

Regulatory Science Symposium

Make Informed Decisions: Key Statistical Principles to Clinical Trial Design

Friday, Feb. 4, 2022

9AM - 4PM PDT

Hasting Auditorium HMR 100

We're Back
In Person!



Registration required | No on-site registration

For more information contact: Annie Ly, Project Administrator, lyannie@usc.edu

Agenda

9:00 AM PDT	Introduction <i>Nancy Pire-Smerkanich, DRSc</i> USC, SC-CTSI, School of Pharmacy Assistant Professor, Dept. of Reg. & Quality Sciences
9:30 AM PDT	Basic Statistical Principles: Validity and Sample Size <i>Steve Snapinn, PhD</i> Advarra Managing Expert
10:15 AM PDT	Break
10:30 AM PDT	Medical Device Trials <i>Frances Richmond, PhD</i> USC, SC-CTSI, School of Pharmacy Professor, Dept. of Reg. & Quality Sciences Director, DK Kim International Center for Regulatory Science
11:15 AM PDT	Pediatric Trials <i>Matthew Borzage, PhD</i> USC, Keck School of Medicine & CHLA Assistant Professor of Research Pediatrics
12:15 PM PDT	Lunch
1:00 PM PDT	CTSI Clinical Study Design Types <i>Wendy Mack, PhD</i> USC, Department of Preventive Medicine, Division of Biostatistics Professor USC, SC-CTSI, Biostatistics, Epidemiology, and Research Design (BERD) Core Director
2:15 PM PDT	Break
2:30 PM PDT	Gender, Race and Ethnicity in Clinical Trials <i>Wendy Mack, PhD</i> USC, Department of Preventive Medicine, Division of Biostatistics Professor USC, SC-CTSI, Biostatistics, Epidemiology, and Research Design (BERD) Core Director
3:30 PM PDT	Wrap-Up <i>Eunjoo Pacifici, PharmD, PhD</i> USC, SC-CTSI, School of Pharmacy Chair & Associate Professor, Dept. of Reg. & Quality Sciences Associate Director, DK Kim International Center for Regulatory Science

Regulatory Science Symposium: *Make Informed Decisions: Key Statistical Principles to Clinical Trial Design*

Speaker Bios

Nancy Smerkanich, DRSc, MS, is an Assistant Professor in the Department of Regulatory and Quality Sciences, School of Pharmacy at USC. Dr. Smerkanich holds a Doctorate and master's degree in Regulatory Science from USC and a Bachelor of Science Degree in Microbiology and a Bachelor of Arts in Russian from the University of Connecticut. Dr. Smerkanich received her faculty appointment after successfully completing her Doctoral Dissertation on "Benefits Risk Frameworks – Implementation in Industry" in 2015. In addition to teaching courses related to drug development and clinical trials, she provides regulatory guidance to industry peers. Nancy brings many years of practical regulatory knowledge and experience to academia where she participated in all regulatory aspects of product development, having served as Regulatory Liaison, US Agent, and Global Regulatory Lead across varied therapeutic areas. Known for her dedication to education and mentoring across industry, Nancy continues to be recognized for her ability to provide accurate, relevant and dynamic instruction on both the technical and strategic aspects of global regulatory affairs and for her service to professional organizations such as the Drug Information Association (DIA) and The Organization for Professionals in Regulatory Affairs (TOPRA). piresmer@usc.edu



Steve Snapinn, PhD, is a managing expert at Advarra. Dr. Snapinn holds a PhD in Biostatistics from the University of North Carolina at Chapel Hill, MS in Bioengineering from Columbia University in the City of New York, and BS in Engineering Science from the University of Virginia. He has over 30 years of experience as a biostatistician in the pharmaceutical industry. Formerly, he was a consultant at Seattle-Quilcene Biostatistics LLC, the Senior Vice President of Biometrics at Alder Biopharmaceuticals Inc., the Vice President of Global Biostatistical Science at Amgen for over 14 years, and the Senior Director of Biostatistics at Merck for over 20 years. He is also the former editor of Statistics in Biopharmaceutical Research and is a fellow of the American Statistical Association. He has shared his expertise about the essential role of statistics in the medical product development process with numerous graduate and doctoral students at USC for 4 years as a guest lecturer. snapinns@gmail.com



Frances Richmond, PhD, is a Professor of Regulatory and Quality Sciences (Teaching Track) and Director of the DK Kim International Center for Regulatory Science at USC. She was educated as a neurophysiologist (BNSc, MSc, PhD) at Queen's University, Kingston, Canada and then completed post-doctoral studies at the Université de Montréal and the U.S. National Institutes of Health (NIH). She possesses numerous years of teaching experience and expertise in research and industry, formerly serving as a professor and Associate Dean of Life Sciences at Queen's University, conducting research as a clinical scientist at the Alfred E. Mann Foundation, consulting at Advanced Bionics



Corporation, and more. She was the first woman to be appointed Director of a research consortium, specifically the MRC Center for Sensory-Motor Research, funded by Canada's Medical Research Council (1995-2000). After joining USC in 1999 as a professor, she founded the Department of Regulatory and Quality Sciences, previously holding the role of Chair. Dr. Richmond is or has been a member of five large US research consortia (NIH Engineering Research partnership, NIH Bioengineering Research partnership, Clinical and Translational Science Institute, Tobacco Center of Regulatory Science, Consortium for Technology and Innovation in Pediatrics). Dr. Richmond and her team have been responsible for the development and oversight of multiple undergraduate and graduate programs in the School of Pharmacy that provide certificate, MS and doctoral training in the regulatory and quality management of foods, dietary supplements, medical devices and drugs. fjr@usc.edu

Matthew Borzage, PhD, is an Assistant Professor of Research Pediatrics, Keck School of Medicine of USC and a faculty researcher at the Children's Hospital of Los Angeles with an interest in the area of neurodevelopment and a focus on flow of cerebrospinal fluid and blood in the brain. He holds a MS and PhD in Biomedical Engineering from USC. His primary training as a biomedical engineer enables him to work with the critical technologies of noninvasive data acquisition of cerebrospinal fluid and near infrared spectroscopy with classic analysis techniques for these technologies and to have a critical working knowledge of the nature of pathophysiologies that commonly impact flow in the in brain, including disrupted cerebrospinal fluid flow via Chiari malformations, spina bifida, post intraventricular hemorrhagic hydrocephalus, and more. Combining his research interests and capabilities enables understanding the physiology that underlines some of the most devastating neurological pathologies. He has authored over 25 research publications. His current work focuses on shunt-responsive hydrocephalus and individual cerebral hemodynamic and oxygenation relationships. borzage@usc.edu



Wendy Mack, PhD, is the Director of the Biostatistics, Epidemiology, and Research (BERD) Core at the Southern California Clinical and Translational Science Institute (SC-CTSI) and a Professor in the Department of Preventive Medicine, Division of Biostatistics in the Keck School of Medicine at USC. She received her doctorate in Biometry from USC. She has over 25 years of experience in directing biostatistical and data coordination activities for multiple single-centered and multi-centered clinical trials and observational studies. She has directed the biostatistical and data coordination activities of randomized clinical trials (the majority being NIH- or PCORI-funded), as well as NIH-funded program projects, and has a wealth of experience and expertise in analysis of longitudinal clinical trial outcomes. With over 30 years of teaching USC students, Wendy remains deeply committed to training the next generation of clinical investigators and biostatisticians. As the former director of the MS programs in Biostatistics and Epidemiology in the Department of Preventive Medicine, she has mentored numerous K-awardees, junior faculty, and graduate students (MS and PhD). wmack@usc.edu



Eunjoon Pacifici, PharmD, PhD, is the Chair and Associate Professor of Regulatory and Quality Sciences and Associate Director of the DK Kim International Center for Regulatory Science at USC. Dr. Pacifici received a BS in Biochemistry from the University of California Los Angeles followed by a PharmD and PhD in Toxicology from USC. She conducted her graduate research in the laboratory of Dr. Alex Sevanian in the Institute for Toxicology where she studied the mechanism of oxidative damage and repair in endothelial cell membrane. Before returning to USC as faculty, Dr. Pacifici worked at Amgen and gained experience in conducting clinical research with a special focus on the Asia Pacific and Latin America regions. She initially worked in the clinical development group managing U.S. investigational sites and central laboratories and then went on to work in the Asia Pacific / Latin America group interfacing with local clinical and regulatory staff in Japan, the People's Republic of China, Taiwan, and Mexico. She represented regional clinical and regulatory views on therapeutic product development teams and led satellite task forces in order to align local efforts with U.S. activities. Her additional professional experiences include community pharmacy practice in various settings and clinical pharmacy practice at the Hospital of the Good Samaritan in Los Angeles. Her current focus is on developing the next generation of regulatory scientists and pharmacy professionals with the knowledge, tools, and skills to expedite the development of innovative, safe, and effective biomedical products. epacific@usc.edu



Regulatory Science Symposium

Make Informed Decisions: Key Statistical Principles to Clinical Trial Design

Introduction

Nancy Pire-Smerkanich, DRSc, MS

Associate Professor, Regulatory and Quality Sciences



Lily Jara, BS
Clinical Research
Supervisor,
COVID-19
Biorepository
Project Manager,
CRS

**Contact
Information:**
crs@sc-ctsi.org

SC CTSI Clinical Research Support (CRS)

A single stop for accessing all services an investigator and research team needs to develop, activate, conduct, and report results for human subject research studies

Initial focus on investigator-initiated trials (non-cancer)

- **Services:**
 - Clinical research coordinators for hire
 - Research navigation
 - Recruitment support
 - Budget preparation support
- **Clinical Trials Unit (CTU):**
 - Skilled research and nursing staff
 - Services to support highly-complex human subjects research studies
 - Specimen processing lab
- **Voucher program:**
 - Awards up to \$3,000 to generate new data for development of clinical and/or community research projects

<https://sc-ctsi.org/about/groups/clinical-research-support>

Clinical Trial Quality Training Series

The screenshot shows the USC School of Pharmacy logo and SCCTSI logo at the top. The main heading is 'CLINICAL TRIAL QUALITY Training Series'. Below this, there are two highlighted sections: 'MODULE 2: AUDITING' and 'CATCH UP WITH MODULE 1: MONITORING'. The 'MODULE 2: AUDITING' section includes a sub-heading 'Clinical Trial Quality Training Series Module II: Auditing of a Clinical Research Site' and a description of the module. The 'CATCH UP WITH MODULE 1: MONITORING' section lists resources such as Quizzes, Templates, Checklists, SOPs, and Resources. A 'TO ACCESS THIS FREE RESOURCE' section provides instructions on how to access the training series.

USCSchool of Pharmacy
Department of Regulatory and Quality Sciences

SCCTSI

CLINICAL TRIAL QUALITY Training Series

MODULE 2: AUDITING

Clinical Trial Quality Training Series
Module II: Auditing of a Clinical Research Site

Now available for public use

CLINICAL TRIAL QUALITY TRAINING SERIES

Brought to you by the University of Southern California (USC) Department of Regulatory and Quality Sciences and Southern California Clinical and Translational Science Institute (SC-CTSI), these self-study modules allow you to learn and familiarize yourself with the concepts of monitoring and auditing of clinical research.

TO ACCESS THIS FREE RESOURCE

1. Go to: <http://uscregsci.remote-learner.net>
2. Sign In/Create a new account
 - a. For new accounts, open your email and confirm
3. Select the module and click "Enroll Me"

CATCH UP WITH MODULE 1: MONITORING

- Quizzes
- Templates
- Checklists
- SOPs
- Resources

Clinical Trial Quality Training Series
Module I: Monitoring of a Clinical Trial Site

1. Go to: <https://uscregsci.remote-learner.net>
2. Click **create new account** (right-hand side)
3. Type in your information and click **Create my new account** (bottom of page)
4. Open your email and click the link to confirm your account
5. Click **courses** (middle of page)
6. Scroll down and click the desired module
7. Click **Enroll me** (middle of page)

Georgia CTSA and SC CTSI: Online Course Catalog

- Free trainings for clinical research workforce
- Free, one-time registration to the first 400 registrants
- Registration provides unlimited access to all courses and programs in the Online Course Catalog
- Participants earn a certificate or badge with contact hours upon completion of a course or program
- Contact hours can be used for CRP certification renewal
- To get started:
<https://twd.ce.emorynursingexperience.com/>



Georgia CTSA Translational Workforce Development Announces Online Course Catalog with Free Trainings for Clinical Research Professionals

The Georgia Clinical and Translational Science Alliance (Georgia CTSA) and the University of Southern California Clinical and Translational Science Institute (SC CTSI) are collaborating on an exciting new educational venture geared toward clinical research professionals at every stage of their professional development. Through this partnership, Georgia CTSA has created a new Online Course Catalog with [free course and program offerings](#) available to clinical research professionals and principal investigators. These courses and programs are created and vetted by experts in cross-disciplinary fields such as instructional design, technology, workforce development, regulatory science, clinical and translational science, and operations.

"We are fortunate to partner with USC SC CTSI to bring such a broad offering of high-quality trainings to our clinical research professionals."

Linda McCauley, RN, PhD, Program Director of the Georgia CTSA Translational Workforce Development and Dean of the Nell Hodgson Woodruff School of Nursing at Emory University

"This joint effort between Georgia CTSA and SC CTSI will create a wonderful resource to support training and career development of clinical research professionals at all levels. It will be a game changer, especially for people working in an academic setting."

Thomas Buchanan, MD, Director & Principal Investigator of the SC Clinical and Translational Science Institute

"It has been a pleasure to partner with Georgia CTSA team in our common goal to promote life-long learning for the clinical research workforce."

Eunjoo Pacifici, PharmD, PhD, Chair and Associate Professor in the Department of Regulatory and Quality Sciences and Associate Director of the DK Kim International Center for Regulatory Science at the USC School of Pharmacy

Participants earn a certificate or badge with contact hours (continuing education) from an accredited provider upon completion of a course or a program (series of courses). Contact hours can be used to meet requirements for CRP certification renewal.

Free, one-time registration to the Georgia CTSA Online Course Catalog is available to the first 400 registrants. Registration provides unlimited access to all courses and programs in the Georgia CTSA Online Course Catalog. View the [Online Course Catalog](#) to get started.

The first program, *Legal Aspects for Conducting Clinical Trials*, is comprised of six courses. Individual courses in all programs receive a certificate, and completing the program earns a badge. The second program, *Clinical Trials with Medical Devices*, is comprised of seven courses of which completion of five of the seven courses will earn a badge. Be sure to check out the dashboard features as you build your professional career.

Stay Tuned for More Courses and Programs as We Develop This Free Online Course Catalog!

Find us on our website: <https://regulatory.usc.edu/>



Department of Regulatory & Quality Sciences

Advancing the Profession

One of the first programs in this dynamic arena, the Department of Regulatory and Quality Sciences remains a global leader in producing professionals with the knowledge and skills to manage regulated biomedical products worldwide. This rapidly growing and increasingly global field encompasses every aspect of pharmaceutical and medical device development, quality assurance and clinical trials oversight—helping shepherd life-improving and often lifesaving advances to the marketplace.

Degree Programs

Five Graduate Streams

- DRSC
- **MS Regulatory Science**
- MS Regulatory Management
- MS Management of Drug Development
- MS Medical Product Quality

Certificates

- Food safety
- Regulatory Science
- Early Drug Development
- **Clinical Design and Management**
- Patient and Product Safety



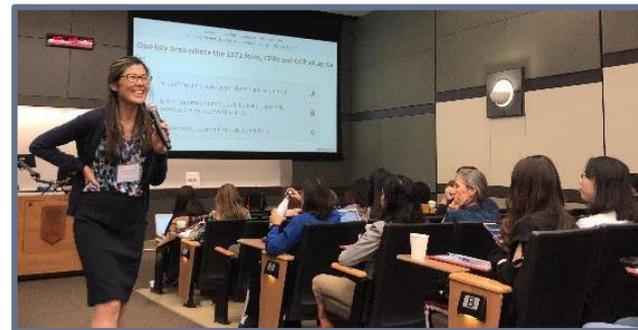
Nancy Smerkanich
DRSc, MS

Assistant Professor
Department of Regulatory
and Quality Sciences

piresmer@usc.edu

Symposiums

- **2015** - Clinical Trial Hurdles
- **2016 Spring** - Clinical Trial Startup
- **2016 Fall** - Monitoring and Auditing
- **2017 Spring** - Clinical Trials in Special Populations
- **2017 Fall** - Clinical Trials in Era of Emerging Technologies and Treatments
- **2018 Spring** - Regulatory Aspects of Clinical Trial Design
- **2018 Fall** - Pharmacovigilance and Safety Reporting
- **2019 Spring** - Patient-Centered Drug Development and Real World Evidence/Data
- **2019 Summer** - Clinical Trials with Medical Devices
- **2019 Fall** - Legal Aspects of Conducting Clinical Trials
- **2020 Spring** - Quality by Design in Clinical Trials
- **2020 Fall** – Diversity in Clinical Trials in the Time of COVID-19
- **2021 Spring** – Clinical Research Career Pathways (half-day)
- **2021 Spring** – Principles of Global Clinical Research for Medical Devices
- **2021 Fall** – Innovation to Translation: Role of Genomics in Medical Product Development
- **2022 Spring** – Make Informed Decisions: Key Statistical Principles to Clinical Trial Design
- **2022 Fall** – TBD



Symposium recordings are easily accessible for viewing on the SC CTSI's online educational library <https://sc-ctsi.org/training-education/courses?audience=researchProfessionals>

Regulatory Science Symposium

Make Informed Decisions: Key Statistical Principles to Clinical Trial Design

Friday, Feb. 4, 2022
9AM - 4PM PDT

Online
via Zoom



Agenda

9:00 AM PDT	Introduction <i>Nancy Pire-Smerkanich, DRSc</i> USC, SC-CTSI, School of Pharmacy Assistant Professor, Dept. of Reg. & Quality Sciences
9:30 AM PDT	Basic Statistical Principles: Validity and Sample Size <i>Steve Snapinn, PhD</i> Advarra Managing Expert
10:15 AM PDT	Break
10:30 AM PDT	Medical Device Trials <i>Frances Richmond, PhD</i> USC, SC-CTSI, School of Pharmacy Professor, Dept. of Reg. & Quality Sciences Director, DK Kim International Center for Regulatory Science
11:15 AM PDT	Pediatric Trials <i>Matthew Borzage, PhD</i> USC, Keck School of Medicine & CHLA Assistant Professor of Research Pediatrics
12:15 PM PDT	Lunch
1:00 PM PDT	CTSI Clinical Study Design Types <i>Wendy Mack, PhD</i> USC, Department of Preventive Medicine, Division of Biostatistics Professor USC, SC-CTSI, Biostatistics, Epidemiology, and Research Design (BERD) Core Director
2:15 PM PDT	Break
2:30 PM PDT	Gender, Race and Ethnicity in Clinical Trials <i>Wendy Mack, PhD</i> USC, Department of Preventive Medicine, Division of Biostatistics Professor USC, SC-CTSI, Biostatistics, Epidemiology, and Research Design (BERD) Core Director
3:30 PM PDT	Wrap-Up <i>Eunjoo Pacifici, PharmD, PhD</i> USC, SC-CTSI, School of Pharmacy Chair & Associate Professor, Dept. of Reg. & Quality Sciences Associate Director, DK Kim International Center for Regulatory Science

Presented by the USC School of Pharmacy International Center
for Regulatory Science and the Southern California Clinical and
Translational Science Institute

This certifies that

Before the end of today's symposium, you will
receive a link to take the program evaluation.

Follow this link to the Survey:
[Take the Survey](#)

Please complete the program evaluation to receive a
certificate of completion by Friday, February 18, 2022.



Eunjoon Pacifici, PharmD, PhD
Director
International Center for Regulatory Science



Thomas A. Buchanan, MD
Director
Southern California Clinical and
Translational Science Institute

USC School of Pharmacy
DK Kim International Center for Regulatory Science



Thank You!



www.sc-ctsi.org

Phone: (323) 442-4032

Email: info@sc-ctsi.org

Twitter: @SoCalCTSI

USC School of Pharmacy

DK Kim International Center for Regulatory Science

regulatory.usc.edu

Phone: (323) 442-3521

Email: regsci@usc.edu

Facebook: @RegSci

Basic Statistical Principles: Validity and Sample Size

Steve Snapinn

Outline

- Statistical Inference, Hypothesis Testing, Type 1 Error, Power and Sample Size
- Randomization and Blinding
- Multiplicity in Clinical Trials
- Types of Endpoint

Hypothesis testing

- Statistical Hypothesis Testing Involves Defining a Null Hypothesis and Determining Whether the Data Support or Refute It
 - H_0 : Mean Difference Between Groups = 0
 - P-Value: Probability the Observed Data (Or Better) Could Have Occurred *By Chance Alone* (ie, If the Null Hypothesis Were True)
 - A Small P-Value Suggests That the Null Hypothesis Is False
 - If the P-Value Is Small Enough (Typical < 0.05) We *Reject* the Null Hypothesis and Conclude the Drug Has an Effect

Statistical significance and clinical significance

- “Statistical Significance” = Rejection of the Null Hypothesis
 - Simply Means That the Treatment Has an Effect > 0
 - Smaller P-Value Means Greater Evidence That Effect > 0
 - Does *Not* Mean Bigger Treatment Effect Size
- “Clinical Significance” Refers to a Sufficiently Large Treatment Effect

Power and Sample size

- Power Is the Probability That a Study Will Detect (ie, Achieve a Statistically Significant Result) a True Treatment Effect of a Pre-Specified Magnitude
 - Typically Want 80% to 90% Power to Detect a Small But Clinically Meaningful Effect
- Sample Size Is Chosen to Provide the Desired Power

Randomization and Blinding

- ***Bias*** is a Systematic Error
 - Potential Source: Allocation of Sicker Patients to a Specific Treatment Arm
 - Avoidance Technique: Randomization
 - Potential Source: Endpoint Assessment Based on Preconceived Beliefs
 - Avoidance Technique: Blinding
- ***Random Error*** is Unpredictable
 - Primary Source: Limited Sample Size

Estimates and confidence intervals

- Clinical Trials Are Done in a Sample of Patients From a Population
- The Results of the Study Allow Us to Draw Inference About the Population From the Sample
- The *Point Estimate* Is the Single Best Estimate of the Average Treatment Effect in the Population
- The *Confidence Interval* Is a Range of Likely Values for the Average Treatment Effect in the Population

Multiplicity

- Clinical Trials Often Include Multiple Hypothesis Tests
- If There Are Multiple Chances to “Win,” the Probability of Type 1 Error Increases
 - Type 1 Error = Statistically Significant Difference When the Treatment Is Ineffective
- Statistical Approaches Can Control the Overall (Experiment-Wise) Probability of a Type 1 Error

PROMISE-1 Key Efficacy Results

P-values for Primary and Key Secondary Endpoints

Endpoint	ALD403 300 mg		ALD403 100 mg		ALD403 30 mg	
	p-value	Decision	p-value	Decision	p-value	Decision
Primary	0.0001	S	0.0182	S	0.0046 ^a	NS
Key secondary: 75% responder rate Weeks 1-4	0.0066	S	0.0112	S	0.0170 ^a	NS
Key secondary: 75% responder rate Weeks 1-12	0.0007	S	0.1126	*	0.0272 ^a	NS
Key secondary: 50% responder rate Weeks 1-12	0.0001	S	0.0085 ^a	NS	0.0064 ^a	NS
Key secondary: day after dosing	0.0159 ^a	NS	0.0312	NS	0.1539 ^a	*

Abbreviations: NS = nominally significant; S = significant; * = not significant.

