of biosimilarity predetermined margin on interchangeability

Hyang Kim (PAREXEL International)

Biological drug products are therapeutic moieties manufactured by a living system or organisms. Generic versions of biological products (or reference) have been produced, referred to as follow-on biologics (or biosimilar) drug products and it is expected to produce the same clinical results as the reference product in any given patient. The fundamental bioequivalence assumption is that the pharmacokinetic similarity between the reference and biosimilar product characteristics is extrapolated to the similarity in efficacy or safety endpoints. In the process of evaluating the similarity with the reference biological product, biosimilarity margin in product characteristic is predetermined. However, depending on the biosimilarity predetermined margin, the interchangeability may be no longer valid. We propose an unbiased test procedure to evaluate the extrapolation of the similarity in product characteristics such as pharmacokinetic responses to similarity in efficacy/safety endpoints in any given patient. It is assumed that similarity in product characteristics is found in assessing biosimilarity and that relationship between a reference product and efficacy/safety endpoint is linear. When predetermined margin for biosimilarity is small, the linear relationship between

biosimilar product and endpoint gets attenuated slightly in slope estimation. However, if the predetermined margin is large, the relationship may not be the same as the relationship between reference product and the endpoint. The proposed unbiased test can detect the distortion of relationship. Unlike small-molecular drug products, the conclusion of therapeutic and interchangeability based on a statement of bioequivalence does not apply to biosimilar drug product. Under a condition that variability in biosimilar product is larger than that in reference product the bioequivalence may not be valid. Detecting the lack of bioequivalence property would be asset in biosimilarity studies.

Simulation Study for Expose-Response(ER) Model in QT study

Junxian Geng (Florida State University), Qianyu Dang (US FDA)

The QT interval, which is measured from the starting of the QRS complex to the end of the T wave, can reflect the duration of ventricular depolarization and subsequent repolarization. Safety concerns arise from significant prolongation of the QT interval. Each pharmaceutical company must conduct as least one thorough QT/QTc (TQT) study when submitting a new drug application, as the ICH E14 clinical guidance for QT assessment was implemented in 2005. Recently, expose-response(ER) model was introduced by clinical pharmacology as an alternative path to quantify QT effects and thereby can replace the TQT study. They pointed out that other than TQT study, which is a resource intensive study designed solely to evaluate the effect on ECG parameter, ER model uses data routinely generated from studies of the clinical development program. They also pointed out that ER model has more power than “by time-point” analysis (TQT) by using the plasma concentration as a continuous covariate. I do some simulation studies to verify the effectiveness of ER model and also make comparison to TQT study based certain criteria.

Estimating dynamic treatment effects from tumor growth studies

Kingshuk Roy Choudhury, Stephen T. Keir, Kathleen Ashcraft, Mary-Keara Boss, Mark W. Dewhirst (Duke)

We present a method for estimating the empirical dynamic treatment effect (DTE) curves from tumor growth delay (TGD) studies. This improves on current common methods of TGD analysis, such as T/C ratio and doubling times, by providing a more detailed treatment effect and overcomes their lack of reproducibility. The methodology doesn't presuppose any prior form for the treatment effect dynamics and is shown to give consistent estimates with missing data. The method is illustrated by application to real data from TGD studies involving three types of therapy: i) radiation therapy ii) combination therapy iii) anti-angiogenic therapy. We show that resulting DTE curves yield new insights into the treatment effect in each case. We discuss how features of the DTE curves should be interpreted and potentially used to improve therapy.

On Analytical Similarity Assessment in Biosimilar Studies

Tongrong Wang (Duke)

For assessment of biosimilarity of biosimilar products, the United States (US) Food and Drug Administration (FDA) proposed a stepwise approach for providing totality-of-the-evidence of similarity between a proposed biosimilar product and a US-licensed (reference) product. The stepwise approach starts with assessment of critical quality attributes that are relevant to clinical outcomes in structural and functional characterization in manufacturing process of the proposed biosimilar product. FDA suggests that these critical quality relevant attributes be identified and classify into three tiers depending their criticality or risking ranking. To assist the sponsors, FDA also suggests some statistical approaches for assessment of analytical similarity for critical quality attributes (CQAs) from different tiers, namely equivalence test for Tier 1, quality range approach for Tier 2, and descriptive raw data and graphical comparison for Tier 3. Analytical similarity assessment for CQAs in Tier 1 is performed based on the equivalence acceptance criterion (EAC) which depends upon the estimate of variability of the reference product. The FDA's recommended approach often underestimates the

variability of the reference product because it does not take the worst possible lots into consideration. In this poster, the statistical properties of the FDA's recommended approach is examined and alternative methods will be proposed in establishing a more accurate and reliable EAC for analytical similarity assessment.