All endpoints were tested with an alpha of 5% as per the serial multiple testing procedure.

^a Unadjusted p-values were presented in the treatment groups for primary and key-secondary endpoints in accordance with Section 9.7.1.2.

Sources of Multiplicity

- The More Statistical Tests Performed, the More Type 1 Errors Are Likely to Occur, and the More Likely It Is That There Will Be At Least One Type 1 Error
- Clinical Trials Often Have Many Statistical Tests, and Therefore Many Opportunities to Make a Type 1 Error
- Sources of Multiple Tests Include
 - Multiple Endpoints
 - Multiple Dose Groups
 - Multiple Timepoints
- PROMISE-1 Had All of These
 - Mean Monthly Migraine Days, 50% Responders, 75% Responders, etc.
 - 300 mg vs PBO, 100 mg vs PBO, 30 mg vs PBO
 - Day 1, Month 1, Months 1-3

Controlling the Experiment-Wise Error Rate

- Requiring $P < 0.05$ Ensures That the Type 1 Error Rate for a Given Test Is $< 5\%$
- However, the Goal in Trials for Regulatory Approval Is to Control the ***Experiment-Wise Error Rate***
- Specifically, We Want No More Than a 5% Chance of Making One or More Type 1 Errors Among ***All*** of the Multiple Statistical Tests
- This Requires a Different Statistical Approach Than Simply Using $P < 0.05$ for All Tests

Main Approaches for Adjusting for Multiplicity

- Bonferroni Method
 - Divide the Alpha Equally Across All Hypotheses
 - eg, If There Are Three Hypotheses, Test Each at $P < 0.0167$
- Improved Bonferroni Methods
 - Holm Procedure: Multi-step step-down procedure: Start with smallest p-value, and continue testing until the first non-significant result (eg, $p_1 = 0.0167$, $p_2 = 0.025$, $p_3 = 0.05$)
 - Hochberg Procedure: Multi-step step-up procedure: Start with largest p-value, and continue testing until the first significant result (eg, $p_1 = 0.05$, $p_2 = 0.025$, $p_3 = 0.0167$)
 - Unequal Alpha Allocation
- Fixed-Sequence Method
 - Test each hypothesis at $p < 0.05$ in an prespecified sequence; stop at the first non-significant result
- Gate-Keeping Approaches, Fallback Approaches, Combination Approaches, etc

Regulatory Claims

- Draft Guidance from FDA in January 2017: Multiple Endpoints in Clinical Trials
- Requires Primary Endpoint Family Which Must Be Significant Before Proceeding to Secondary Endpoint Family
 - “Positive results on the secondary endpoints can be interpreted *only* if there is first a demonstration of a treatment effect on the primary endpoint family.”
- Requires Strong Control of the Type 1 Error Rate to Demonstrate Additional Effects
 - “The Type I error rate should be controlled for the entire trial, defined in section II.C as strong control.”

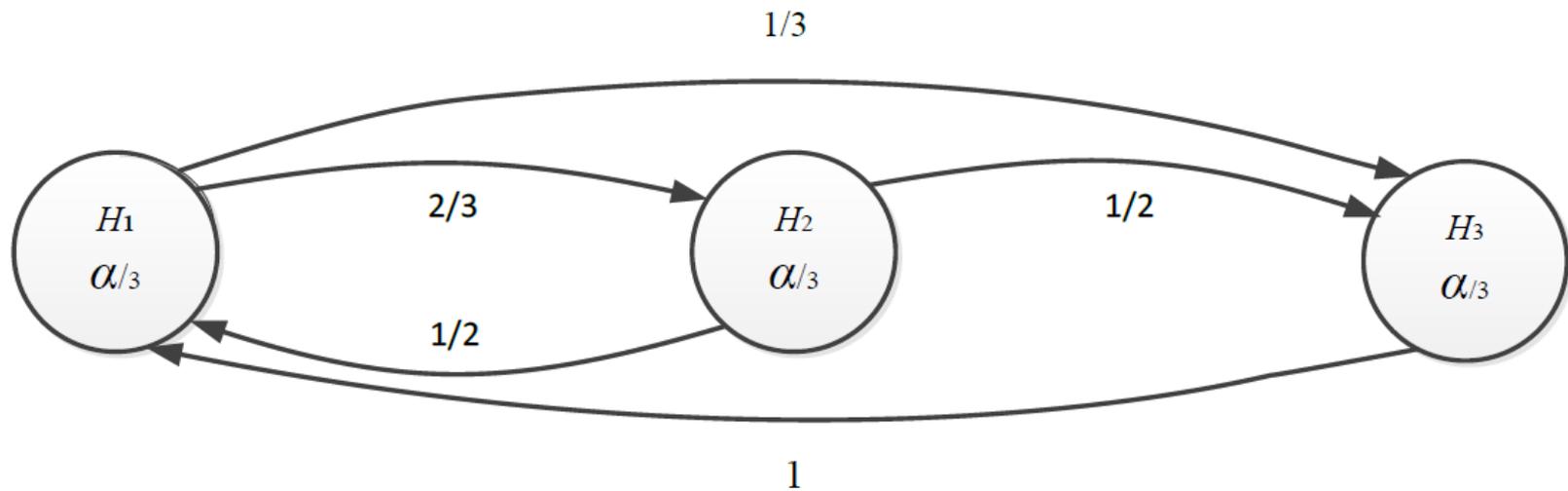
Adjusted P-Values

- Ordinary P-Values Can Only Be Interpreted in the Context of the Multiple Comparisons Procedure
- Adjusted P-Values Solve This Problem
- Adjustment Accounts for the Multiple Comparisons Procedure Such That All P-Values Have the Same Interpretation and $P < 0.05$ Is Always Significant

Multiple Endpoints vs Co-Primary Endpoints

- Ordinary Multiplicity Problem Is When the Drug Is Declared Successful if **Any** Null Hypothesis Is Rejected
- The Opposite Situation Is When the Drug Is Declared Successful Only If **All** Null Hypotheses Are Rejected
- Co-Primary Endpoints Do Not Require Multiplicity Adjustment
- Example: Primary Endpoints for Acute Migraine
 - Pain Freedom at 2 Hours
 - Freedom from Most Bothersome Symptom at 2 Hours

Graphical Approach for Adjusting for Multiplicity



Types of Endpoint

- Continuous
 - e.g., blood pressure (mmHg)
 - Can be post-treatment measurement or change from baseline
- Ordinal
 - e.g., (mild, moderate, severe)
- Dichotomous or Binary
 - Presence or absence of a condition
- Survival or Time-To-Event
- Composite Endpoint
 - Combination of continuous variables
 - Time to first of various events

Continuous Variables

- Example: Change in DBP from Baseline to End of Study
- Treatment Effect Typically Measured As Mean Difference Between Groups
- Simple Analysis Approach: T-Test
- When Distributions Are Not Normal:
 - Consider normalizing transformation
 - Use nonparametric methods
 - Use of median as measure of treatment effect

Dichotomous/Binary Variables

- Example: Objective Response (CR or PR) in Oncology Study
- Simple Analysis Approach: Chi-Square Test
- Various Approaches to Measure Treatment Effect Size
 - Risk difference (NNT)
 - Relative risk
 - Odds ratio

Ordinal Variables

- Example: Patient Assessment of Treatment Benefit
 - Excellent, Good, Fair, Poor
- Calculation of Score
 - E.g., Excellent=4, Good=3, Fair=2, Poor=1
 - Analyze like a continuous variable
- Model-Based Analysis
 - E.g., Proportional Odds Model
 - Treatment effect measured as odds ratio

RENAAL* Results

- Comparison of Losartan and Placebo in Patients with Diabetic Nephropathy
- Primary Endpoint is Occurrence of Doubling of SCr, ESRD or Death
- Rate of Primary Composite Endpoint
 - Losartan: $327/751 = 43.5\%$
 - Placebo: $359/762 = 47.1\%$
- The Difference Was Statistically Significant ($p = 0.02$)
- But, How Big Is It?

*Brenner NEJM 345:861-869; 2001

Available at: <http://www.nejm.org/doi/pdf/10.1056/NEJMoa011161>

Measures of Effect Size for Binary Variables

- Risk Difference
 - $47.1\% - 43.5\% = 3.6\%$
- Number Needed to Treat (NNT)
 - $1 / 3.6\% = 28$ Patients
- Relative Risk
 - $43.5\% / 47.1\% = 0.924 = 7.6\%$ reduction
- Odds Ratio
 - $(43.5\%/56.5\%) / (47.1\%/52.9\%) = 0.865 = 13.5\%$ reduction

Which Measure is the Best?

- Absolute Difference and NNT May Be Best from Public Health or Health Economic Perspective
- Relative Risk and Odds Ratio Are Less Sensitive to Risk of Study Population
- Odds Ratio is Complicated but Mathematically Appealing
- If Rates are Low, Relative Risk and Odds Ratio are Similar

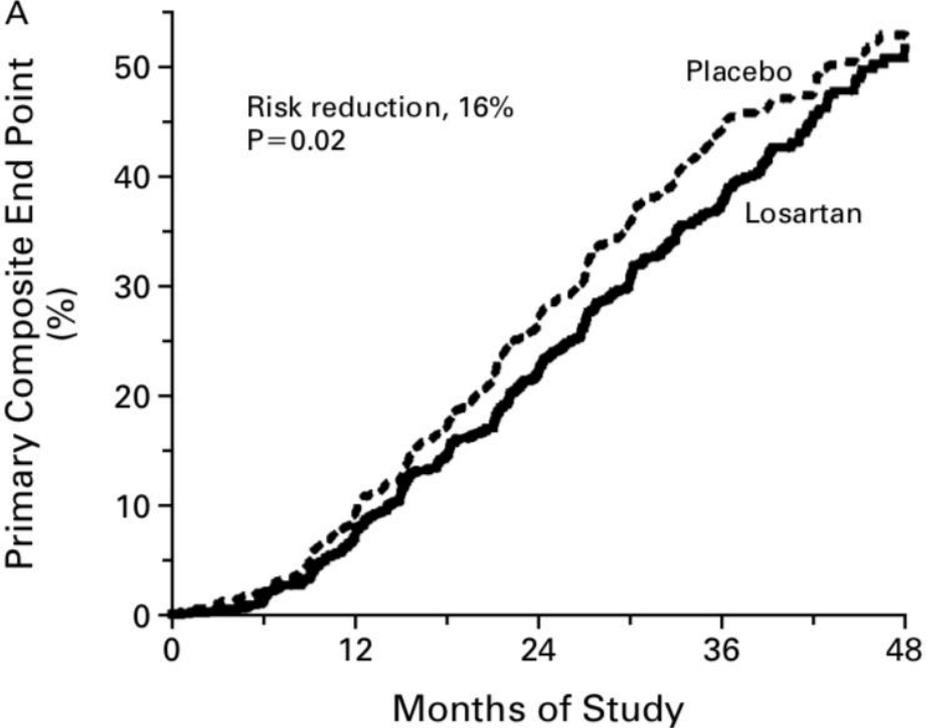
Survival/Time-To-Event Variables

- For Studies Where Endpoints Are Assessed Continuously or Periodically Throughout the Trial
- Example: Survival Time in an Oncology Trial
- Treatment Effect Size Typically Measured By Hazard Ratio

Survival Analysis Accounts for Time-to-Event

- Cox Regression is the Most Common Type of Survival Analysis
- Measure of Effect is the Hazard Ratio
 - Sometimes called relative risk – Can lead to confusion
- Accounts for Censoring (Incomplete Follow-Up)
- RENAAL Result
 - Hazard Ratio = 0.84, or a 16% Reduction

Kaplan-Meier Curves



Kaplan–Meier Curves of the Percentage of Patients with the Primary Composite End Point (Panel A)

NO. AT RISK					
Placebo	762	689	554	295	36
Losartan	751	692	583	329	52

Figure 1A: Brenner NEJM 345:861-869; 2001
<http://www.nejm.org/doi/pdf/10.1056/NEJMoa011161>

Responder Analysis

- Motivation: Slight Change Can Be Significant
 - As Sample Size Increases, Trivial Difference (e.g., 0.1 mmHg) Can Be Statistically Significant
- Want To Ensure Meaningful Drug Effect
- Classify Subjects Into “Responders” and “Non-Responders” Based on Magnitude of Response
- Examples
 - Achievement of ≤ 80 mmHg in DBP in Hypertensive Patients
 - ACR_{20} in Rheumatoid Arthritis
 - Platelet Count $> 50,000$ in ITP
 - Hemoglobin Level > 11 g/dL in Anemia

Composite Endpoints

- Continuous Variables: Calculate an Overall Score
- Clinical Composite: First Occurrence of Any of a Set of Events
 - Major Adverse Cardiovascular Event
 - Myocardial Infarction, Stroke, Cardiovascular Death

RENAAL Study Results

TABLE 3. INCIDENCE OF THE PRIMARY COMPOSITE END POINT AND ITS COMPONENTS.*

END POINT	LOSARTAN GROUP (N=751)		PLACEBO GROUP (N=762)		P VALUE	RISK REDUCTION % (95% CI)
	no. (%)	no./100 patient-yr	no. (%)	no./100 patient-yr		
Primary composite end point†	327 (43.5)	15.9	359 (47.1)	18.1	0.02	16 (2 to 28)
Doubling of serum creatinine concentration	162 (21.6)	7.9	198 (26.0)	10.0	0.006	25 (8 to 39)
End-stage renal disease	147 (19.6)	6.8	194 (25.5)	9.1	0.002	28 (11 to 42)
Death	158 (21.0)	6.8	155 (20.3)	6.6	0.88	-2 (-27 to 19)
End-stage renal disease or death	255 (34.0)	11.7	300 (39.4)	14.1	0.01	20 (5 to 32)
Doubling of serum creatinine concentration and end-stage renal disease	226 (30.1)	11.0	263 (34.5)	13.2	0.01	21 (5 to 34)

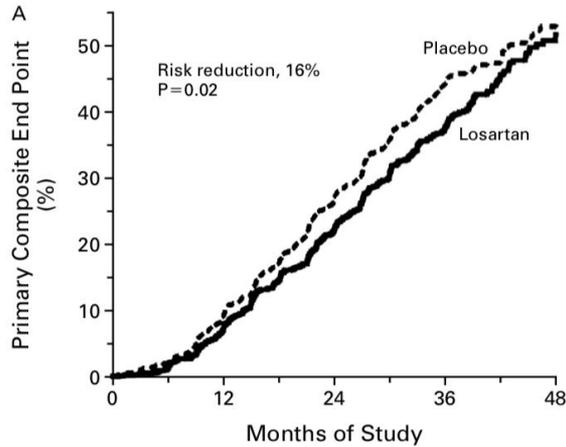
*In end-point trials, there is often a difference between the risk reduction as determined on the basis of the Cox regression model and the risk reduction as determined on the basis of the crude rates of events. The difference results in part from the fact that the Cox regression model accounts for the time at risk — i.e., the longer average follow-up in the losartan group than in the placebo group. To address this aspect of the difference, we present the numbers of events per 100 patient-years of follow-up. In addition, the Cox model accounts for the base-line level of proteinuria (which was a stratification factor) and the geographic region, as prespecified in the data analysis plan. CI denotes confidence interval.

†The primary end point was a composite of a doubling of the serum creatinine concentration, end-stage renal disease, or death.

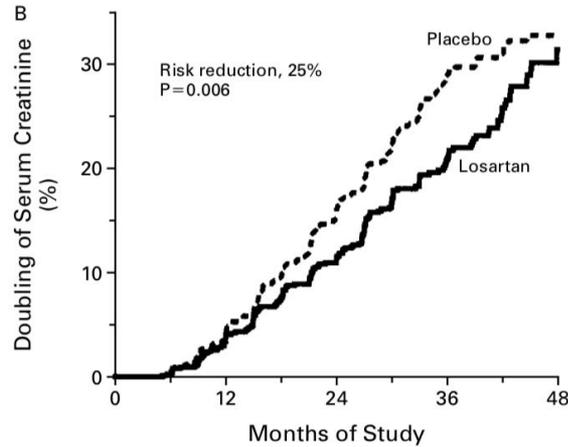
Table 3: Brenner NEJM 345:861-869; 2001

Available at: <http://www.nejm.org/doi/pdf/10.1056/NEJMoa011161>

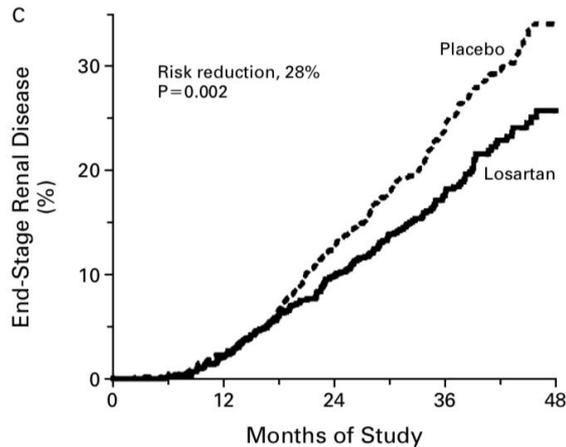
RENAAL Study Results



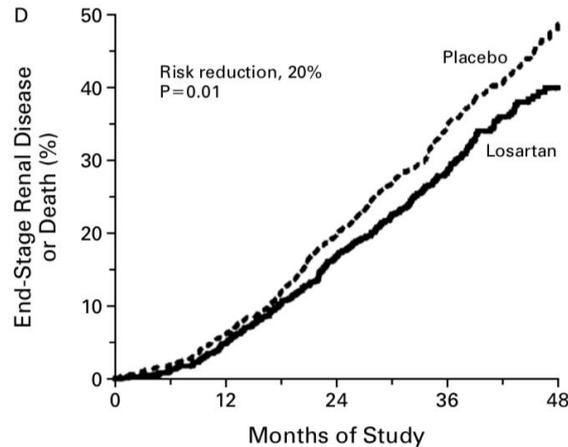
NO. AT RISK					
Placebo	762	689	554	295	36
Losartan	751	692	583	329	52



762	689	554	295	36
751	692	583	329	52



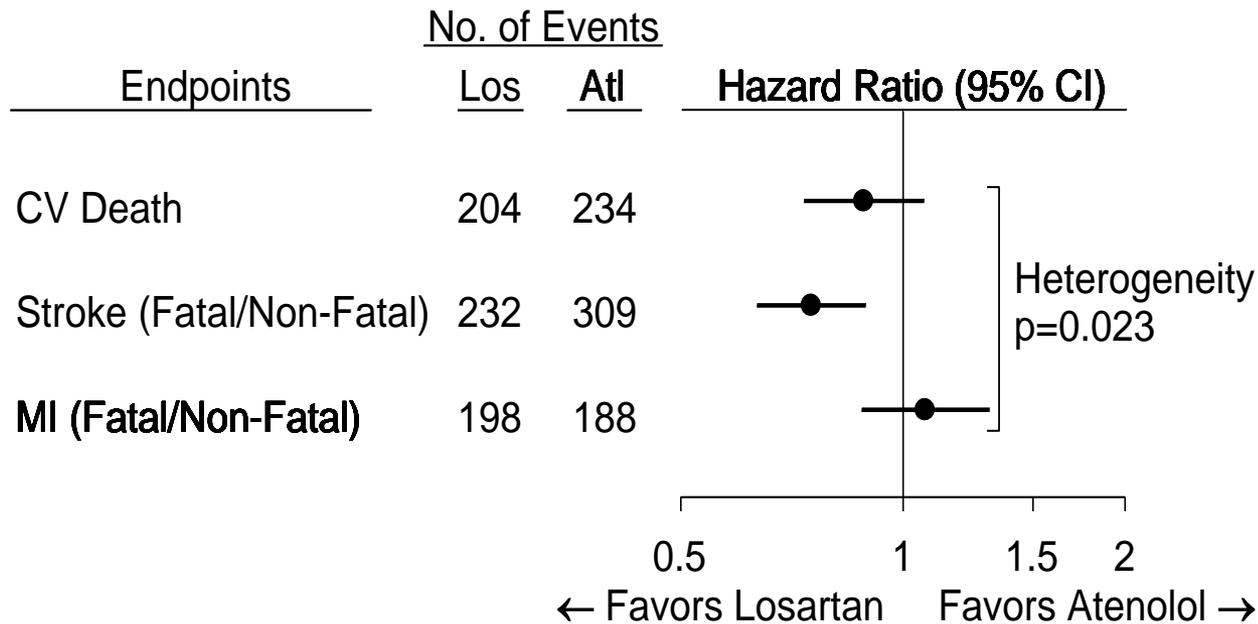
NO. AT RISK					
Placebo	762	715	610	347	42
Losartan	751	714	625	375	69



762	715	610	347	42
751	714	625	375	69

Kaplan–Meier Curves of the Percentage of Patients with the Primary Composite End Point (A) and Its Individual Components, a Doubling of the Serum Creatinine Concentration (B), End-Stage Renal Disease (C), and the Combined End Point of End-Stage Renal Disease or Death (D). The mean follow-up time was 3.4 years (42 months).

LIFE: Secondary Component Endpoints



Clinical and Surrogate Endpoints

- A Clinical Endpoint Is Defined as a Measure of How a Patient *Feels, Functions, or Survives*
- A Surrogate Endpoint Is Believed to Predict Clinical Benefit And Is Often Used in Place of a Clinical Endpoint
 - Examples: Serum Cholesterol, Blood Pressure, HbA1C
- Advantage of a Surrogate Endpoint: Can Be Measured More Quickly and More Easily Than a Clinical Endpoint
 - Reduces Study Sample Size and Duration
- Many Examples of Surrogate Endpoints That Failed to Predict Clinical Benefit
 - CAST Trial in Patients With Ventricular Arrhythmias
- Surrogate Endpoints Should Be *Validated* Prior to Use in Clinical Research

Missing Data and Bias

- In Virtually All Clinical Trials, Some Endpoint Measurements Are Missing
 - Most Common Reason Is Early Discontinuation From the Study
- Missing Data Can Introduce Bias
 - Eg, Patients Who Don't Benefit Withdraw From the Study
- Statistical Analysis Techniques Can Only Partially Adjust for the Bias

Some Opinions in the Literature on Handling Missing Data

- “... by far the best course is to avoid the problem to the extent possible.”
 - O’Neill & Temple, *Clinical Pharmacology & Therapeutics*, 2012
- “... the best method to handle non-ignorable data is to prevent it.”
 - Hardy et al, *J Am Geriatr Soc.*, 2009
- “The best way to deal with missing data is to avoid it.”
 - Sainani, *American Academy of Physical Medicine and Rehabilitation*, 2010
- “... the preferred and often only satisfactory approach to addressing missing data is to prevent it.”
 - Fleming, *Annals of Internal Medicine*, 2011
- “Of course, the best way to handle missing data is to avoid it ...”
 - Siddique et al, *Psychiatr Ann*, 2008

Handling Missing Data

- Types of Missingness
 - Missing Completely at Random
 - Missing at Random
 - Non-Ignorable Missingness
- Techniques for Handling Missing Data
 - Imputation
 - Last Observation Carried Forward
 - Multiple Imputation
 - Statistical Modeling
 - Tipping Point Analysis

Thank You!