Compare of four drug-bioequivalence criterion assessing drug interchangeability

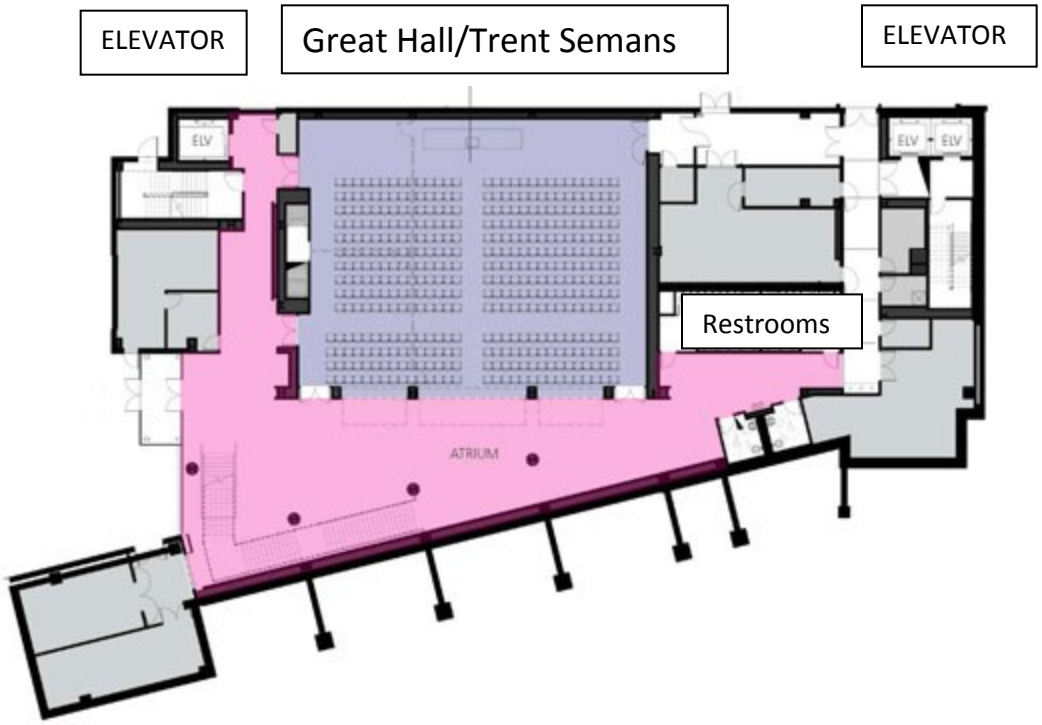
Meng Chen, Shein-Chung Chow (Duke)

As stated by the United States Food and Drug Administration (FDA), an approved generic drug, which is regulated according to average bioequivalence, can be used as a substitute for the innovative drug. However, FDA does not indicate that two generic copies of the same innovative drug can be used interchangeably, even though they are bioequivalent to the same brand-name drug. Along with such undetermined fact, it attracts increasing concerns about whether the approved generic drug has the same therapeutic effect as the brand-name drug. That is concerning about whether they can be used interchangeably. Four criterions are reviewed in this poster for assessment of bioequivalence of generic drug products. The criterion include: average bioequivalence criterion (ABE), a $\sigma_D{}^2$ (the variance due to subject-by-drug interaction) related criterion, and a scaled average bioequivalence (SCDI) criterion. In addition, by extending the idea of reverse of test and reference product, a new criterion for the assessment of interchangeability is proposed.

Floor Plan for Trent Semans Center

Great Hall

Great Hall is located on the ground floor and it has a prominent entrance on Research Drive.



Thursday, Oct. 22, 2015

Great Hall will be divided into two sections. Short Courses 1 & 3 take place in Section A; Short Courses 2 & 4 take place in Section B; Registration/Breakfast & Lunch in Atrium.