Designing Device Trials

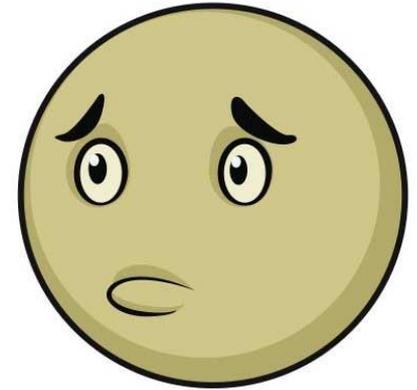
When do we do device trials?

Only when the information is **necessary** to prove safety/ efficacy in a way that

- performance testing
- animal testing
- previous literature

....cannot!

Typically device trials are done on implants



But I worry that....

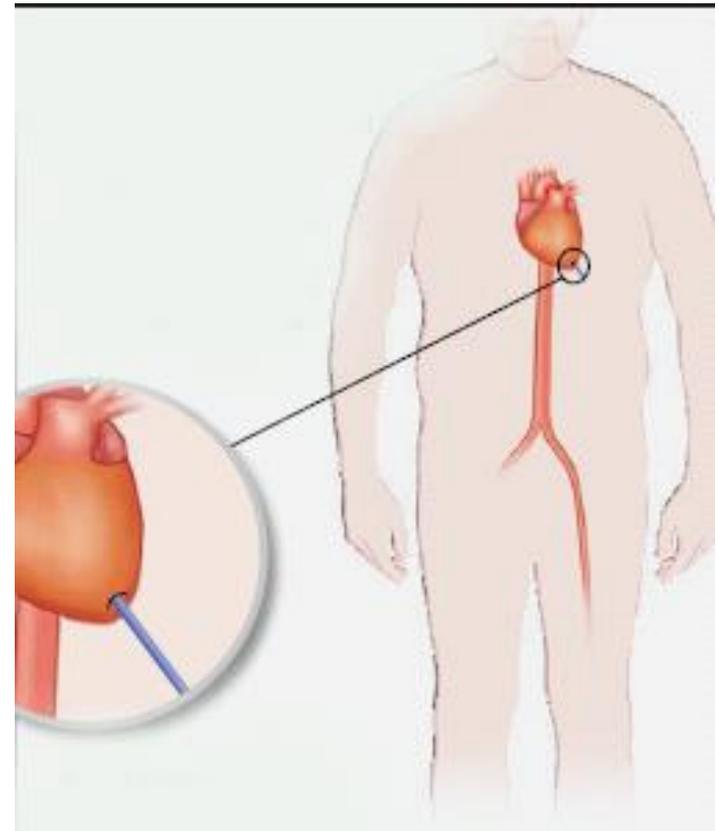
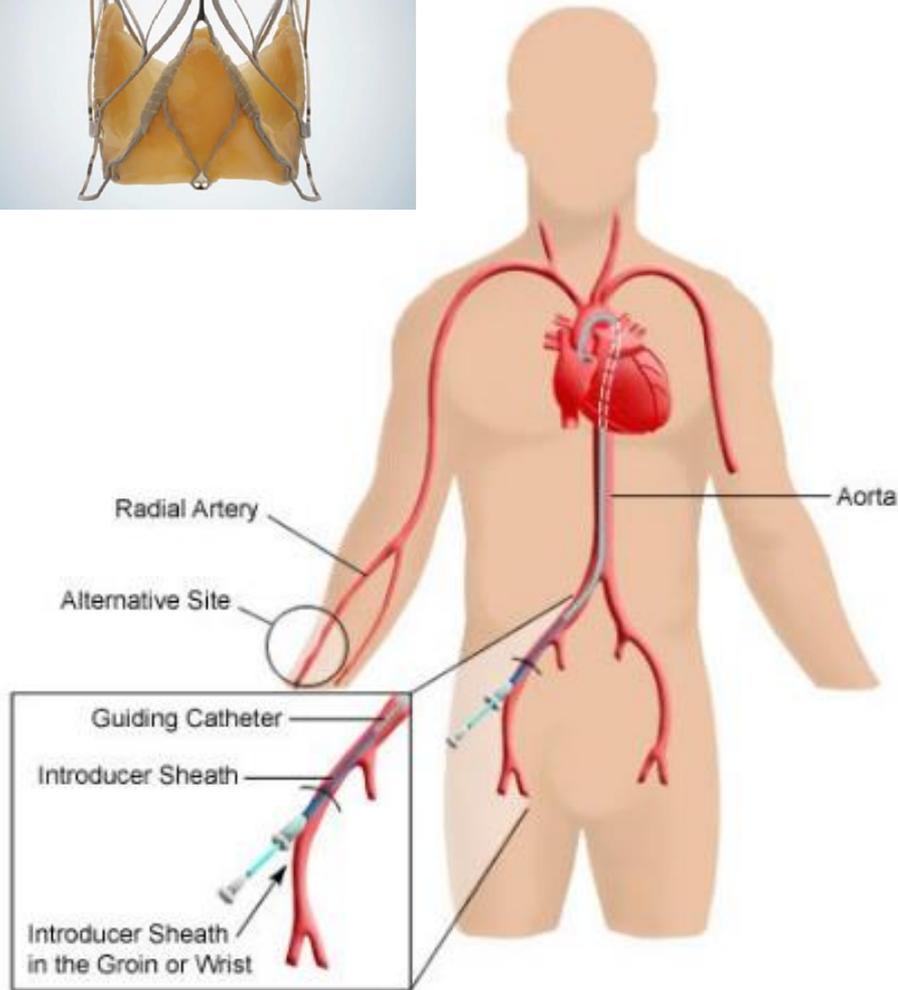
- device cost will limit numbers
- cannot do sham or control operations
- cannot easily implement active control arms
- device size may be problematic
- surgery itself carries risk
- Post-market trials become important for reimbursement, physicians
- Clinical utility is key

Strong partnerships are needed

- Collaboration needed to design trial
- Delivery and programming of device may require participation of company staff
- Multicenter trials will have steering committees for outcomes and publication



Transcatheter versus transapical implants



Edwards SAPIEN Transcatheter Heart Valve Model 9000TFX and RetroFlex 3 and Ascendra Delivery Systems

FDA Review of P110021

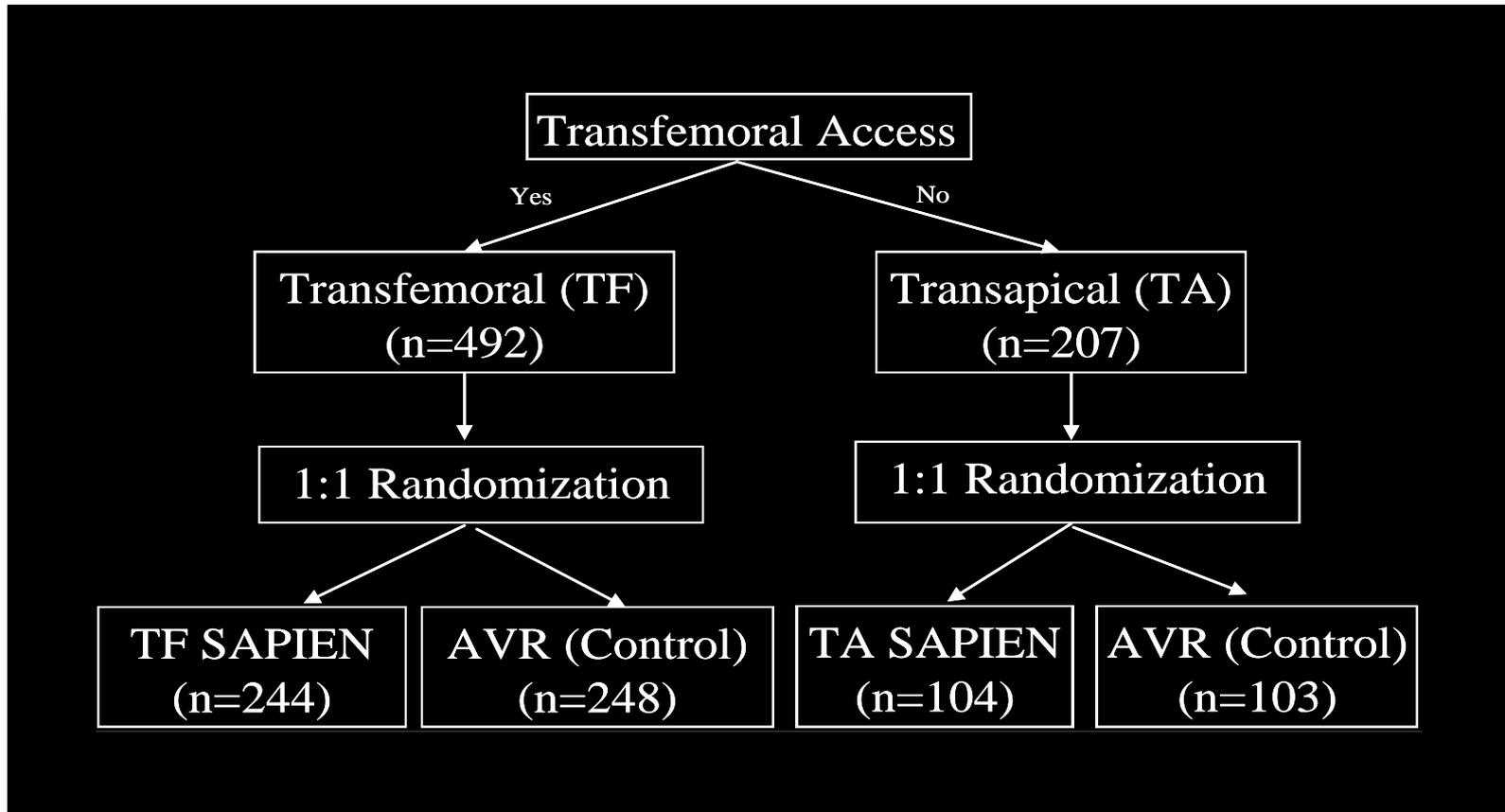
Lisa Kennell

Division of Cardiovascular Devices
Office of Device Evaluation
Food and Drug Administration

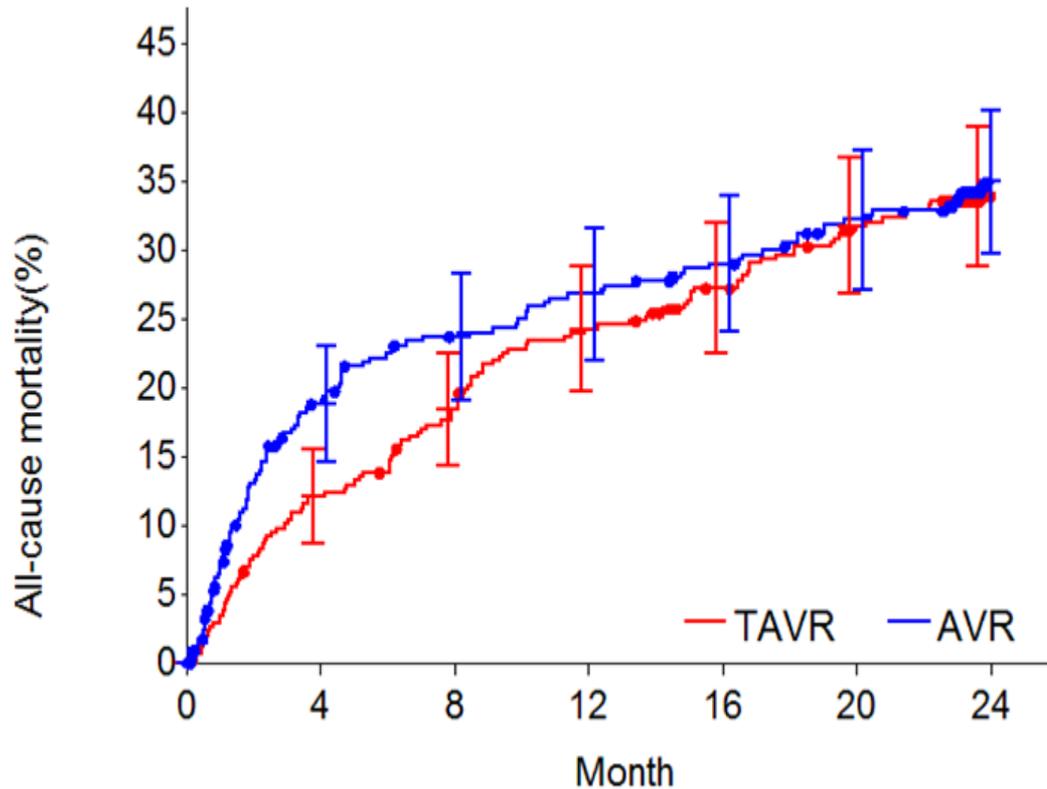
**Circulatory System Devices Panel Meeting
June 13, 2012**

A relatively sophisticated trial design

Prospective, non-blinded, randomized, controlled, multi-center clinical trial with non-inferiority outcome



Primary Safety/Effectiveness endpoint



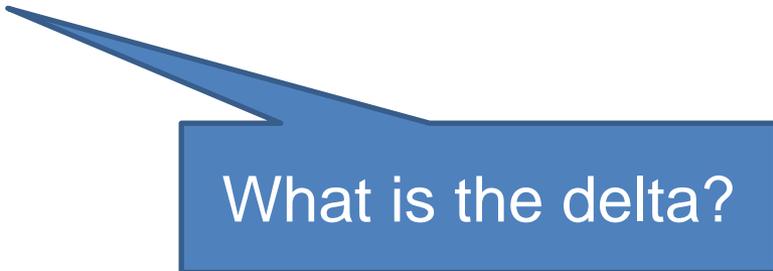
**At one year,
89/351 AVR patients died
84/348 TAVR patients died**

Primary hypothesis: survival at day 365

Non-inferiority study with delta 7.5%

Control AVR: 89/351 died; 26.8%

Exptal: 84/348 died; 24.3%



What is the delta?

Which one would you recommend?

Adaptive Trials

An adaptive design is one that “allows for prospectively planned modification based on accumulating study data without undermining the study’s integrity and validity”

FACTOR	FIXED TRIAL	ADAPTIVE TRIAL
Number of patients	Fixed	Variable
Patient population	Fixed	Can be narrowed
Randomization	Constant probability	Can be adjusted
Primary hypothesis	Fixed	Can be changed
Decision rules	Simple	Complex

Advantages of Adaptive Trials

- Can we stop early?
- Can we enrich the experiment?
- Can we use the control arm?
- Can we justify early termination?

How?

GUIDANCE DOCUMENT

Interacting with the FDA on Complex Innovative Trial Designs for Drugs and Biological Products

Guidance for Industry
DECEMBER 2020

**Medical
Studies**
**Industry and Food
Administration Staff**

Document issued on July 27, 2016.

Sample Size Adjustment

- Group sequential designs
- Sample size reestimation
- Adaptive recruitment
- Adaptive randomization
- Changes to eligibility
- Drop/add/change treatment arms
- Changes to statistics or endpoints
- Study duration

Usually determined by stages or endpoints

- Adaptation must be determined in detail before unblinding any data
- Often relies on data safety monitoring boards to protect the data to manage bias and statistical validity



Disdvantages:

- Damage statistical validity?
- Affect bias of clinical managers?
- Must be predicted in advance of approval



CDRH experience

- 251 adaptive studies 2007-2013*
 - Mostly designs (IDEs)
 - Some (32) product submissions (PMA, 510(k))
- Overwhelming majority are sample size related adaptations
 - Frequentist (156/176 sample size related adaptations)
 - Group sequential, sample size re-estimation
 - Bayesian (67/75 sample size related)
 - Sample size re-estimation
 - Adaptive recruitment

*Yang et. al., 2016. "Adaptive Design Practice at the Center for Devices and Radiological Health (CDRH), January 2007 to May 2013" Therapeutic Innovation and Regulatory Science 1-8.

Let us look at a pressure- ulcer trial using microstimulators



Fig.1: Gluteus Maximus¹²

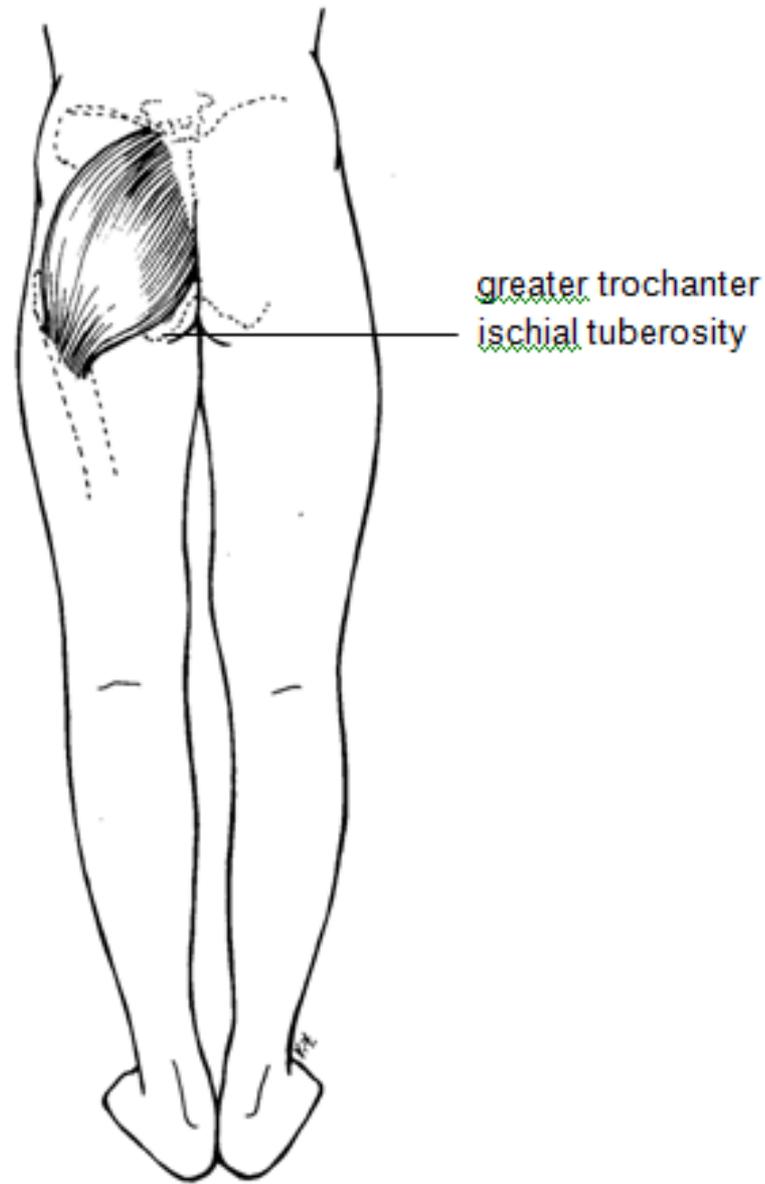
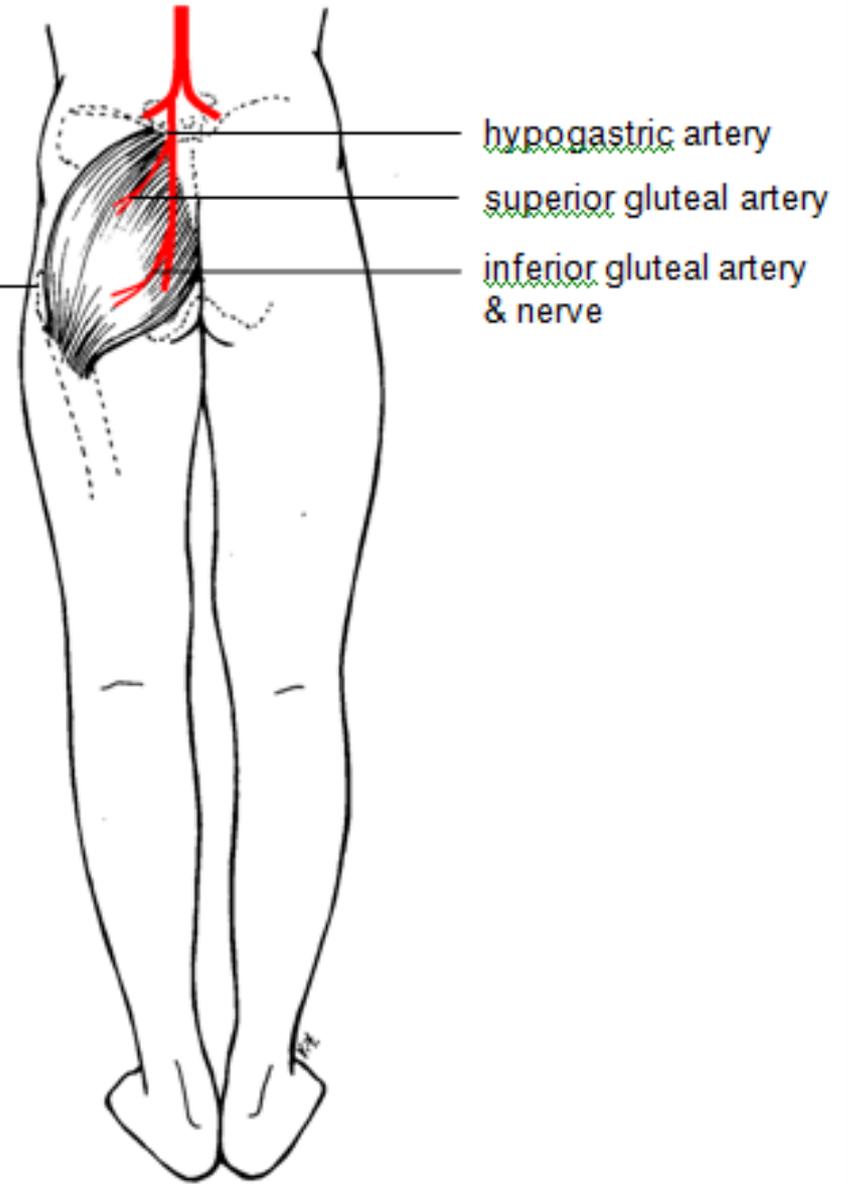
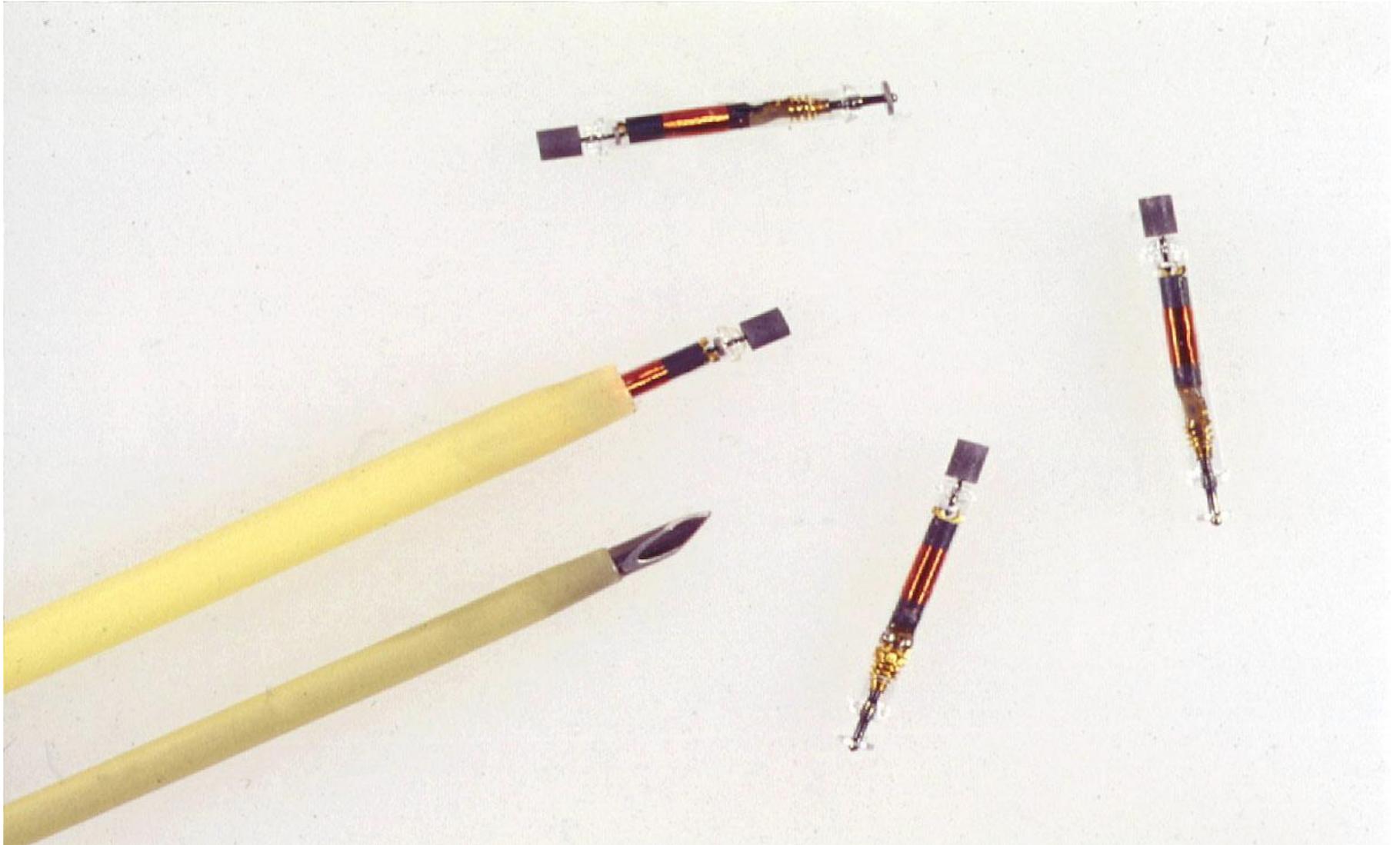