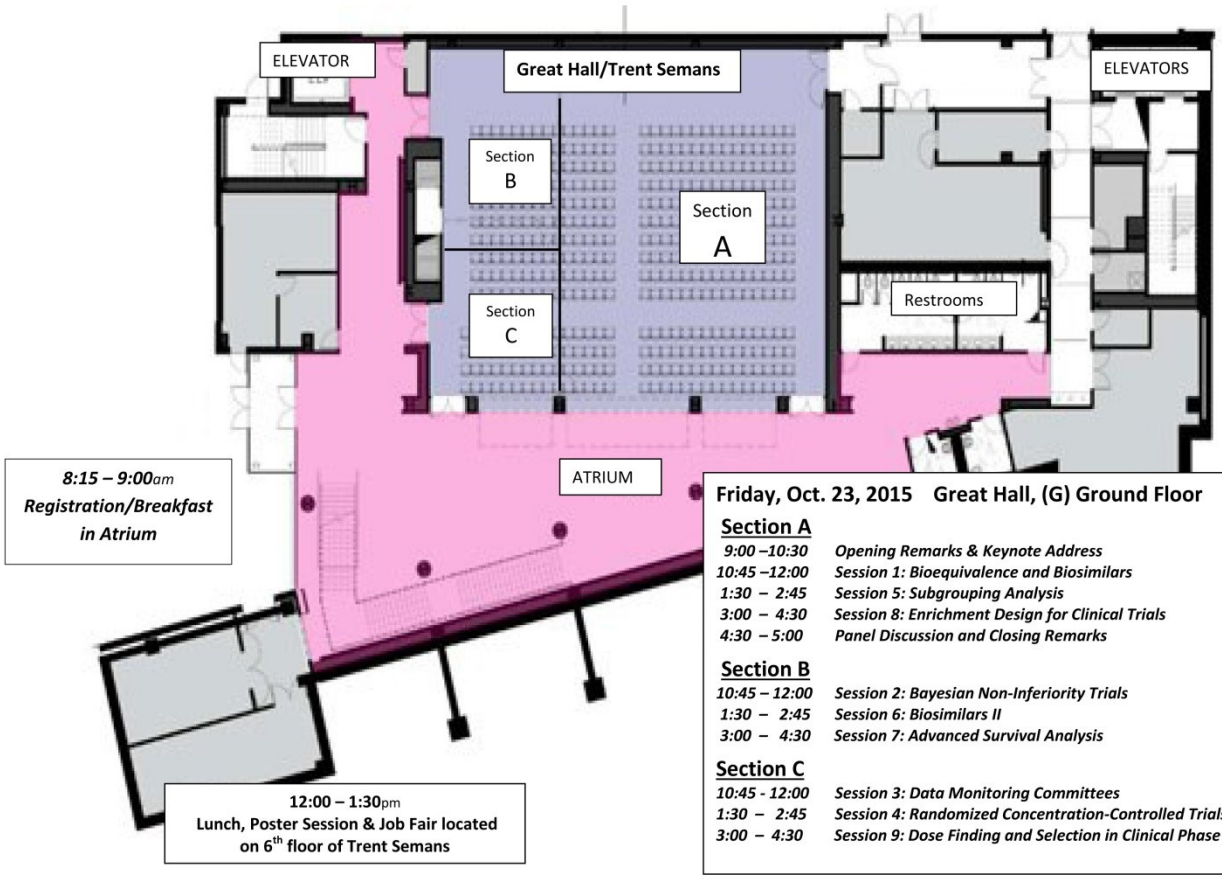
Friday, Oct. 23, 2015

Great Hall will be divided into three sections. Opening, keynote speech and closing sessions will be held in Section A. Parallel sessions will be held in Sections A-C. See Program Schedule for details.

Sixth Floor

This newly completed area has 10,000 square feet of space furnished with chairs and tables. This floor will be used for social mixer on Oct. 22, 2015 and for poster session, job fair and lunch on Oct. 23, 2015.

Friday, Oct. 23, 2015



8:15 – 9:00am
Registration/Breakfast in Atrium

12:00 – 1:30pm
Lunch, Poster Session & Job Fair located on 6th floor of Trent Semans

Friday, Oct. 23, 2015 Great Hall, (G) Ground Floor

Section A

9:00 –10:30	*Opening Remarks & Keynote Address*
10:45 –12:00	*Session 1: Bioequivalence and Biosimilars*
1:30 – 2:45	*Session 5: Subgrouping Analysis*
3:00 – 4:30	*Session 8: Enrichment Design for Clinical Trials*
4:30 – 5:00	*Panel Discussion and Closing Remarks*

Section B

10:45 – 12:00	*Session 2: Bayesian Non-Inferiority Trials*
1:30 – 2:45	*Session 6: Biosimilars II*
3:00 – 4:30	*Session 7: Advanced Survival Analysis*

Section C

10:45 - 12:00	*Session 3: Data Monitoring Committees*
1:30 – 2:45	*Session 4: Randomized Concentration-Controlled Trials*
3:00 – 4:30	*Session 9: Dose Finding and Selection in Clinical Phase*