Fig.2: Neurovascular pedicles



Bions and Insertion tool

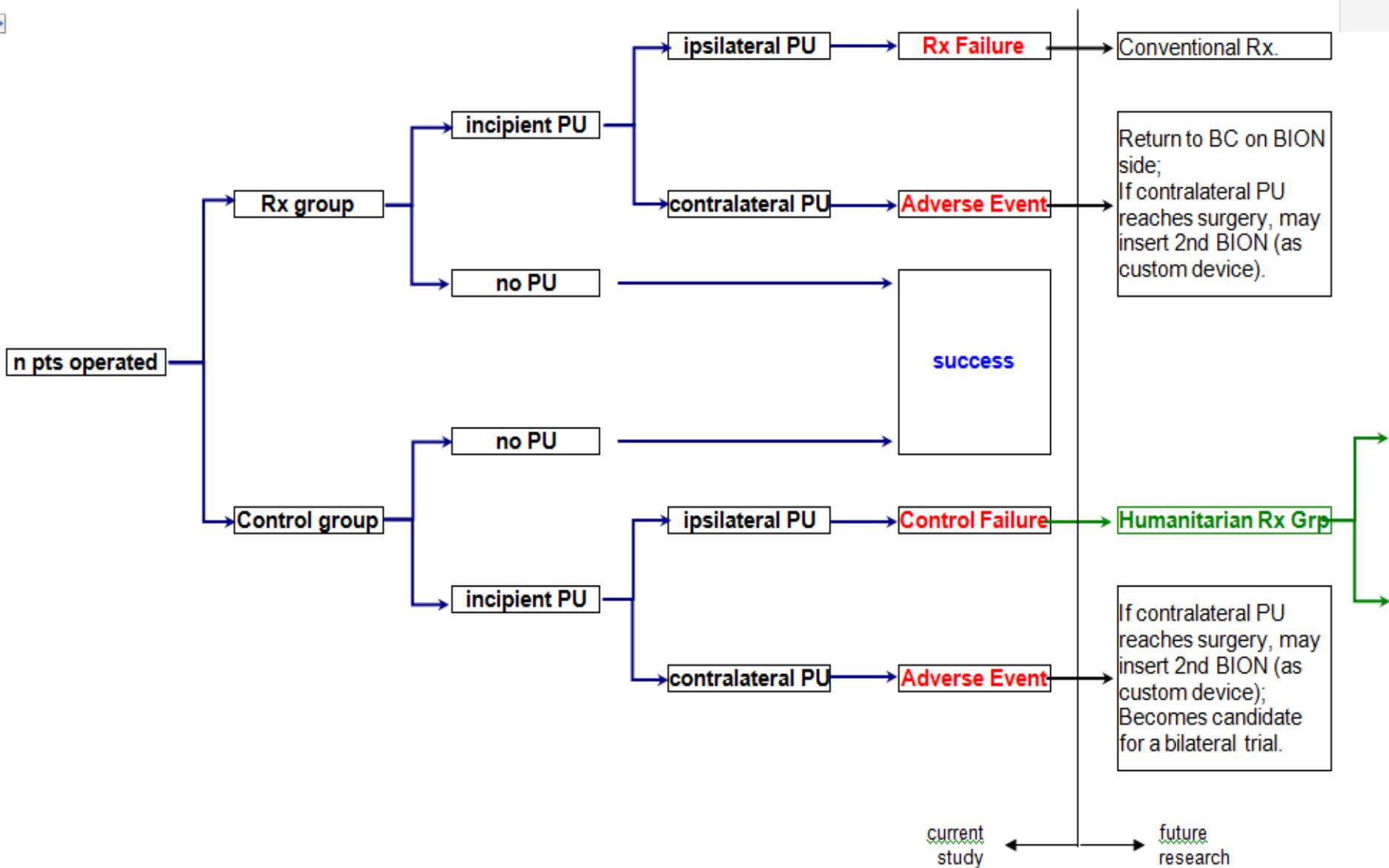


Inclusion Criteria

1. Subject has had a spinal cord injury and has bilateral lower limb paralysis.
2. Subject cannot contract voluntarily his gluteus muscle.
3. Subject is between 18 and 70 years old.
4. Subject sits in a wheelchair for at least 5 hours per day.
5. Attending physician considers the subject in general good health (other than SCI and PU wound).
6. Subject is having gluteal rotation flap surgery for PU treatment in which the inferior gluteal pedicle will be exposed but not damaged or sacrificed in any way (for any PU wound).
7. Subject is mentally capable of understanding the goals and the application of therapy.
8. Subject is able to apply the therapy (with or without help) once discharged from Rancho Los Amigos.
9. Subject is willing and capable of giving informed consent.
10. Subject is willing and capable of traveling to testing center at the schedule described above.

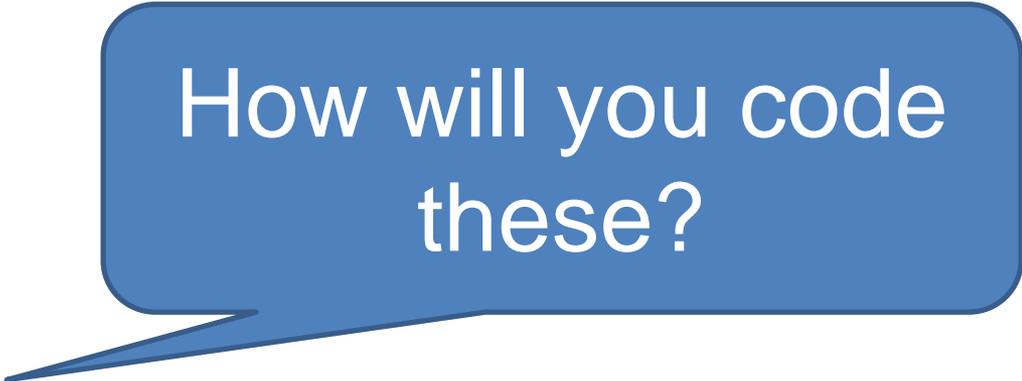
Exclusion Criteria

1. Subject is pregnant, nursing, or planning on becoming pregnant in the next 12 months
2. Subject has an electronic implant (ex: heart pacemaker, etc)
3. Subject has large metallic implant (e.g.: plates, hip joints) in the buttock/pelvic area (small metal implants, such as bone screws and metal sutures are acceptable).
4. Subject has any condition associated with wound healing abnormality (e.g.: connective tissue disorder, immune disorder, diabetes, clinical obesity)
5. Subject is malnourished
6. The attending physician has concerns about the healing of this subject (e.g.: heavy smoking, excessive and poorly-managed incontinence)
7. Subject has concurrent concomitant condition affecting the buttock/pelvic area, including other pressure wound not corrected by the flap surgery.
8. Claustrophobia or fear of having an MRI scan done.
9. Subject has damage to the inferior or superior gluteal neurovascular pedicles.



History

1. Sex
2. Age
3. Weight
4. Height
5. Side of PU
6. Stage of PU
7. History of PU
8. Patterns of tobacco, alcohol or caffeine use
9. Nutritional Status (eating well/poorly)
10. Allergies/Hypersensitivities
11. Drugs/medications
12. Other illnesses/conditions
13. Clinical trials in which the individual has participated previously



How will you code these?

Typical exercise program

The screenshot displays the Clinfit V2.0 software interface. The main window is titled "Program View: Medium 1" and shows a configuration for a "Medium 1" program. The program name is "Medium 1". The pulse trains frequency is set to "off", with a color-coded bar showing frequencies from 1 to 50 Hz. The selected muscle is "Deltoid (middle) : Left : Proximal". The waveform shows a blue trapezoidal pulse train. Below the waveform, the muscle "Supraspinatus : Left : Distal" is also visible. The "Cycles (sec)" control is set to 10. A "Channel Information [VIEW]" panel on the right provides detailed settings for the BION channel, including address, muscle, pulse shape, threshold, and pulse train parameters.

Channel Information [VIEW]

BION

Address: 16

Muscle: Deltoid (middle)

Pulse Shape

Current: 12.0 milliamps

Width: 104 microseconds

Threshold

Current: 4.0 milliamps

Width: 100 microseconds

Pulse Train

Intensity: 3.12 times-threshold

Train Delay: 1.0 seconds

Ramp Duration: 1.0 seconds

Train Duration: 5.0 seconds

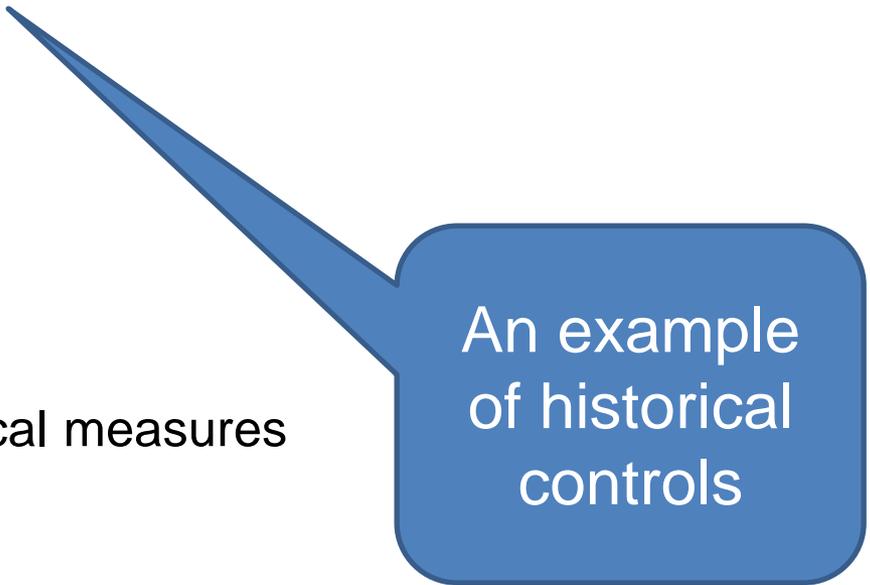
Recharge: 100 uA

OFFLINE 1:02 PM O:Guest P:W B

Another Study: Shoulder Subluxation

Outcome Measures

- subluxation - clinical and radiological measures
- muscle bulk - ultrasound
- muscle tone - modified Ashworth scale
- muscle power - manual muscle testing
- shoulder range of motion (active and passive)
- pain - visual analogue scale (at rest and with movement)
- arm/hand function - COVS (10a & b)



An example
of historical
controls

BT1-001 Experimental Patient



BT1-005 Control Patient

(opted for Bion therapy at the end of the trial)

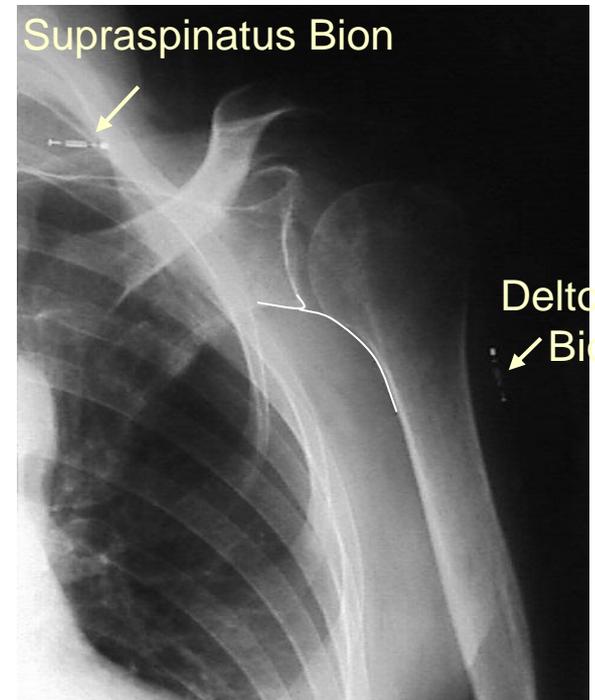
Initial Subluxation

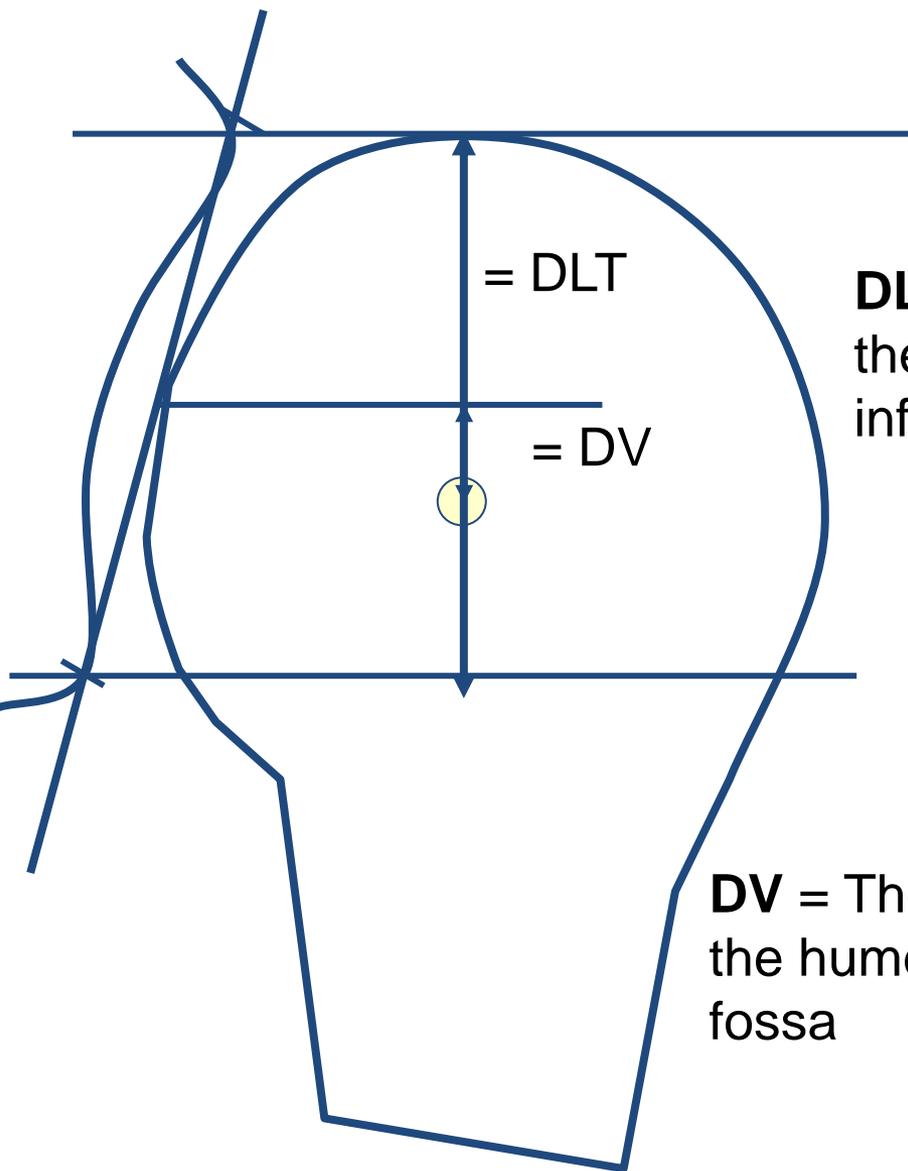


6 weeks conventional therapy



8 weeks after Bion therapy

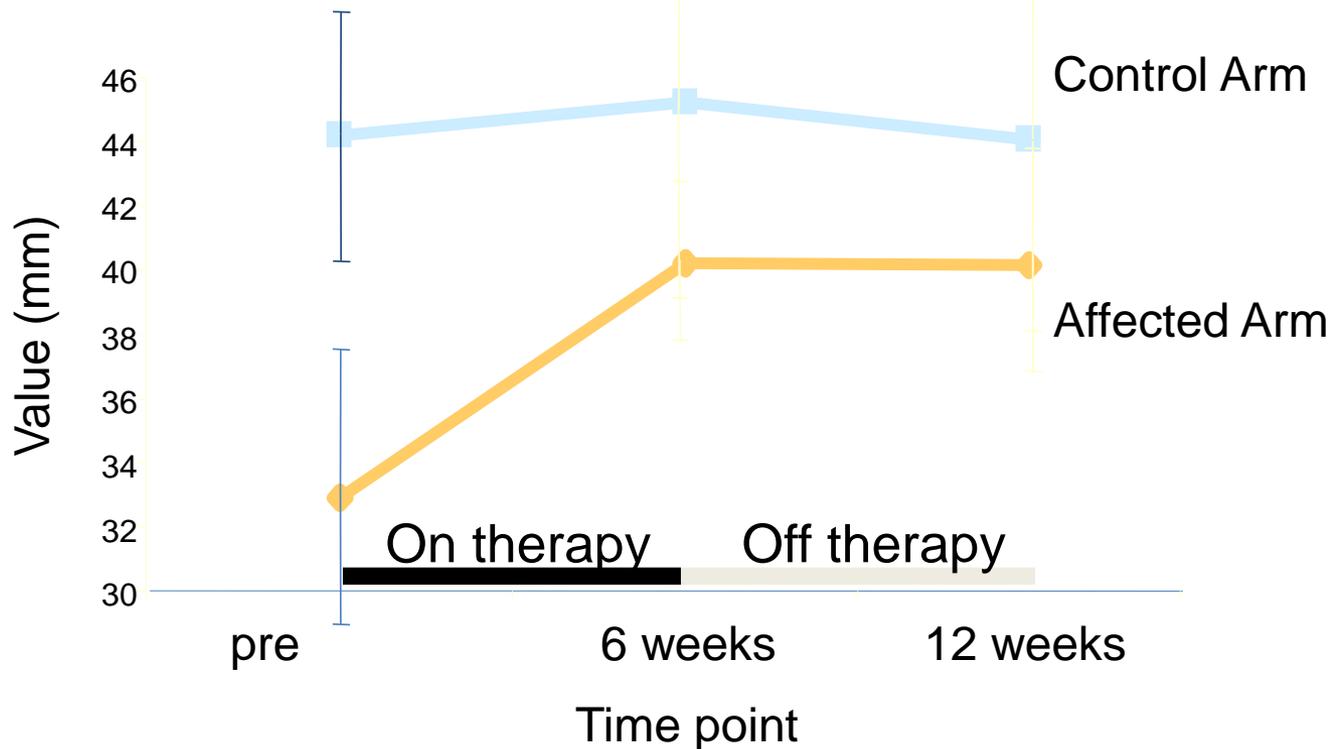




DLT = The vertical distance separating the apex of the humeral head and the inferior margin of the glenoid fossa

DV = The vertical distance linking the center of the humeral head to the center of the glenoid fossa

Average DLT measures for Patients undergoing Bion Therapy

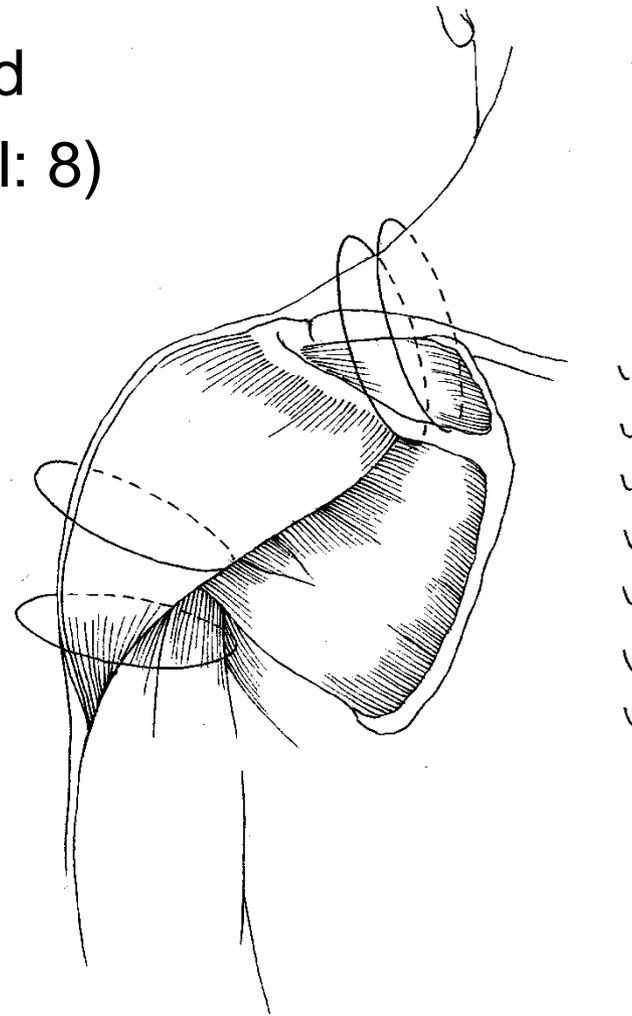


n=5

DLT = the vertical distance separating the apex of the humeral head and the inferior margin of the glenoid fossa

Measurements of muscle thickness using ultrasound

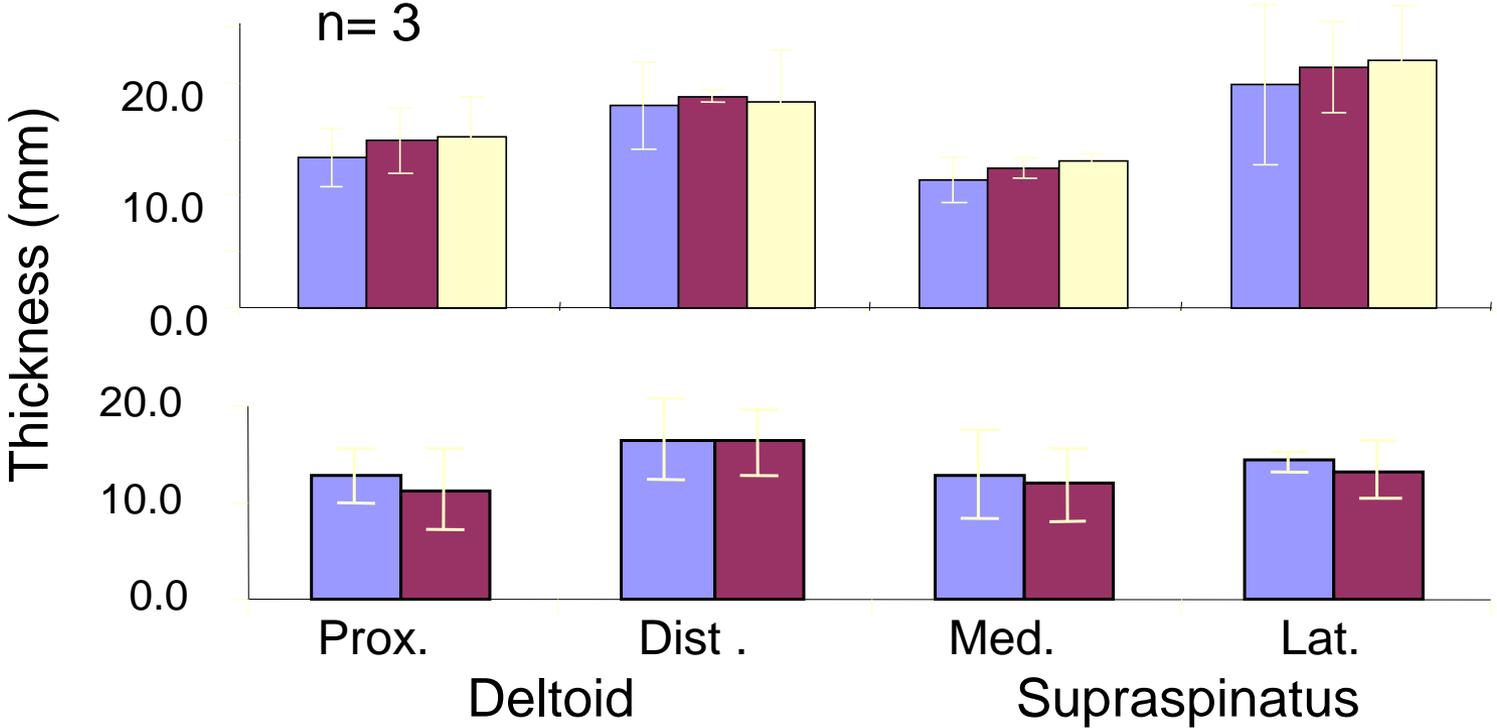
- 2 locations on each of deltoid and supraspinatus on each side (total: 8)
- use metal disk as marker



Average U/S Measures of Muscle Thickness on the Affected Arm

Experimental Patients

n= 3



Control Patients

n=3

- Pre
- 6 Wk On therapy
- 6 Wk Off therapy

Registries versus trials

Real world data is a rich source of information

Types of Registries and Their Main Uses

PATIENT

- Collect data regarding the health status of patients and their care
- Evaluate outcomes, best practices, and treatment guidelines
- Established by patient foundations and pharmaceutical organizations

SPECIALTY

- Focus on advancing care outcomes across a medical specialty or subspecialty
- Aim to develop guidelines and decision support tools and advance research
- May serve as QCDRs to allow clinicians to report to CMS under MIPS

POPULATION

- Focus on entire patient populations, spanning specialty care and specific diseases
- Seek to capture comprehensive population-level health status data

DEVICE

- Focus on tracking the safety and effectiveness of medical devices
- Support post-market surveillance
- Established by medical specialty organizations and medical device companies

PAYER

- Focus on improving outcomes and reducing costs
- Aim to measure and enhance value
- Established by healthcare payer organizations

QUESTIONS?



Pediatric Trials

Matthew Borzage, PhD

Fetal and Neonatal Institute, Division of Neonatology Children's Hospital Los Angeles, Department of Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles, CA

PREA

Pediatric Research Equity Act

Mandatory studies initiated by application for new dosing, route, API, or indication.

Waiver or Partial Waiver if: drug doesn't help children, is unsafe for children, or studies are impossible

Pediatric Study Plan (PSP) outline studies

- <60 days of End of Phase 2 meeting
- Before Phase 3
- >210 days before submitting application
- Deferral possible, with rationale

Required

BPCA

Best Pharmaceuticals for Children Act

Voluntary studies initiated by FDA's **Written Request (WR)**. WR's may be prompted by **Proposed Pediatric Study Request (PPSR)**.

6 months of exclusivity

Rewarded

Would it be
ethical to
exclude
patients
based on ...

Their Race or Ethnicity?

Their Sex or Gender?

Being Old?

Children are persons who have not attained the legal age for consent to treatments or procedures involved in the research, under the applicable law of the jurisdiction in which the research will be conducted.
(45 CFR 46.402(a)).



Child

Neonatal
Protections



Child

Just a Child
😊



Child

Might also
be an adult?

Extra Protections



Neonate (and Child)

Neonates are children with extra protections

From day of birth = day 0

Through through the 28th day after birth

- Viable
- Nonviable
- Uncertain Viability



45 CFR 46 Subpart B Additional Protections for Pregnant Women, Human Fetuses and Neonates Involved in R

45 CFR 46 Subpart D Additional Protections for Children Involved as Subjects in Research

AAP: 10.1542/pcco_book092_document002



Neonate and Child

Q: Must the child consent?

A: Normally they can't.

Q: Can the parents consent?

A: No, they can't either.

Permission means the agreement of parent(s) or guardian to the participation of their child or ward in research. (45 CFR 46.402(c)).



Q: Must the child consent?

A: Normally they can't.

Q: Can the parents permit?

A: Yes, (46.408 (b); 50.51).

Q: Must BOTH parents permit?

One Parent: no greater than minimal risk (46.404, 50.51(a))

OR greater than minimal risk with a prospect of direct benefit (46.405, 50.52)

Two Parents: greater than minimal risk, with no prospect of direct benefit, but likely to yield generalizable knowledge about the underlying condition or disorder (46.406, 50.53)





Q: Must the child consent?

A: Normally they can't.

Q: Can the parents permit?

A: Yes, (46.408 (b); 50.51).

Q: Must BOTH parents permit?

A: Depends on the risks



Q: Must the child consent?

A: Normally they can't.

Q: Can the parents permit?

A: Yes, (46.408 (b); 50.51).

Q: Must BOTH parents permit?

A: Depends on the risks

Q: Must the child assent?



21 CFR 50 Subpart D Additional Safeguards for Children in Clinical Investigations

45 CFR 46 Subpart D Additional Protections for Children Involved as Subjects in Research

45 CFR 46.408

Assent means a child's affirmative agreement to participate in research. Mere failure to object should not, absent affirmative agreement, be construed as assent. (45 CFR 46.402(b)).



Q: Must the child consent?

A: Normally they can't.

Q: Can the parents permit?

A: Yes, (46.408 (b); 50.51).

Q: Must BOTH parents permit?

A: Depends on the risks

Q: Must the child assent?

A: Yes unless:

- They are too limited to assent
- There's a benefit only available if they assent
- Informed consent for an adult would be waived

21 CFR 50 Subpart D Additional Safeguards for Children in Clinical Investigations

45 CFR 46 Subpart D Additional Protections for Children Involved as Subjects in Research

45 CFR 46.408

How do logistics and recruitment bias data?

Do I recruit across IRB-imposed age thresholds?

Consenting vs Assenting vs Permission yields different cohorts.

Will I have greater data loss at different time points and is it a confound?

E.g. progressive intellectual disability from phenylketonuria.

How are the dates and times of study activities biasing data?

Time of day (vs school), day of week (vs social), month of year (vs holidays)

Does study design bias recruitment?

Is an adult leaving work a different burden than their child missing school?

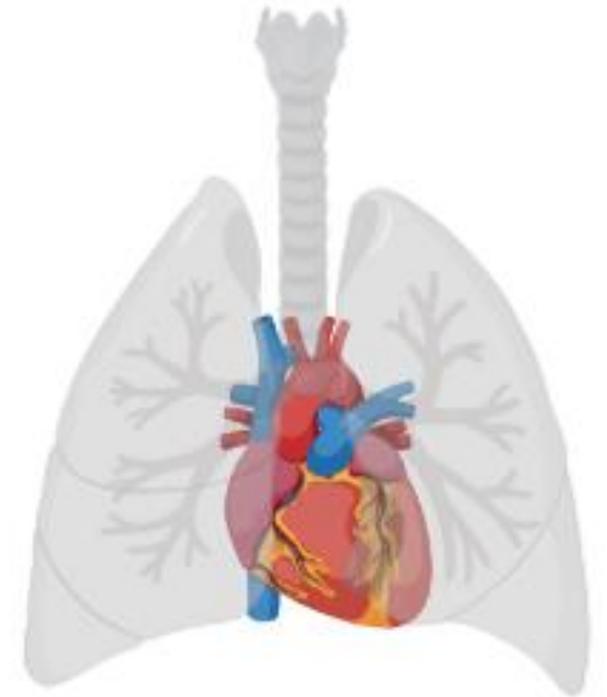
Yes. Homework, note to school, extracurricular activities, etc.

Do parents accept different benefits/risks for themselves vs their children?

Yes. The benefit/risk assessment is completely different.



Caffeine



- ↓ apnea
- ↓ hypoxemia
- ↓ mechanical ventilation
- ↓ heart defects

Design an experiment with a washout to get drug down to 1%,
with $\tau = 16$ hours $N(t) = N_0 \left(\frac{1}{2}\right)^{(t/\tau)}$
T=106 h

Multiple Variables & Their Interaction Moderate ADME!

Non-Responders?

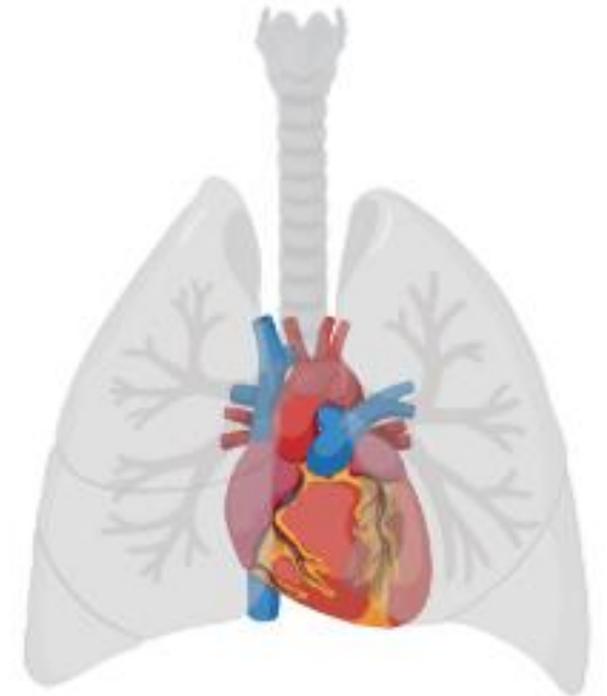
A) 0-4 weeks	& breast fed	40%
B) 22-26 weeks	& breast fed	26%
C) 0-4 weeks	& formula fed	22%
D) 22-26 weeks	& formula fed	1%



Caffeine

Breastfeeding vs
Formula

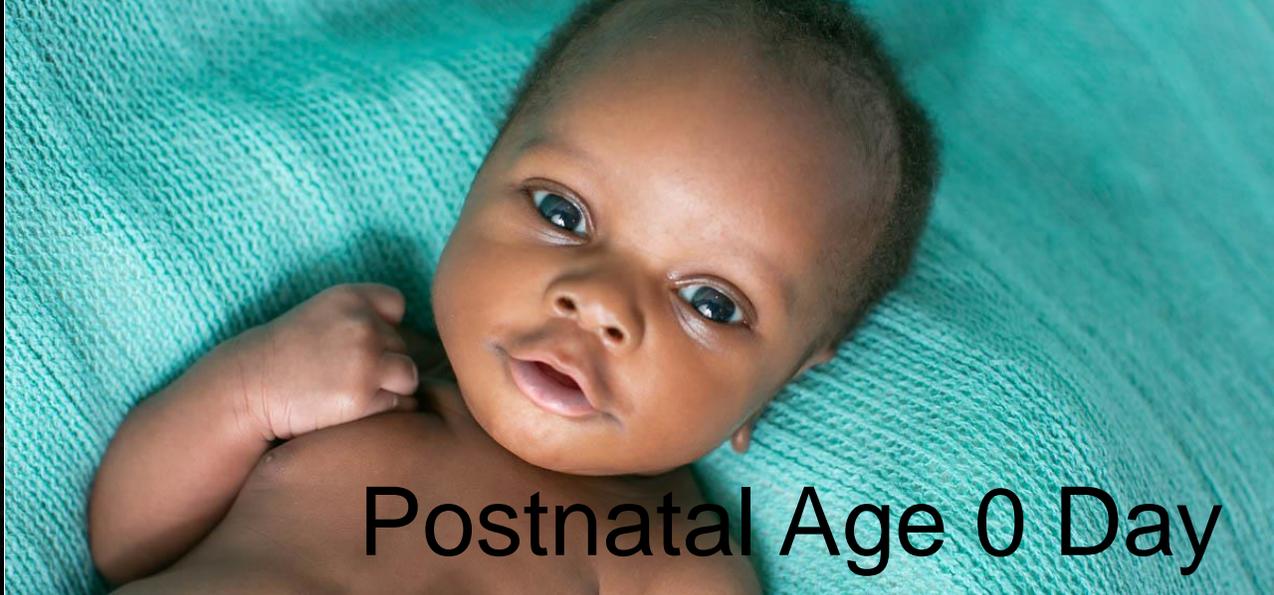
Postnatal Age



- ↓ apnea
- ↓ hypoxemia
- ↓ mechanical ventilation
- ↓ heart defects



21 Weeks Post Menstruation



Postnatal Age 0 Day

40 Weeks Post Menstruation



Postnatal Age 147 Days

Post Menstrual Age (PMA)

- Approximately 16% of deliveries do not have an accurate PMA

Increased variance in some of the data!

- Demographics affect which women have accurate or inaccurate PMA

Data has nonrandom subgroups with uneven variance = Homoscedasticity!

- At least 12.9% of birth certificate dates are affected by digit preference.

The data is not distributed continuously.

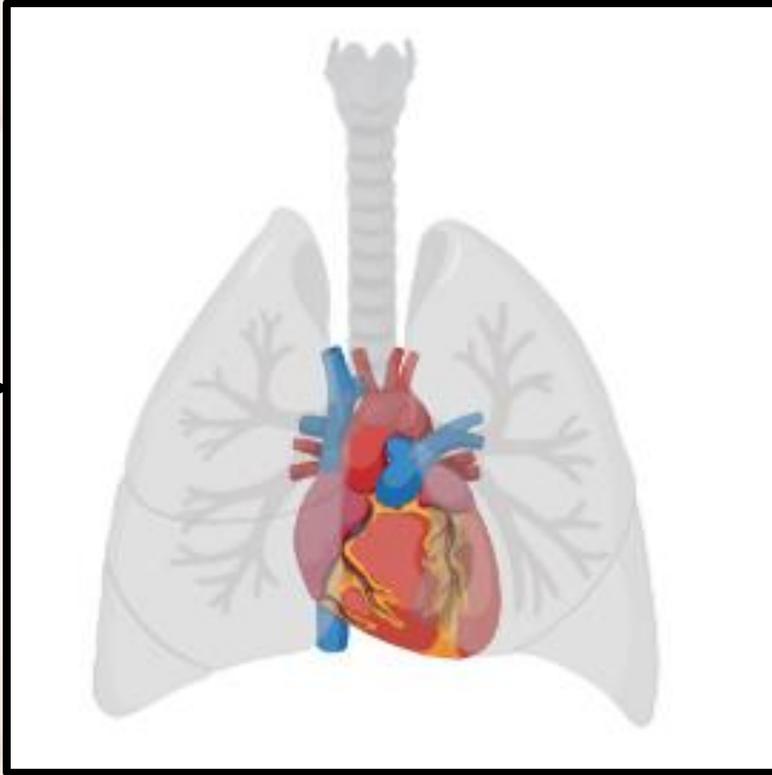
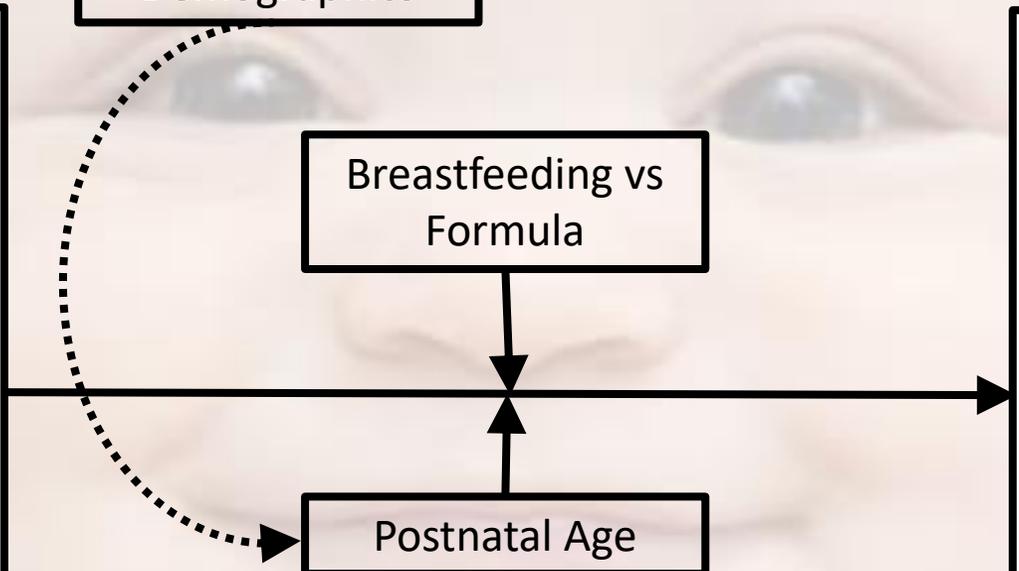