Directions to Trent Semans Center

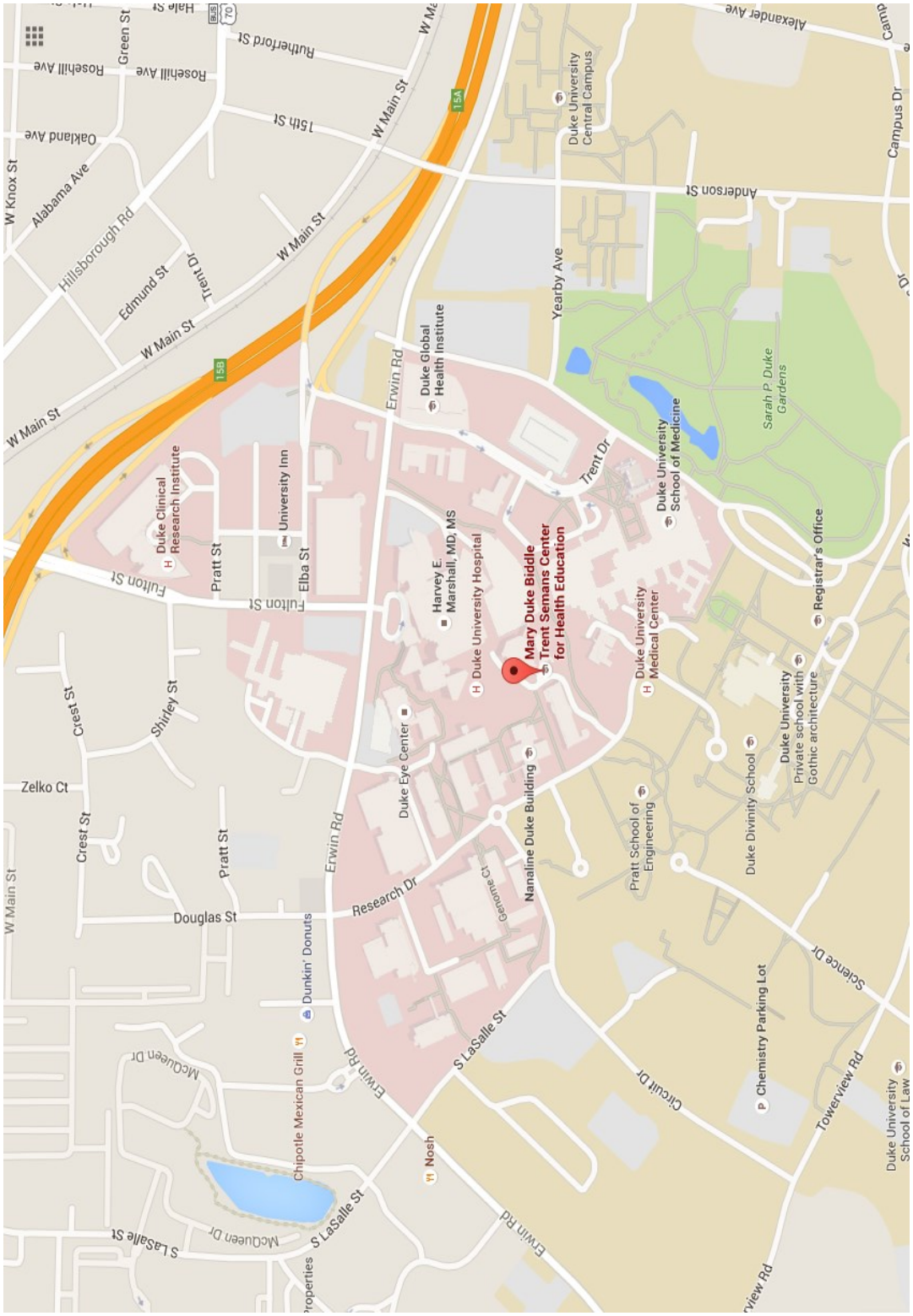
Campus Directions to Trent Semans Center/Great Hall

The symposium will be held in the Great Hall (on level "0") and the 6th floor of the Mary Duke Biddle Trent Semans Center for Health Education located on 8 Searle Center Drive, Durham, NC 27710. The center is behind the Bryan Research building and next to the Searle Center/Medical Center Library & Archives.