Caffeine

Demographics

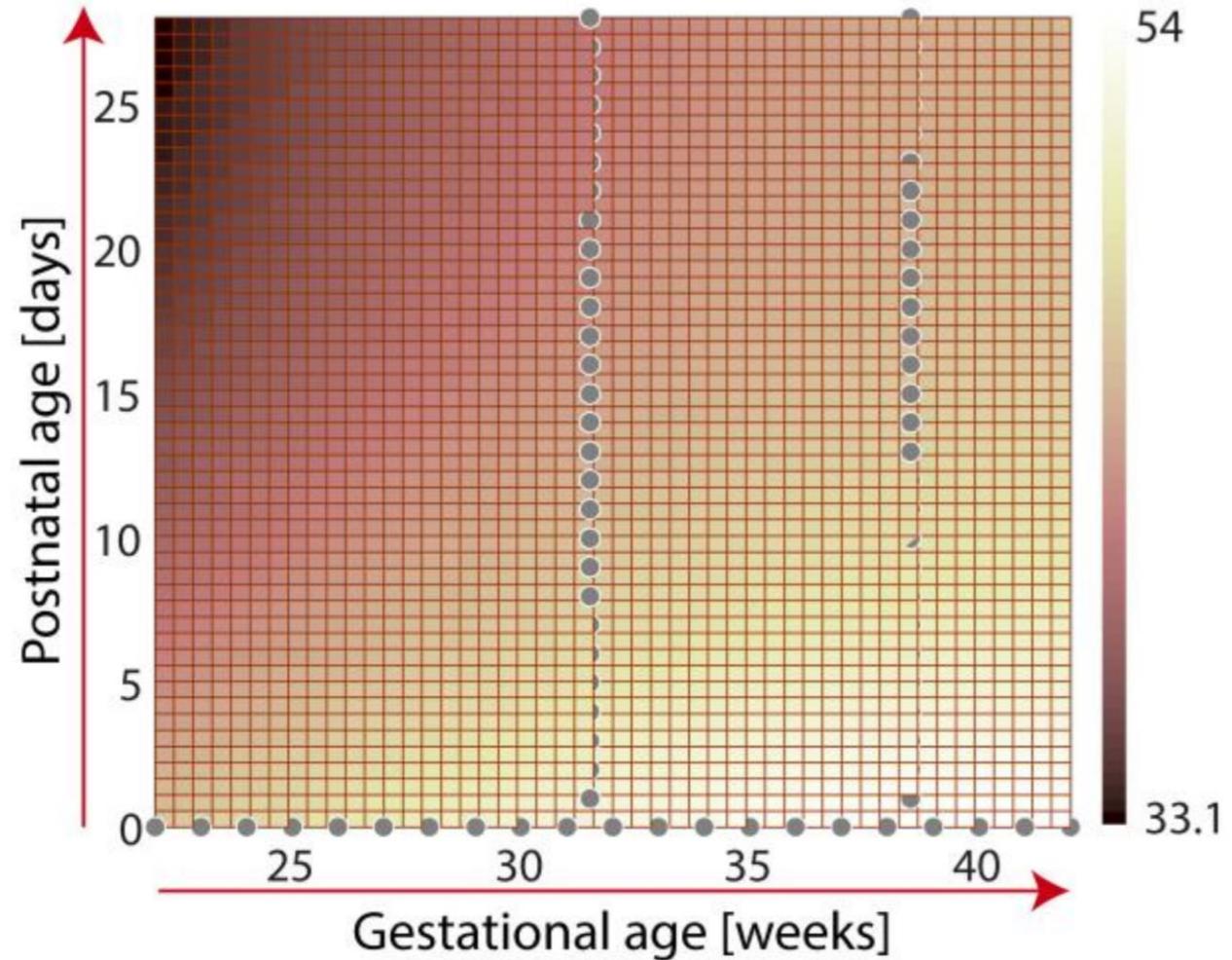
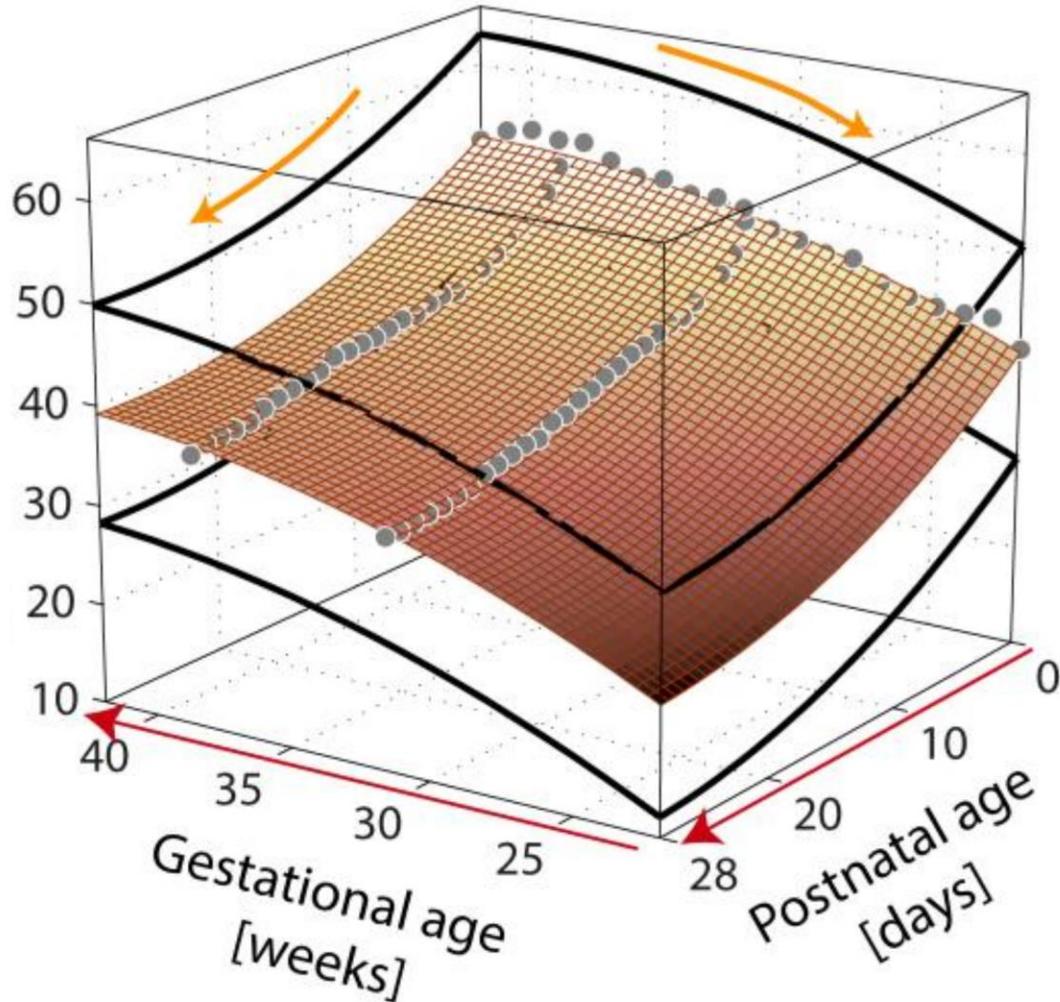
Breastfeeding vs Formula

Postnatal Age



- ↓ apnea
- ↓ hypoxemia
- ↓ mechanical ventilation
- ↓ heart defects

Multiple Variables (PMA & Postnatal Age) and their Interaction all Affect Physiology





Caffeine

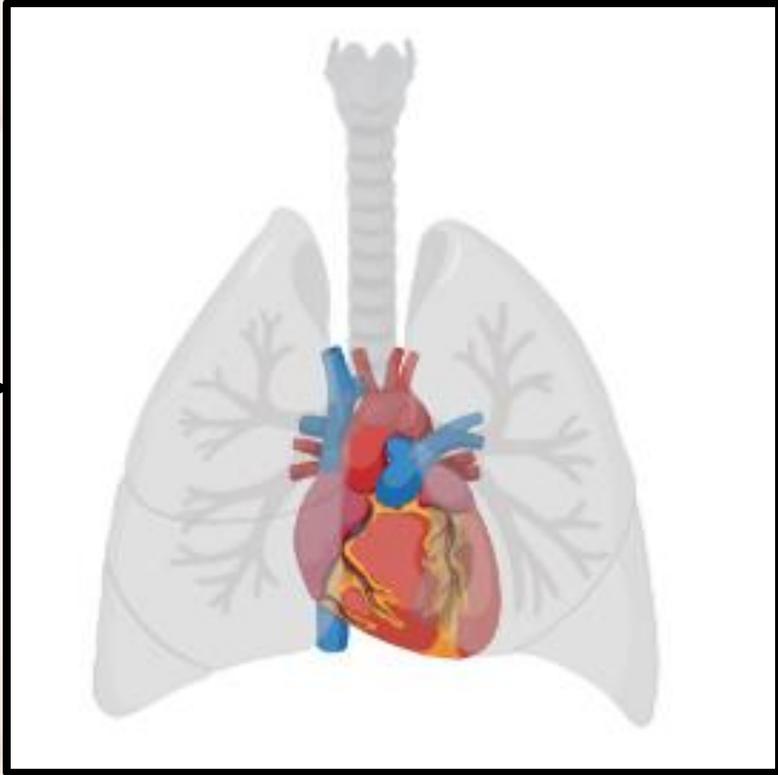
Demographics

Breastfeeding vs Formula

Postnatal Age

PMA

Physiology



- ↓ apnea
- ↓ hypoxemia
- ↓ mechanical ventilation
- ↓ heart defects

Demographics of Mothers Breastfeeding at 6 Months

Race

Education

Mother's Age

Maternal demographics influence feeding (moderating variable)!

27% Black

36% HS

19% <20 years

43% White

56% College

49% >30 years

45% Hispanic

52% Asian

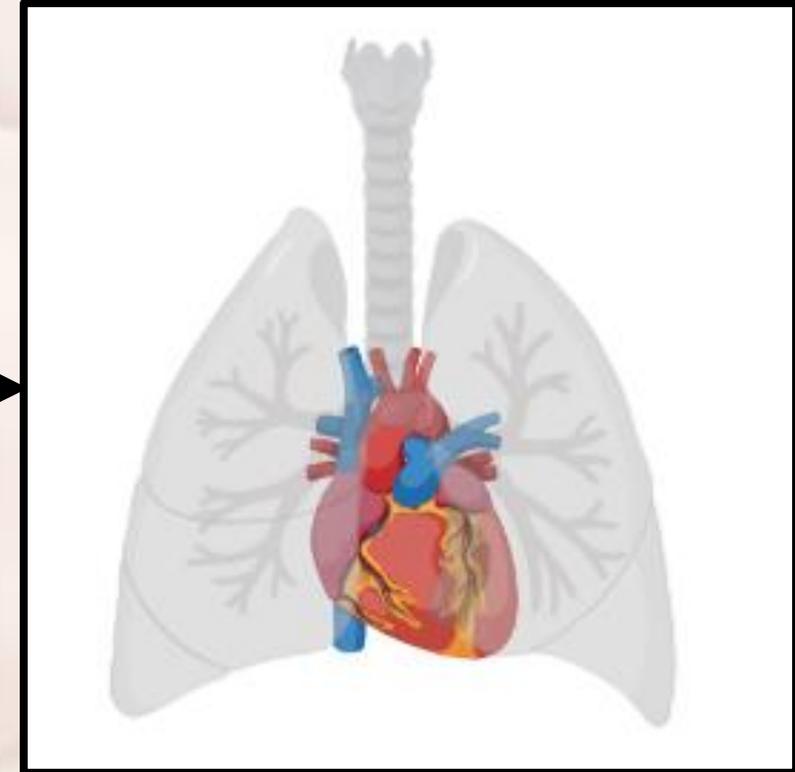
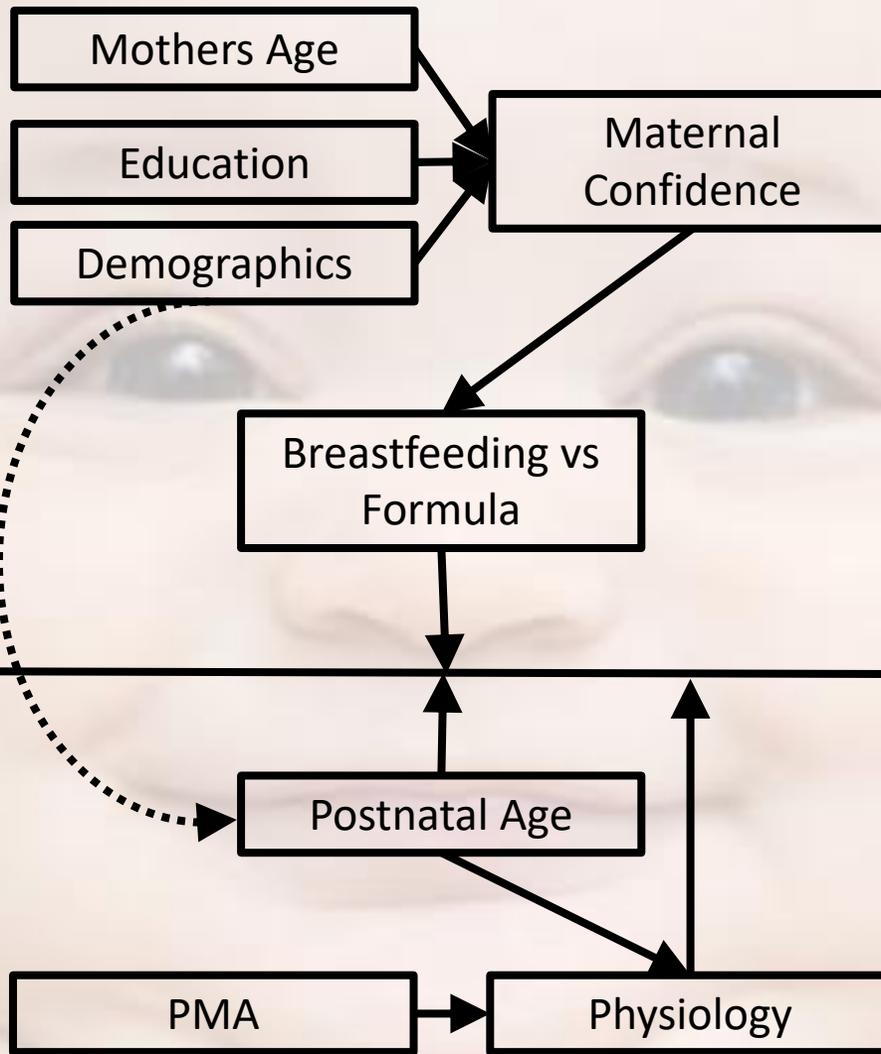
But the hidden (latent) factor was confidence!

But what's the best way to predict who would continue trying breastfeeding?

Women who were confident (27%) versus those who were not (5%)



Caffeine



- ↓ apnea
- ↓ hypoxemia
- ↓ mechanical ventilation
- ↓ heart defects



Expensive
Small effects
Large sample sizes
Fixed by randomization?

Inexpensive
Large effects
Small sample size
Unfixable by randomization

Ignore

Address

Studying Caffeine for Newborns

Age Range (marketing, PK/PD, accuracy of variables, uneven variance, chronological vs developmental)

Moderating Effects (demographics)

Mediating Effects (physiology)

Latent Factor (maternal confidence)

Recruitment and Allocation Strategy (which attributes to control?)

Are My Outcomes Measurable in this Child?

Imaging

Cranial Ultrasound (younger), MRI (young & older), X-Ray (older)

Bloodwork & Laboratory Measures

Volume of blood (limited in younger)

Questionnaires

Cognition (language, age-specific tests)

Behavior (parent, child, or teacher)

Pain (visual scale or numeric)

What Else is Changing?

- Cardiopulmonary: heart goes from parallel to series circulation
- Skeletal: plate fusion in the skull, long bones
- Hormones: infant and adolescent
- Environmental: parents move, new schools

Did I miss a developmental milestone or fail to appreciate all its effects?

Puberty – hormonal, behavioral, fetal exposure, hematocrit
brain blood flow

Take Away:
Your data are complex,
and your intuition is
wrong.

?

Borzage@usc.edu

CTSI Clinical Study Design Types

Wendy Mack, PhD

Southern California Clinical and Translational Science Institute:

Biostatistics, Epidemiology and Research Design (BERD)

Biostatistics, Epidemiology and Research Design (BERD)

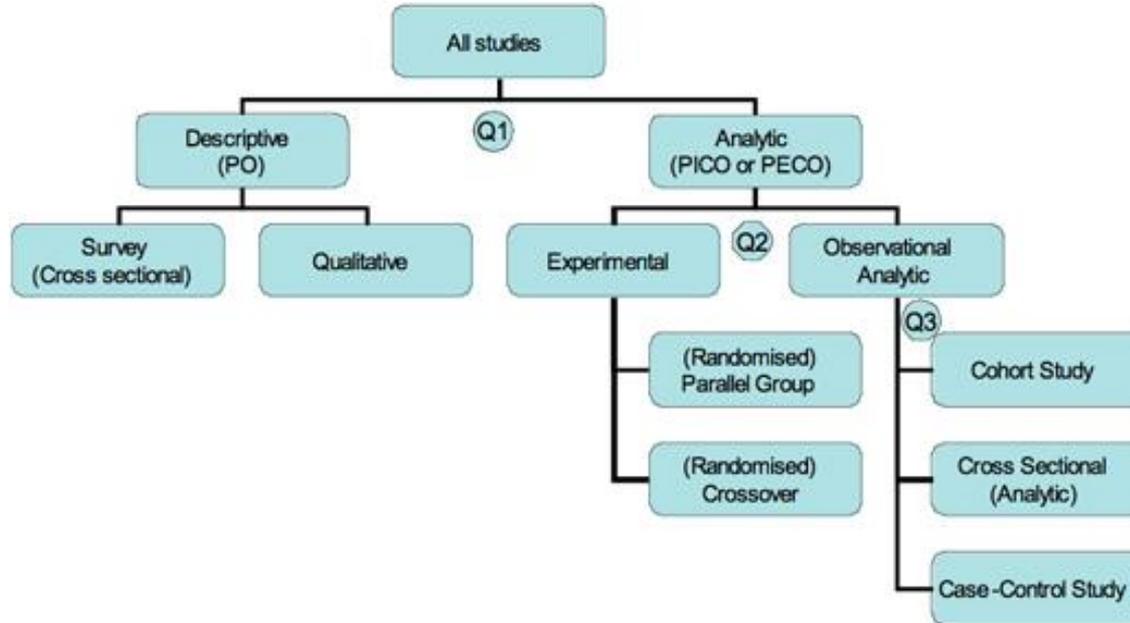
Objectives

- Review of study designs, including clinical trials
- Examples of study designs and clinical trials
- Alignment of study designs with:
 - Research question
 - Data collection
 - Statistical analysis: What statistical methods are appropriate for study design and data collected?

PICOT Criteria to Develop the Research Question

- **P Population**
What specific population will you test the intervention in?
- **I Intervention (or Exposure)**
What is the intervention/exposure to be investigated?
Intervention (clinical trial); Exposure (observational study)
- **C Comparison Group**
What is the main comparator to judge the effect of the intervention?
- **O Outcome**
What will you measure, improve, affect?
- **T Time**
Over what time period will outcome be assessed?

Spectrum of Study Designs



From Center for Evidence-Based Medicine (CEBM), University of Oxford
<http://www.cebm.net/study-designs/>

Descriptive vs. Analytic Study

- Descriptive Study: Research question involving “PO” (Population, Outcome)
- What is the survival rate following hip fracture in community-dwelling postmenopausal women?

(P: Community-dwelling postmenopausal women with hip fracture; O: Survival)

Descriptive vs. Analytic Study

- Analytic Study: Research question adds I/E and C (Intervention/Exposure, Comparator Group).

Questions of association and/or effect (I/E on O).

Comparator group: Not “exposed”, does not get “intervention”

- Does the survival rate following hip fracture differ in postmenopausal women who live with others vs alone? (E:live with others; C:live alone)

Observational Study Defined

- Clinicaltrials.gov: A clinical study in which participants identified as **belonging to study groups** are assessed for biomedical or health **outcomes**. Participants may receive diagnostic, therapeutic, or other types of interventions, but the investigator **does not assign** participants to specific interventions (as in a clinical trial). Exposures (interventions) are **self-selected**.
- Associations between exposures/interventions and outcomes may be biased (confounded) by characteristics that differ between those that choose exposure vs. no exposure.

Cohort Study

The Association Between Inhaled Nitric Oxide Treatment and ICU Mortality and 28-Day Ventilator-Free Days in Pediatric Acute Respiratory Distress Syndrome

Anoopindar K. Bhalla, MD^{1,2}; Nadir Yehya, MD³; Wendy J. Mack, PhD^{4,5}; Melissa L. Wilson, MPH, PhD^{4,5}; Robinder G. Khemani, MD, MSCI^{1,2}; Christopher J. L. Newth, MD, FRCPC^{1,2}

Cohort Study

Select persons **free of outcome**, including persons **with and without exposure**. Follow forward in time to determine outcome.

Does the proportion of persons with outcome (or rates of outcome) differ in persons with versus without the exposure?

Population: Children in pediatric ICU for acute respiratory distress syndrome

Exposed (E+): Inhaled nitric oxide treatment

Not exposed (E-, C): No inhaled nitric oxide treatment

Outcome: ICU mortality; 28-day ventilator-free days



Cohort Study

What types of statistics did we use to compare ICU mortality?

Chi-square (comparing proportions who died in each group)

Logistic regression with dichotomous outcome (mortality) to control for confounders

Higher-level statistics (**propensity scores**) to control for factors related to who gets iNO vs not

Results:

Unadjusted mortality **higher** in patients receiving iNO (25.2%) compared to those not receiving iNO (16.3%); $p = 0.02$. Mortality did not differ with adjustment.

TABLE 3. The Association Between Treatment With Inhaled Nitric Oxide and Mortality

iNO Treatment	Unadjusted Analysis ($n = 499$) ^a		Matched Analysis ($n = 176$) ^{b,c}		Inverse Probability Weighting ($n = 464$) ^{a,c}		Stratification Analysis ($n = 464$) ^{a,c,d}	
	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p	OR (95% CI)	p
iNO treatment	1.7 (1.1–2.8)	0.02	1.3 (0.56–3.0)	0.54	2.2 (0.59–8)	0.24	1.6 (0.85–3.1)	0.14

Case-Control Study

VRE in cirrhotic patients

Melissa Barger¹, Emily Blodget^{2*}, Sol Pena³, Wendy Mack⁴ and Tse-Ling Fong⁵

BMC Infectious Diseases 2019; <https://doi.org/10.1186/s12879-019-4352-1>

Case-Control Study

Select persons with (cases) and without (controls) outcome (O); determine their **past exposure** (E; i.e., BEFORE the outcome occurred).

Does the proportion of persons who were exposed differ in cases and controls?

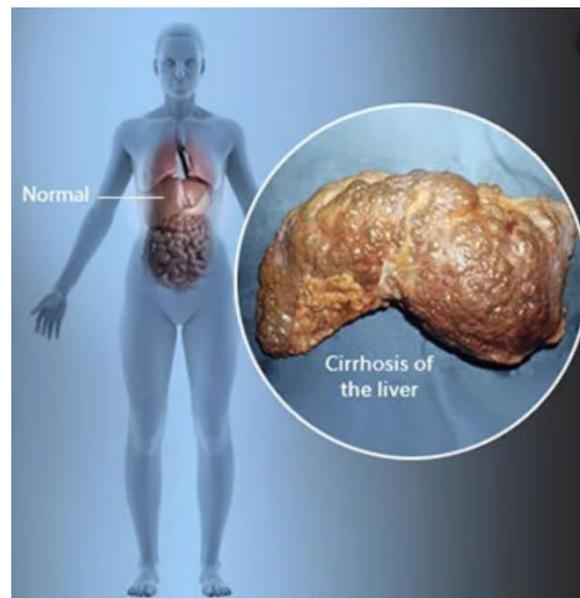
Population: Patients with liver cirrhosis

Cases (O+): Infected with Vancomycin-Resistant Enterococci bacteria (VRE)

Controls (O-, C): Not infected with VRE

Exposure(s): Demographic and clinical factors

Compare: The proportion or mean exposure levels between cases vs controls.



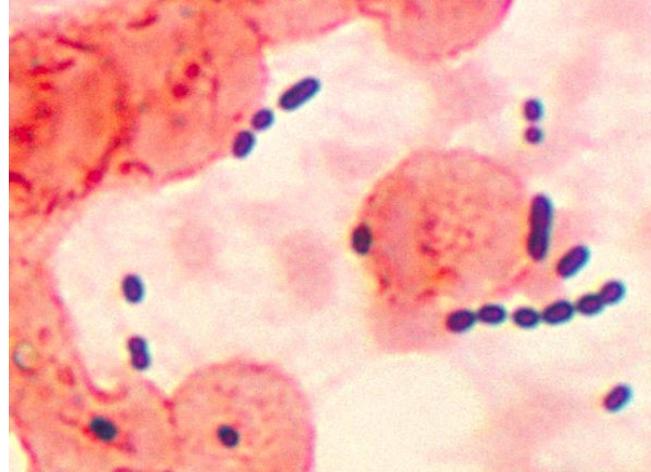
Case-Control Study

What types of statistics did we use to compare VRE+ and VRE-?

Chi-square test for categorical E (e.g., proportion male/female by VRE group)

t-test or non-parametric Wilcoxon rank sum for continuous E (e.g., mean age by VRE group)

Logistic regression for dichotomous outcome (VRE+, VRE-) to control for confounding biases.



Case-Control Study

Table 3 Multivariable Associations with VRE infection

Variable	OR (95% CI)	<i>p</i> -value
a. VRE Infected vs. VRE Negative, Gram Positive		
Female	3.73 (1.64, 8.49)	0.002
Child Pugh A or B (vs C)	0.37 (0.16, 0.84)	0.018
Ascites	9.43 (3.22, 27.65)	< 0.001
Dialysis	3.31 (1.21, 9.04)	0.02
b. VRE Infected vs. VRE Negative, Sterile		
Age		
< 47	1.0	0.041
47–53	0.32 (0.12, 0.88)	
54–60	0.57 (0.24, 1.36)	
≥ 60	0.32 (0.14, 0.75)	
Female	5.60 (2.90, 10.80)	< 0.001
Child Pugh A or B (vs C)	0.42 (0.21, 0.83)	0.013
Ascites	4.86 (1.74, 13.61)	0.003
Dialysis	3.19 (1.48, 6.87)	0.003
Any Antibiotic	2.37 (1.27, 4.42)	0.007

Clinical Trial Defined

Clinicaltrials.gov: A **clinical study** in which participants are **assigned** to receive one or more **interventions** (or **no intervention**) so that researchers can evaluate the effects of the interventions on **biomedical or health-related outcomes**. The assignments are determined by the study protocol. Participants may receive diagnostic, therapeutic, or other types of interventions.