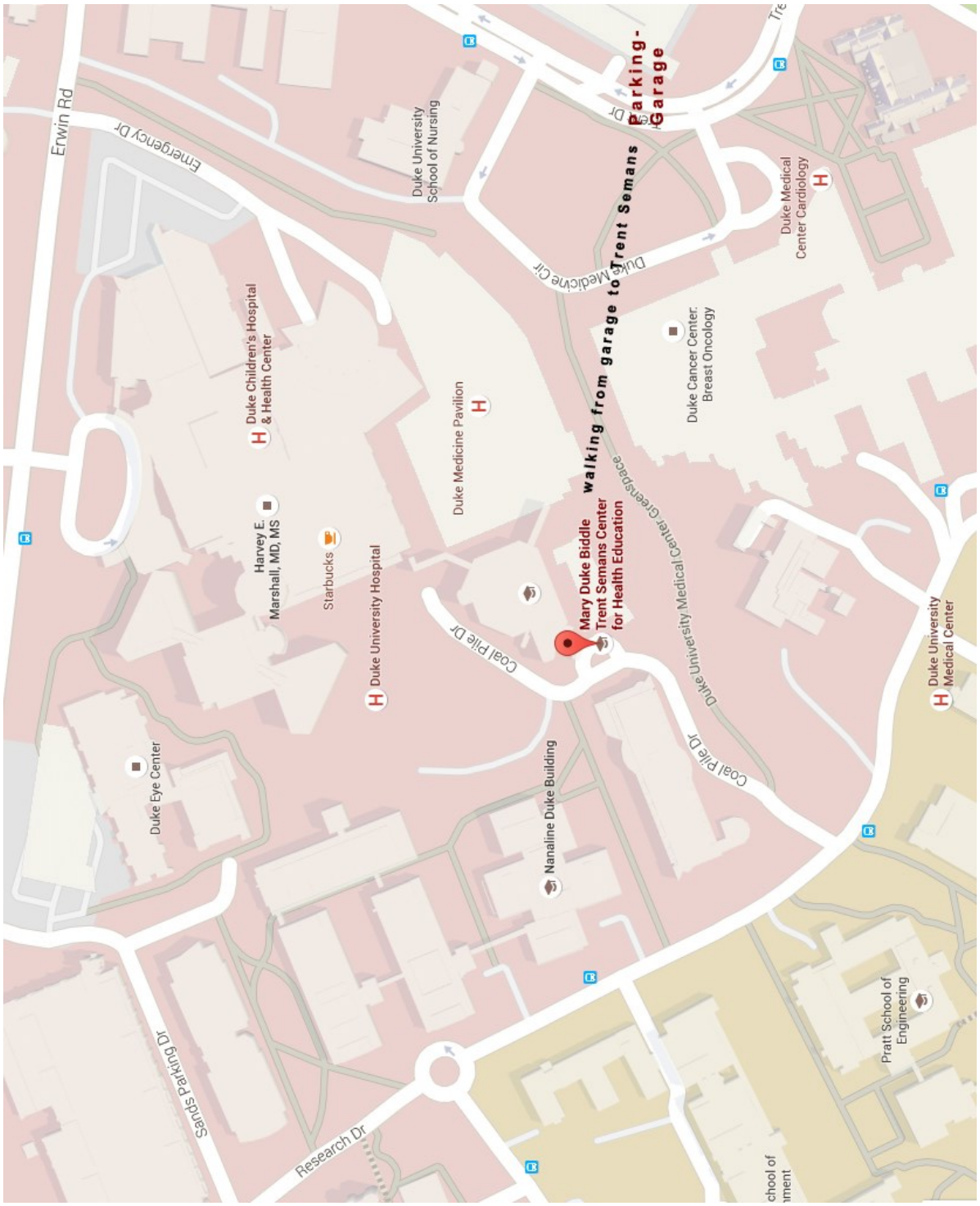
Driving Direction from RDU Airport to Duke Campus

Go 1.2 miles NW on Terminal Blvd. Stay straight onto Airport Blvd. Take the I-40 W for RTP/Chapel Hill. Take the NC-147 N exit, Exit 279B. Take Exit 14, toward Ninth St/Duke Univ. Follow sign to Duke Campus and Trent Drive.

Map and Directions to Trent Semans Center



Map and Directions to Trent Semans Center



Parking in Duke Garage (Trent Drive)

The new Duke PARCS system is operational in this parking garage on Trent Dr. and will require:

1. VERY IMPORTANT - take your parking entry ticket with you (do not leave in car). You will pay for parking before returning to your car.

2. Pick up a validation ticket when you pick up your name tag in Trent Semans

3. Pay/use the validation ticket and the entry ticket at a paystation.

The paystations are located in the level 1 and 2 elevator lobbies of the parking garage.

- Insert your entry ticket in slot
- Pay with validation ticket
- Receive a receipt and you will use this receipt to raise gate and exit.

If you have any difficulties during any of these procedures, press the “call for assistance” button at the paystation.