A cohort study where persons are “assigned” to exposures (interventions) and followed for ascertainment of outcomes.

Clinical trials are not feasible when assignment to an exposure/intervention is not ethical.

Quasi-Experimental Study

Assign groups without randomization.

“Natural” quasi-experiments often occur in healthcare settings, where one can evaluate the possible impact of newly introduced interventions or practices (at patient or systems level).

Pre-post comparisons (the comparator is pre-intervention).

Quasi-Experimental Study

The Effect of Utilization Review on Emergency Department Operations

Shoma Desai, MD*; Phillip F. Gruber, MD; Erick Eiting, MD, MMM; Seth A. Seabury, PhD; Wendy J. Mack, PhD; Christian Voyageur, BA; Veronica Vasquez, MD; Hyung T. Kim, MD; Sophie Terp, MD, MPH

Ann Emerg Med 2017;70:623-631. PMID: 28559030

Quasi-Experimental Study

Evaluated impact of utilization review software implementation (I) designed to reduce hospital admissions.

Outcomes: Effectiveness (numbers of hospital admissions); safety (re-visits to ED)

Comparator (C): Pre-implementation period.



Quasi-Experimental Study

What statistics did we use to compare pre- and post-implementation?

A **regression model for counts** (e.g., # admitted per 30-day period)

Table 2. Patient dispositions.

Variables	Preimplementation Mean (95% CI)	Postimplementation Mean (95% CI)
ED LOS, h	8.1 (7.8–8.5)	8.5 (8.1–9.0)
Hospital LOS, days	6.5 (6.1–6.9)	6.5 (6.2–6.7)
Observation unit LOS, h	24.4 (23.5–25.2)	27.1 (26.4–27.8)
Admission rate (total)	14.2 (13.6–14.8)	12.8 (12.3–13.4)
Admission rate (unmonitored)	10.9 (10.6–11.3)	9.3 (9.0–9.6)
Admission rate (monitored)	3.3 (3.0–3.6)	3.5 (3.2–3.8)
Discharge rate	82.4 (81.8–82.9)	83.4 (82.8–83.9)
Transfer rate	0.9 (0.9–1.0)	0.5 (0.5–0.6)
Observation rate	2.5 (2.3–2.7)	3.4 (3.1–3.6)
30-day revisit rate, total	20.4 (19.9–20.9)	24.4 (23.8–25.0)
30-day admission rate, total	3.2 (3.1–3.4)	2.8 (2.7–3.0)

Clinical Trial Designs: Parallel Group

- **Parallel group:** Each participant is assigned to one (and only one) of the trial interventions. Standard approach for most clinical trials

Adipose tissue inflammation in breast cancer survivors: effects of a 16-week combined aerobic and resistance exercise training intervention

Christina M. Dieli-Conwright¹  · Jean-Hugues Parmentier² · Nathalie Sami¹ · Kyuwan Lee¹ · Darcy Spicer³ · Wendy J. Mack⁴ · Fred Sattler³ · Steven D. Mittelman^{2,5}

Breast Cancer Res Treat 2018;168:147-157. PMID: 29168064

Clinical Trial Designs: Parallel Group

Population: Obese female breast cancer survivors

Randomized Intervention: 16-week endurance and resistance exercise program

Randomized Comparator: Delayed exercise program

Outcomes: Body composition; cardiometabolic risk measures; systemic inflammation

Clinical Trial Designs: Parallel Group

What statistics did we use to compare these randomized groups?

Changes in outcomes (pre/post intervention) were not normally distributed:

Within group: **Non-parametric Wilcoxon signed rank**

Between group: **Non-parametric Wilcoxon rank sum**

For future study planning purposes: **Intervention effect sizes**

Clinical Trial Designs: Parallel Group

Results example:

Table 3 Effects of Exercise on body composition

	EX group <i>N</i> = 10	CON group <i>N</i> = 10	<i>p</i> value ^a	Effect size ^b
Weight, kg				
Pre	84.99 ± 10.53	84.53 ± 10.54		
Change	- 3.69 ± 2.12	0.47 ± 0.67	0.62	
Within group <i>p</i> value ^c	0.002	0.062	0.0002	1.58
Lean mass, kg				
Pre	53.78 ± 7.90	53.69 ± 8.37		
Change	1.80 ± 2.30	- 0.71 ± 2.30	0.85	
Within group <i>p</i> value ^c	0.049	0.065	0.03	0.97

Clinical Trial Designs: Crossover

- **Crossover:** Each participant receives both the experimental and comparator interventions, usually in randomized order, with a washout period between interventions

Perfect matching – each participant acts as their own control – requires fewer subjects

Disadvantages: Greater likelihood of dropout; must be a stable disease under study; only appropriate for interventions that wash-out and have short-term (not permanent) outcomes

Appetite-Regulating Hormones Are Reduced After Oral Sucrose vs Glucose: Influence of Obesity, Insulin Resistance, and Sex

Alexandra G. Yunker,^{1,2} Shan Luo,^{1,2,3} Sabrina Jones,^{1,2} Hilary M. Dorton,^{2,4} Jasmin M. Alves,^{1,2} Brendan Angelo,^{1,2} Alexis DeFendis,^{1,2} Trevor A. Pickering,⁵ John R. Monterosso,^{3,4} and Kathleen A. Page^{1,2}

Journal of Clinical Endocrinology and Metabolism, 2020

Clinical Trial Designs: Crossover

Brain and behavior responses to sweetened beverages

Population: Young adults (18-35 years) in three groups (lean, obesity prone, obese)

Interventions (order randomized): glucose and sucrose drinks (75g)
Sucrose: most common sugar we consume (“table sugar”)

Outcomes: short-term responses in brain activation (fMRI), **hunger/satiety hormones (reported here)**, behavioral ratings of hunger/desire for food

Clinical Trial Designs: Crossover

How did we analyze these data to compare the outcomes following glucose vs. sucrose drinks?

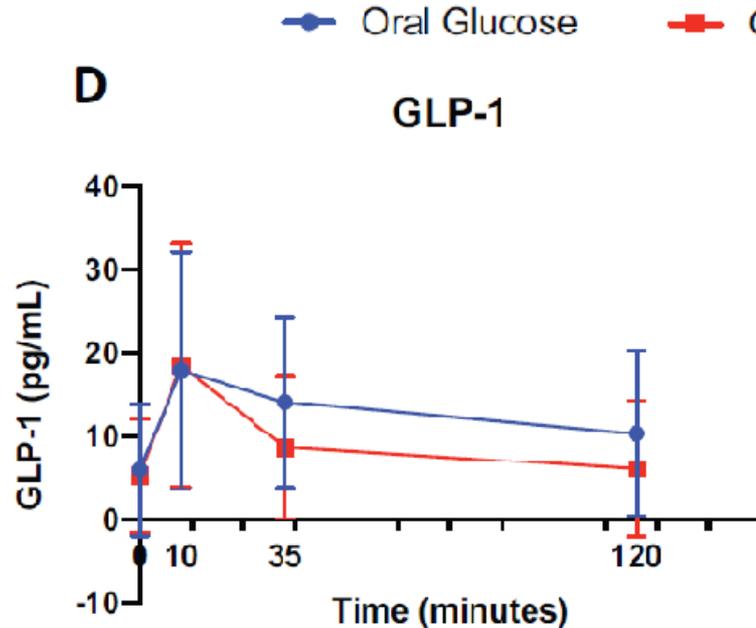
Need to consider the **correlated data** (each subject provides outcomes for each of the three drinks).

Paired t-tests

Mixed effects regression model: Repeated measures analysis (extension of **paired** testing), with **order of the interventions** considered.

Clinical Trial Designs: Crossover

Result example: GLP-1 (hormone that triggers post-meal satiety)



Oral sucrose showed blunted GLP-1 response compared to oral glucose ($p < 0.001$)

Clinical Trial Designs: Cluster Randomized

Cluster randomized: The unit of randomization is a group of persons, rather than a single individual.

Common in testing complex interventions in primary care, health promotion, community/public health settings

Advantage: Avoids **contamination of intervention effects** with cluster-related effects (e.g., private more affluent hospitals have more resources for programs to prevent hospital-acquired pneumonia)

Disadvantages: (1) Requires more subjects to account for the correlation of the outcomes between persons in the same cluster. (2) Blinding is usually not possible

Clinical Trial Designs: Cluster Randomized

Pilot study to examine the effects of indoor daylight exposure on depression and other neuropsychiatric symptoms in people living with dementia in long-term care communities

Clin Interv Aging 2018;13:1071-1077. PMID: 29881263.



Clinical Trial Designs: Cluster Randomized

Population: Persons with dementia diagnosis residing in memory care facility

Randomization occurred at the facility level: 4 facilities to intervention, 4 to comparator.

Intervention: 12-week daily exposure to morning natural daylight in the facility

Comparator: Standard activity locations

Outcomes: Validated measures of depression and behavior

Clinical Trial Designs: Cluster Randomized

What statistics did we use to evaluate the efficacy of this intervention?

We needed to account for the clustering within facility (i.e., persons within facility may have similar outcomes simply due to facility factors, or the types of persons residing within particular facilities): **Mixed effects linear regression models** to estimate intervention group differences

Clinical Trial Designs: Cluster Randomized

Table 2 Baseline and endpoint comparisons by treatment

	Daylight	Control	p-value between treatments
N	46	31	
Baseline outcomes			
NPI-NH	16.2 (3.2)	16.1 (2.7)	0.97
CSDD	4.2 (1.9)	3.9 (1.9)	0.93
Endpoint outcomes			
NPI-NH	13.4 (4.2)	19.1 (4.4)	0.35
CSDD	2.7 (2.1)	5.3 (2.0)	0.39
Change (endpoint minus baseline) outcomes			
NPI-NH	-2.8 (2.9)	3.1 (3.2)	0.17
p-value within group	0.33	0.33	
CSDD	-2.0 (0.9)	1.5 (1.0)	0.01
p-value within group	0.025	0.13	
mLux _{AVG}	159.3 (13.8)	42.3 (3.1)	<0.0001

Clinical Trial Designs: Equivalence or Non-Inferiority Trials

Equivalence or non-inferiority trials: Rather than hypothesizing that one intervention group will show superior outcomes to another group, these trials hypothesize that the **new** intervention groups will demonstrate trial outcomes that are **the same as (equivalency)** or **no worse than (non-inferiority)** a currently standard intervention.

These trials are appropriate if there is a standard and effective intervention available. The investigators want to show that the new intervention (which may be cheaper, have less side effects, or be more readily applied in the population) works just as well as the standard intervention.

Advantage: Allows the testing of the efficacy of interventions that may increase the adoption of an effective intervention in the population.

Disadvantages: Requires the naming of an “**equivalence or non-inferiority margin**” that should have a clear clinical rationale. The margin: **how different** (equivalence) or **how much worse** (non-inferiority) can the new intervention be, and still be comfortable stating that the new intervention is no different (or no worse) than the standard intervention? Can require large sample sizes.

Clinical Trial Designs: Equivalence or Non-Inferiority Trials

Effectiveness of Online vs In-Person Care for Adults With Psoriasis A Randomized Clinical Trial

April W. Armstrong, MD, MPH; Cindy J. Chambers, MD, MAS, MPH; Emanuel Maverakis, MD; Michelle Y. Cheng, MD; Cory A. Dunnick, MD; Mary-Margaret Chren, MD;

TRIAL REGISTRATION ClinicalTrials.gov Identifier: [NCT02358135](https://clinicaltrials.gov/ct2/show/study/NCT02358135)

JAMA Network Open. 2018;1(6):e183062. doi:[10.1001/jamanetworkopen.2018.3062](https://doi.org/10.1001/jamanetworkopen.2018.3062)

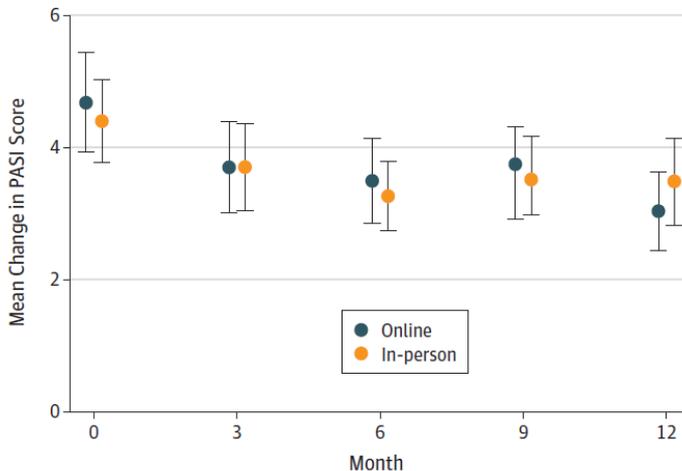


Clinical Trial Designs: Equivalence or Non-Inferiority Trials

In this study, we evaluated an innovative, collaborative connected-health model in which patients and PCPs can access dermatologists online directly and asynchronously via a pragmatic trial. The primary aim of this pragmatic trial was to determine whether this online, collaborative connected-health model results in equivalent improvements in psoriasis disease severity compared with in-person care.

Clinical Trial Designs: Equivalence or Non-Inferiority Trials

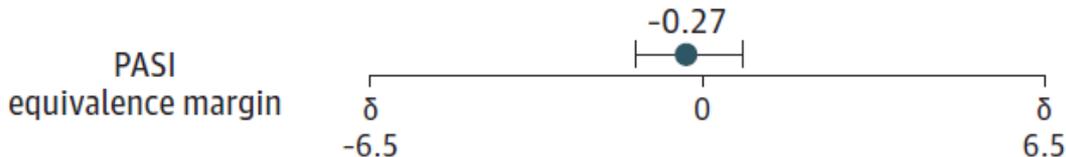
A PASI score



Differences in psoriasis score (PASI) by treatment group

Mean group difference (-0.27) and confidence limits are within the bounds of the pre-specified equivalence margin (± 6.5)

D Summary data on equivalency evaluation



Statistical Analysis Plan

- Ties directly back to your research question, aims, and hypotheses

You MUST be able to answer the following questions

- What are my **dependent** (outcome) variables? How are they measured? What type of variable are they? Am I measuring them just once (cross-sectional) or multiple times (longitudinal, repeated measures)?
- What are my **independent** variables (experimental interventions, control variables)? How are they measured? What types of variables are they?
- Given the above, what are appropriate methods of analysis?

Testing Differences Among Groups

- Group comparisons by data type
- For **categorical** data, groups are compared with **chi-square tests** (testing if the proportions of subjects in categories differs between groups)
- For **continuous** data, groups are compared with parametric or non-parametric tests (depending on normality of data)
 - **Parametric** (normal outcome data): t-tests (2 groups), analysis of variance (>2 groups)
 - **Non-parametric** (non-normal): Wilcoxon rank sum

Testing Differences Among Groups

- Group comparisons for **matched/repeated** measures
- For **categorical** data, groups are compared with **chi-square tests** that incorporate the matching (McNemar's test for proportions)
- For **continuous** data, groups are compared with parametric or non-parametric tests, incorporating the matched data
 - **Parametric** (normal outcome data): paired t-tests (2 groups), repeated measures analysis of variance (>2 groups)
 - **Non-parametric** (non-normal): signed rank test

Survival Time Data

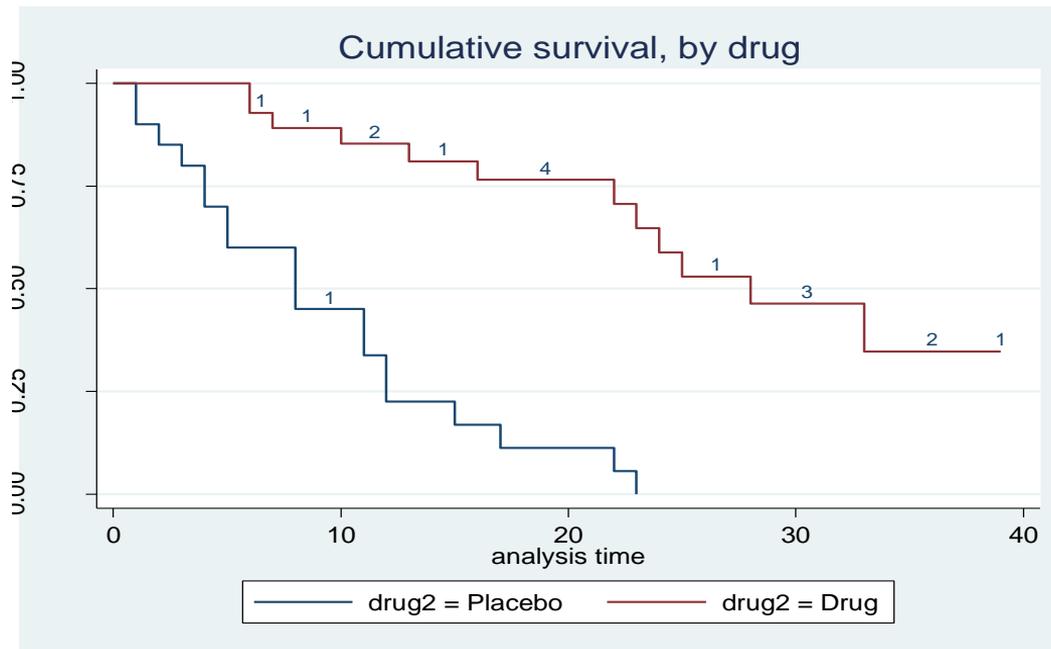
- Survival time data: Contains two components
 - 1) If the subject had the event (did the subject die?)
 - 2) The last time the subject was observed

E.g., Subject died at age 82

Subject was alive at age 53 (last age observed on-study)

Subject died 2.5 years after lung cancer diagnosis

Lifetable Group Comparisons: Graph Survival Over Follow-up



Linear Regression (and Other Regression Models)

- We often want to compare groups on our trial outcome variable, adjusting for other variables.

For example, in a clinical trial we might want to compare a cognitive therapy to a medication group on a depression measure following 6 months of treatment. However, we want to adjust for each person's level of depression when they first started the trial. Regression models will allow us to do this

- Linear association model with a continuous outcome (dependent) variable, multiple independent variables

$$Y = a + b_1X_1 + b_2X_2 + \dots$$

e.g. Depression score (6 months) = $a + b_1(\text{Treatment group}) + b_2(\text{Baseline depression score})$

- Coefficient of determination (R^2) is the proportion of variation in Y that can be explained by all of the X independent variables

Other Regression Models

- There are many types of such regression models. The type of regression model used depends on what type of data the **outcome (dependent)** variable is. **You must select the correct regression approach to match your dependent variable!**
- **Continuous** outcome: **linear** regression – do independent (X) variables relate to **the levels** of Y? (e.g., Depression score at 6 months)
- **Dichotomous** outcome: **logistic** regression – do independent (X) variables relate to **the probability** that $Y=1$ (vs $Y=0$)? (e.g., Has a participant reported suicidality within 6 months of starting a trial treatment)

Other Regression Models

- **Ordinal** categorical outcome: **ordinal** logistic regression – do independent (X) variables relate to **the probability** that Y = higher compared to lower level? (e.g., After 6 months of trial treatment, has a participant's depression score decreased, not change, or increased)
- **Nominal** outcome (not ordered): **multinomial** logistic regression – do independent (X) variables relate to the probability that Y = category 1 (vs category 2, 3, etc.)? (e.g., After 6 months of trial treatment, is a trial participant employed, on medical disability, or unemployed)

Other Regression Models

- **Count** outcome: **Poisson or negative binomial** regression – do independent (X) variables relate to the count Y (e.g., How many emergency/urgent care visits did a participant make in the 6 months following start of trial treatment)
- **Survival** outcome: Cox (proportional hazards) or other “survival” regression – do independent (X) variables relate to the event rate? (e.g., Does the rate of accidental injuries differ in persons randomized to a cognitive therapy versus medication for major depression)

Gender, Race, and Ethnicity in Clinical Trials

Wendy Mack, PhD

Southern California Clinical and Translational Science Institute:

Biostatistics, Epidemiology and Research Design (BERD)

Objectives

- NIH, FDA policies on inclusion and reporting by sex, race, ethnicity
- Rationale for inclusion and reporting by sex, race, ethnicity
- Examples of heterogeneity of intervention effects by sex, race, ethnicity
- Implications for trial design and reporting



NIH Policies on Inclusion

NIH Policy and Guidelines on The Inclusion of Women and Minorities as Subjects in Clinical Research

As required by federal law ([42 USC 289a-2](#)) and NIH policy, applications that propose to involve human subjects must address the inclusion of women, minorities, and children in the proposed research.

The NIH Inclusion Policy and Guidelines are updated to provide additional guidance on reporting analyses of sex differences, race, and ethnicity differences by study intervention effects for all NIH-defined Phase III clinical trials. The valid analyses, or stratified results reporting, should be conducted for each primary outcome measure by sex/gender and by race and/or ethnicity. Annually, applicants/investigators must provide an update on their progress in meeting the NIH-funded objectives, including providing the number of individuals enrolled in research study, broken out by sex/gender, race, and ethnicity. Valid analysis reporting in ClinicalTrials.gov is required for Applicable NIH-defined Clinical Trials (ACTs).

Learn more at: https://grants.nih.gov/grants/funding/women_min/guidelines.htm

Read background information: <https://orwh.od.nih.gov/research/clinical-research-trials/nih-inclusion-policy/including-women-and-minorities-clinical>

Policy link for NIH Guide Notice on Valid Analysis: [NOT-OD-18-014](#)

NIH Policies on Inclusion: Key Points

- All NIH applications involving human subjects research must consider inclusion of women, minorities, and children
- Phase III clinical trials must provide valid analyses by sex, race and ethnicity. Those analyses must be reported in clinicaltrials.gov.
- If prior studies suggest that there may be heterogeneity of intervention effects by sex, race, ethnicity, then the trial should be designed (i.e., sufficient sample size) to test effects within relevant subgroups.
- Annual reporting of participant recruitment by sex, race, ethnicity

NIH Policies on Inclusion: Enrollment Table

Planned

Racial Categories	Ethnic Categories				Total
	Not Hispanic or Latino		Hispanic or Latino		
	Female	Male	Female	Male	
American Indian/Alaska Native	0	0	0	0	0
Asian	0	0	0	0	0
Native Hawaiian or Other Pacific Islander	0	0	0	0	0
Black or African American	0	0	0	0	0
White	0	0	0	0	0
More than One Race	0	0	0	0	0
Total	0	0	0	0	0

Policies Mirrored by FDA: For example...

Collection of Race and Ethnicity Data in Clinical Trials

**Guidance for Industry and Food and
Drug Administration Staff¹**

FDA Policy: Key Points

- Standardized collection of sex, race and ethnicity data for clinical trial submissions
- Consistent collection and reporting of subpopulation data (trials are presenting data on the same subgroups)
- Reporting requires tabulation of participant numbers by age, sex, race
- Expected to enroll participants by age, sex, race, ethnicity that reflect the demographics of the clinical condition under study
- E.g. Multi-regional clinical trials to address possible heterogeneity of treatment effects

FDA Policy: Key Points

Two-Question Format

In order to be consistent with OMB and other recommended best practices, FDA recommends using the two-question format for requesting race and ethnicity information, with the ethnicity question preceding the question about race. Example:

Question 1 (answer first): Do you consider yourself Hispanic/Latino or not Hispanic/Latino?

Question 2 (answer second): Which of the following five racial designations best describes you? More than one choice is acceptable.

NIH and FDA Policies

- Recommendations, not mandates
- NIH applications are limited by geography! (Numbers below are approximate)

	White	Black	Asian	Hispanic
Portland, OR	77%	6%	8%	9%
Los Angeles	50%	9%	14%	48%
Chicago	45%	32%	5%	28%
US	61%	13%	5%	18%

Rationale

- Incidence and survival of many diseases varies by sex, race, and ethnicity. Examples are diabetes, stroke and HIV/AIDS.

Disparities are related not only to genetic differences, but also lifestyle, environmental, and socioeconomic factors.

Addressing health disparities requires data collection and reporting in racial/ethnic populations

- Variable drug effects on persons based on sex, race, ethnicity (and age). Examples:

Whites more likely to have low levels of an enzyme (CYP2D6) that metabolizes certain antidepressants, antipsychotics and beta blockers.

Racial differences in skin structure and physiology can alter responses to dermatologic and topically-applied agents.

Rationale

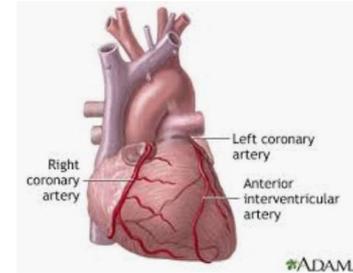
- Clinical trial participation is historically dominated by white (and highly educated) participants
- Historically some trials were limited by sex (e.g., cholesterol-lowering trials dominated by male participants)
- Reduction of health disparities requires that all demographic groups be represented in clinical trials
- Requires representation in early phases of intervention development and evaluation (e.g., effective drug dose may vary by sex, race)

Subgroup Reporting

- FDA Demographic Rule: Submission of safety and efficacy data by age, gender, race
- Population level PK studies to evaluate differences in safety, efficacy by gender and race/ethnicity

Sex and Clinical Trials

- Historically many trials were limited to men. Why?
Worry about hormonal fluctuations and effects on drug metabolism.
Worry about safety in pregnant women (excluding ALL women of child-bearing age)
Perceptions that certain conditions are “male” diseases.
- Example: Coronary Drug Project
One of the earliest RCTs for CHD prevention, conducted before the development and approval of cholesterol-lowering medications.
8431 men, aged 30-64 (young!) post-myocardial infarction.
Tested 5 medications known to alter blood cholesterol: Two doses of estrogen!
Both estrogen groups terminated early due to adverse effects.



Sex and Clinical Trials

Many adverse effects are more evident in females than males

Given simple differences in size, medical device trials are particularly prone to sex-biases and should include sufficient numbers of both sexes.

Sex vs. Gender in Clinical Trials

- Clinical trials generally use the two terms interchangeably, and in reporting guidelines

Standard case report forms have two choices: male/female

- More current guidelines encourage differentiation

Sex: biogenetic and physiologic differences distinguishing males and females

Gender: socially constructed roles by which society differentiates men and women

Proposed expansion of case report forms to include two-step question (sex at birth, current gender identity). May be very relevant to particular drug evaluations.

Sex vs. Gender in Clinical Trials

Clayton JA,
Tannenbaum C.
Reporting sex, gender,
or both in clinical
research?
JAMA 2016;316:1863-
1864

Table. Suggested Approach for Reporting Demographic Characteristics of Study Participants and Outcome by Sex and Gender (N = 59)

Demographic Characteristics	
Total No.	59
Age range, y	18-90
Sex, No. ^a	
Male participant	27
Female participant	32
Gender, No. ^b	
Men	26
Women	33
Outcome, No. (%) ^c	50 (85)
Males	20 (40)
Females	30 (60)
Outcome, No. (%) ^d	
Male	20 (74)
Female	30 (94)

^a Ascertained by genotyping of blood sample.

^b Ascertained by self-report.

^c The number (%) occurring in males and females of the total outcomes (n = 50).

^d Number (%) of outcomes occurring within the subgroups of males (20/27) and females (30/32).

Sex and Gender in Clinical Trials

- Both sex and gender influence how/what particular treatment a person selects, how they adhere to it, and how they metabolize a drug.
- As appropriate, trials should include equal numbers of men and women and analyze report results separately for each. Why?

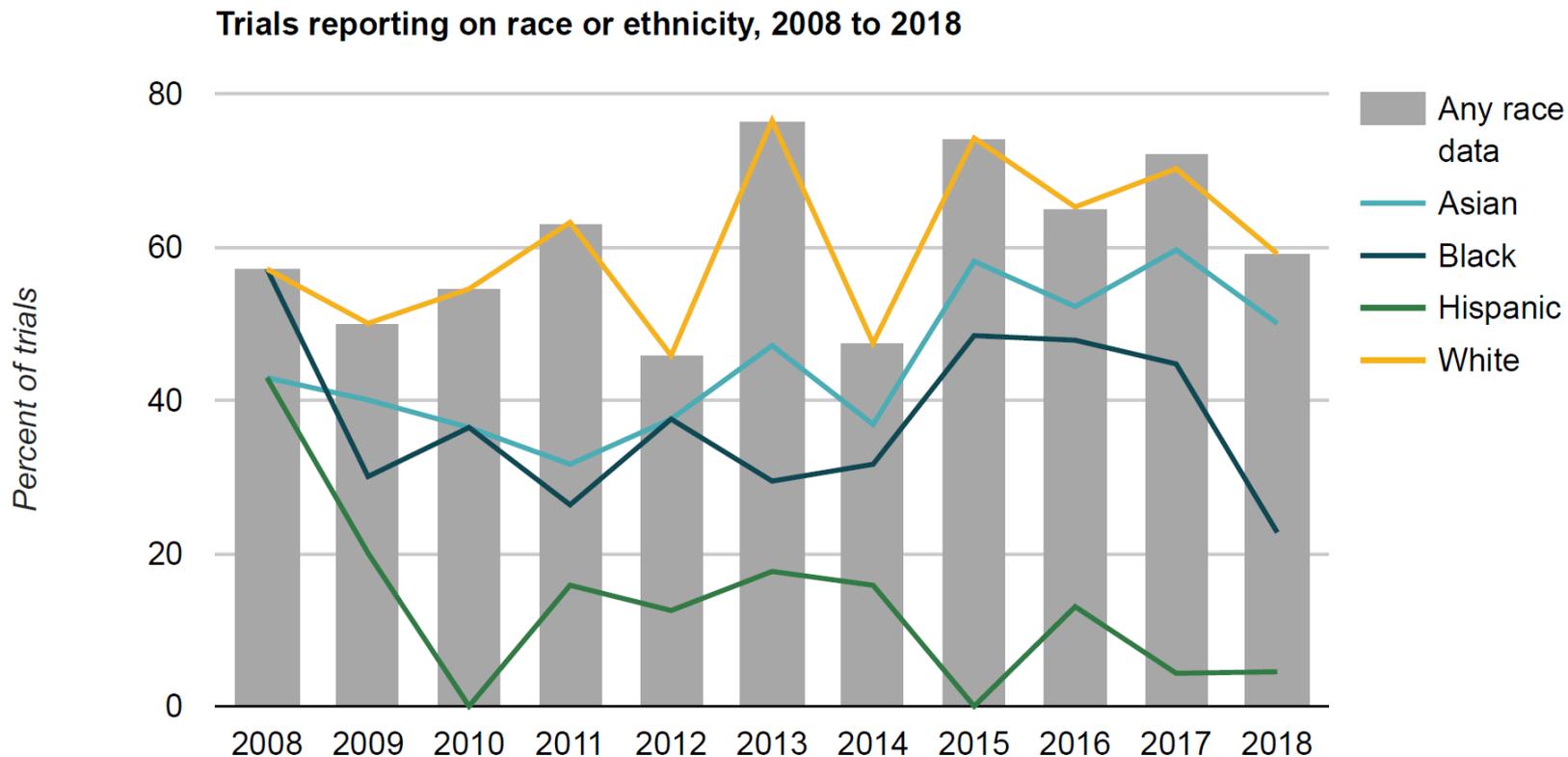
The total sample analysis may mask sex/gender differences in efficacy or safety.

Although there usually is not sufficient sample size for adequate statistical power to test efficacy by sex/gender, reporting the sex-specific results can be used in later meta-analyses that can combine results over multiple trials to achieve adequate power for sex-specific analyses.

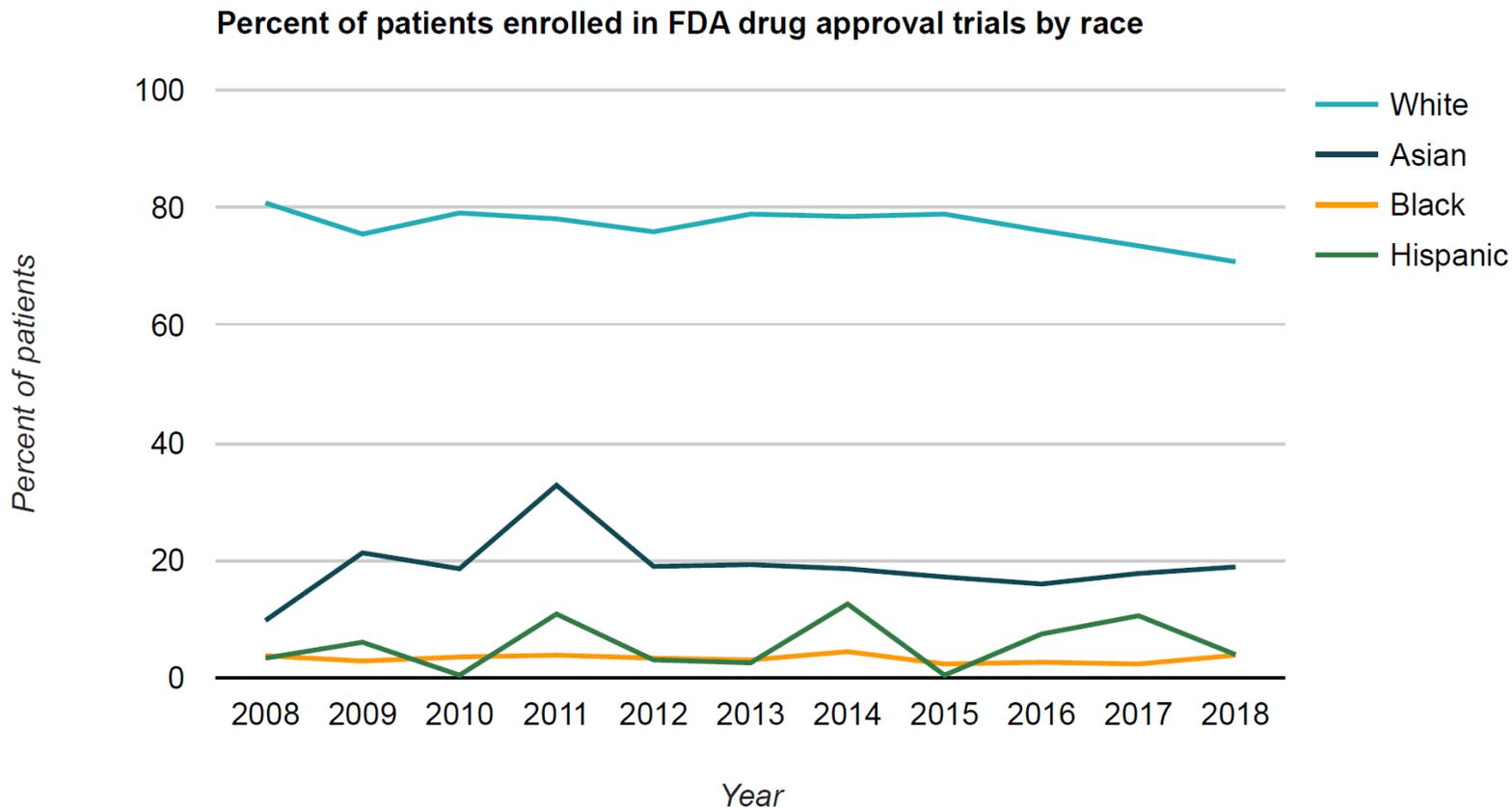
Race (Mis)reporting and Participation in Cancer Clinical Trials

- Loree et al; Disparity of race reporting and representation in clinical trials leading to cancer drug approvals from 2008 to 2018. JAMA Oncology 2019; 5(10). Reviewed 230 clinical trials involving 112,293 participants.
- 145 reported on one race; only 18 reported 4 groups (white, Asian, black, Hispanic)
- Black patients represented in 3% of trial participants.
- Hispanic patients represented in 6% of trial participants.
- No major changes in representation over the 11-year period.

Race (Mis)reporting in Cancer Clinical Trials

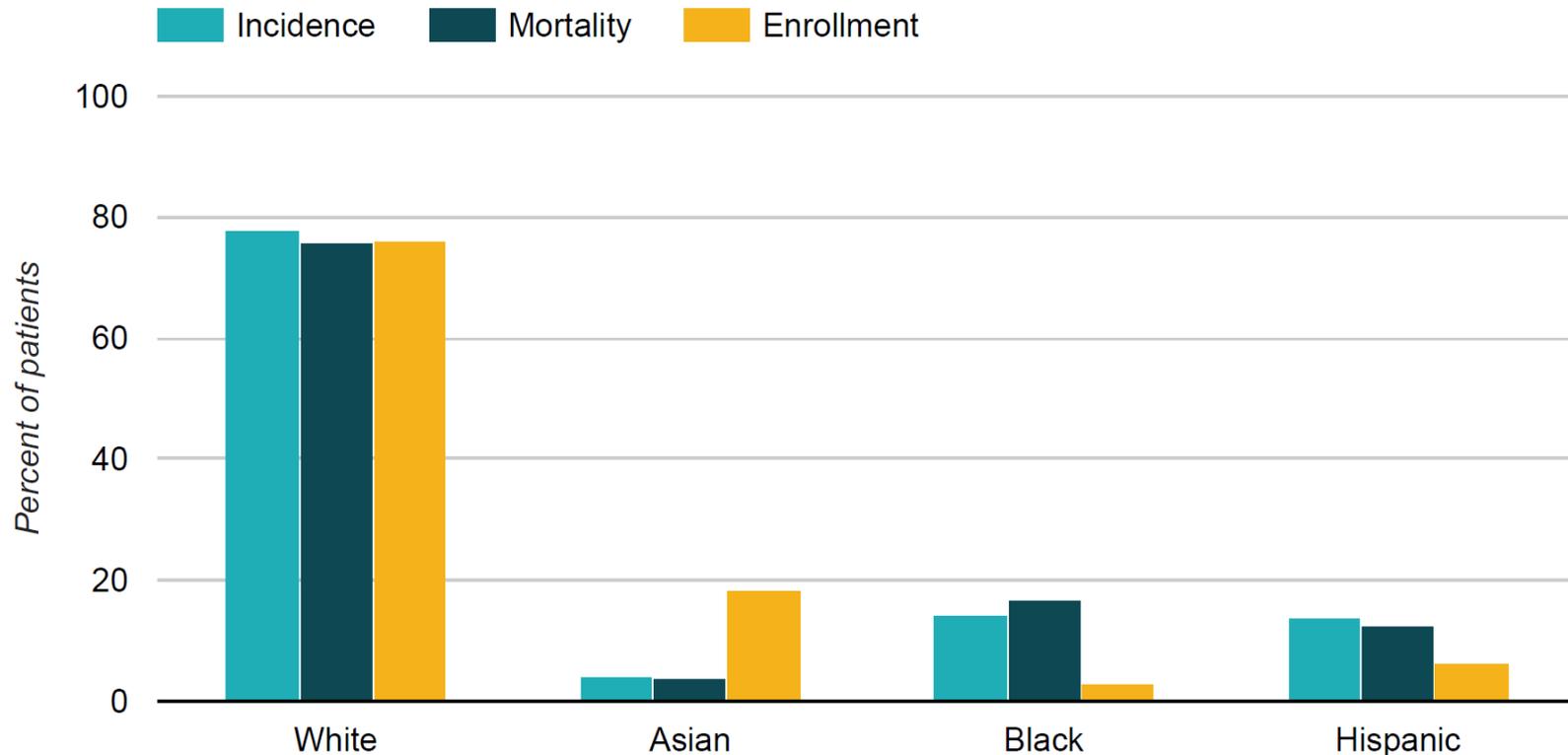


Race (Mis)reporting in Cancer Clinical Trials



Race (Mis)reporting in Cancer Clinical Trials

Incidence and mortality rates of patients with cancer compared to trial enrollment



Other Examples of Race/Ethnic Differences in Clinical Trials

African-American Heart Failure Trial (A-HeFT)

Race-focused RCT

Causes of heart failure differ in white and black populations.

Black population shows less effective blood pressure response than white population to ACE inhibitors and beta blockers (standard treatments for heart failure)

Post hoc analyses of heart failure trial data suggested black/white differences that could be tested. Notably, **many heart failure trials had not included sufficient numbers of women and minorities to even conduct subgroup analyses.**

First heart failure trial focused in black population. Added an isosorbide dinitrate/hydralazine combination to standard therapy. 43% improvement in survival, 33% reduction in heart failure hospitalizations.

Other Examples of Race/Ethnic Differences in Clinical Trials

African-American Study of Kidney Disease and Hypertension (AASK)

Race-focused RCT

Hypertension major cause of end-stage renal disease. Black hypertensive population has much higher risk of progression to dialysis-dependent renal disease than white hypertensive population.

Prior RCTs had not recruited sufficient numbers of hypertensive black population

Objective: select optimal antihypertensive regimen to reduce progression of renal disease

Demonstrated efficacy of ACE inhibitors for this purpose, identified optimal level of blood pressure reduction for renal protection

Other Examples of Race/Ethnic Differences in Clinical Trials

Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack (ALLHAT)

RCT specifically designed to evaluate ethnic differences in treatment response

Compared newer blood pressure reducing agents to diuretics in reducing CVD risk

Recruitment goal to obtain sufficient numbers of black participants to analyze efficacy by race (recruited >15000)

Overall and in non-black participants, diuretics at least as effective as newer agents in reducing blood pressure and CVD outcomes.

In black participants, efficacy of ACE inhibitors was dramatically worse than diuretics in terms of blood pressure reduction and CVD outcomes.

Notably, ALLHAT findings occurred >25 years after introduction of ACE inhibitors.

Implications for Trial Design and Analyses

- Addressing and understanding health disparities requires that clinical trials enroll sufficient numbers of participants in sex/race/ethnic subgroups to look for differences in efficacy, safety, pathophysiology (e.g., ALLHAT).
- Inclusion of adequate numbers by sex and race/ethnicity and analysis and reporting of results by these subgroups will allow post hoc trial analyses and meta-analyses to identify possible differences in efficacy and safety.
- When differences are found, follow-up trials should focus on particular population subgroups (e.g., A-HeFT, AASK)

Implications for Trial Design and Analyses

- Clinical trials in the past have recruited homogeneous groups (to reduce variability) and applied results to other populations. The paucity of subgroup-focused and adequately powered trials to assess subgroup differences contributes to ongoing health disparities.
- Clinical trials should also recruit participants to whom the intervention will be applied. For example, there is a tendency to exclude older individuals with co-morbid conditions, the very population to which most of the tested medications will be applied.

Implications for Trial Design and Analyses

- RECRUITMENT!

Identify and respond to barriers to recruitment and retention.

Example: Research shows minority patients engage physicians of like race. Inclusion of female and race/ethnic minority physicians as co-investigators in trials.

- Move clinical trials out of institutions and into communities, for greater use of community-based participatory research.
- Multi-site trials: geographic site-selection based on race/ethnic distribution of populations and disease maps
- Incorporation of technology (e.g., web-based pragmatic trials)

Subgroup Statistical Analyses

- If a known gender/race/ethnicity effect of an intervention exists, then a trial design should include sufficient numbers to conduct valid (i.e., **sufficient sample size**) statistical analyses **within each relevant group**. For example, a separate analysis and reporting of treatment efficacy and safety in males and females.
- If there are not known gender/race/ethnicity heterogeneity of effects, subgroup analyses can still be conducted to compare the magnitude of effects over gender/race/ethnic subgroups. However, there should be sufficient sample size in subgroups to obtain **reasonably precise estimates of effects**.
- To formally test for subgroup differences in effects, statistical models will include a product interaction term (e.g., treatment x gender). Such interaction analyses typically require large sample sizes.

Thank you!



Regulatory Science Symposium

Make Informed Decisions: Key Statistical Principles to Clinical Trial Design

Wrap-Up!

Eunjoo Pacifici, PharmD, PhD

Chair and Associate Professor, Regulatory and Quality Sciences
Associate Director, DK Kim International Center for Regulatory Science

Resource

Common Statistical Terms Used in Clinical Trials

Bias

Systematic tendency of any factors associated with the design, conduct, analysis and evaluation of the results of a clinical trial to make the estimate of a treatment effect deviate from its true value

Equivalence Trial

Trial with primary objective of showing that the response to two or more treatments differs by an amount which is clinically unimportant

Full Analysis Set

Set of subjects that is as close as possible to the ideal implied by the intention-to-treat principle

Generalizability

Extent to which the findings of a clinical trial can be reliably extrapolated from the subjects who participated in the trial to a broader patient population and a broader range of clinical settings

Independent Data Monitoring Committee

May be established by the sponsor to assess at intervals the progress of a clinical trial, safety data, and critical efficacy endpoints, and to recommend whether to continue, modify, or stop a trial

Intention-To-Treat Principle

Principle that asserts that the effect of a treatment policy can be best assessed by evaluating on the basis of the intention to treat a subject rather than the actual treatment given.

Per Protocol Set

Set of data generated by the subset of subjects who complied with the protocol sufficiently to ensure that these data would be likely to exhibit the effects of treatment, according to the underlying scientific model

Presented by the USC School of Pharmacy International Center
for Regulatory Science and the Southern California Clinical and
Translational Science Institute

This certifies that

Before the end of today's symposium, you will
receive a link to take the program evaluation.

Follow this link to the Survey:

[Take the Survey](#)

Please complete the program evaluation to receive a
certificate of completion by Friday, February 18, 2022.


Eunjoo Pacifici, PharmD, PhD
Director
International Center for Regulatory Science

USC School of Pharmacy
DK Kim International Center for Regulatory Science


Thomas A. Buchanan, MD
Director
Southern California Clinical and
Translational Science Institute



Thank You!



www.sc-ctsi.org

Phone: (323) 442-4032

Email: info@sc-ctsi.org

Twitter: @SoCalCTSI

USC School of Pharmacy

DK Kim International Center for Regulatory Science

regulatory.usc.edu

Phone: (323) 442-3521

Email: regsci@usc.edu

Facebook: @RegSci