Overview of Clinical Study Design

Laura Lee Johnson, Ph.D.
Statistician
National Center for Complementary and Alternative Medicine
US National Institutes of Health
Get to Know Each Other Online: Have you ever....

1. Taken a class or have a degree in
   i. Biostatistics
   ii. Epidemiology
   iii. Research design
2. Used logistic regression?
3. Designed some type of public health research project outside of a class environment?
4. Run a study?
5. Done a data analysis outside of a class environment?
Get to Know Each Other Online: Have you ever....

6. Written a clinical research protocol?
7. Read a clinical research protocol?
8. Read a clinical journal article?
Overview

• Many different people take the course
  – Some first introduction, some advanced
• Tricks, tips, and concepts
• General objectives
  – Better consumer of medical and scientific literature
  – Enhance conversations inside research team [with study statisticians and epidemiologists]
Easy to Write

- The study will use a randomized, double-blind, controlled parallel arm design and an intent to treat analysis.
Not Easy

• To implement and maintain the integrity of
  – Randomization
  – Blinding/masking
  – Multiple study arms
  – Data collection so can do reasonable intent to treat analyses
Tonight’s Objectives

• Identify study designs used in clinical studies, epidemiology and public health research
• Discuss masking/blinding, interventions, comparison groups

• Chapters 19, 29, 3rd Edition
Outline

➢ Where to start

• Taxonomy and examples
  – Some general vocabulary
  – Observational and interventional studies

• General good study design
  – Intervention and dose
  – Comparisons

• Conclusions

• Confounding and effect modification
Cervical Cancer

• Cultural
• Access
• Vinegar
• Diagnostic biomarker
  – High negative predictive value
  – Low cost
  – Fast
• Time table for treatment
Other Examples

• Cardiovascular disease
  – Weight
  – Hypertension
• Infectious diseases
• Data collection/mode of administration
  – Mobile health
• Combination of interventions
What is the question of interest?

- Interpreting work into new population
- Making decision about individual case
- Looking at changing a population
  – Diabetes management
- Differences between groups in a study
- Biomarker development? What kind of biomarker?
- Develop a new outcome?
- Level of evidence
- Evaluation of evidence
Analysis Follows Design

Questions → Hypotheses → Experimental Design → Samples → Data → Analyses → Conclusions

• Take all of your design information to a statistician early and often
  – Guidance
  – Assumptions
How a Statistician Sees a Research Study

• Everything impacts the statistical analysis
Outline

✔ Where to start
  ➢ Taxonomy and examples
    ➢ Some general vocabulary
    ➢ Observational and interventional studies
  • General good study design
    – Intervention and dose
    – Comparisons
  • Conclusions
  • Confounding and effect modification
Study Design Taxonomy

- Intervention vs. Observational
- Longitudinal vs. Cross-sectional
- Prospective vs. Retrospective
- Blinded/Masked or Not Blinded/Masked
  - Single-blind, Double blind, Unblinded
- Randomized vs. Non-Randomized
Two Types of Research Studies

• Observational
  – Goal is to observe and collect data on characteristics of interest without influencing the participant, environment or disease course

• Experimental
  – Researcher deliberately influences course of events and investigates effects of an intervention on a carefully selected population of subjects
  – Experimental studies done on human subjects are referred to as clinical trials or clinical studies
Ideal Study - Gold Standard

- Treatment / control
- Parallel groups
- Superiority
- Prospective
- Double blind / masked
- Randomized
• Non-invasive versus invasive respiratory support in preterm infants at birth: systematic review and meta-analysis

• Effectiveness of telemonitoring integrated into existing clinical services on hospital admission for exacerbation of chronic obstructive pulmonary disease: researcher blind, multicentre, randomised controlled trial
• Risk of moderate to advanced kidney disease in patients with psoriasis: population based cohort study

• Prevalent abnormal prion protein in human appendixes after bovine spongiform encephalopathy epizootic: large scale survey
Observational Studies

- Case Reports
- Case Series
- Cross-Sectional or Prevalence Surveys
- Case-Control Study
- Cohort Study (longitudinal)
- Natural History Studies
- Ecological Studies (data on population rather than individual level)
Intervention Based Research Spectrum

- Epidemiology
- Quasi-experimental
- Pre-clinical studies
- Phase 0
- Phase I
- Early/Late Phase II
- Phase III
- Phase IV
- Dissemination and Implementation
- Comparative or Cost Effectiveness
Distinguish

“Observational studies are often analyzed as if they had resulted from a controlled study, and yet the tacit assumption of randomness can be crucial for the validity of inference.”

Quasi Experimental, One/Single Arm, or Non-Randomized Experimental Studies

• No control group
  – Early in investigation
• Concurrent control “group”
  – Treatment assignment not by randomization
• Historically controlled
  – Missing data
  – Poor data
  – Non-comparability of groups
**Non-Randomized**

- Can ONLY show Association
- You will never know all possible confounders!

**Randomized**

- Can show Association AND Causality
- Well done non-adaptive randomization → unknown confounders should not create problems
Types of Randomized Studies

- Parallel Group – classic
- Sequential Trials – physical sciences
- Group Sequential trials – classic
- Cross-over – intervention washout
- Factorial Designs – independence
- Adaptive Designs – gaining popularity
- Enriched Enrollment – regression to the mean
- Cluster Randomized Designs
Parallel Group

• Randomize patients to one of X treatments
• Response
  – Measure at end of study
  – Change (delta) or % change from baseline
  – Repeated measures
  – Function of multiple measures
Variations on Parallel Group Designs

• Dose titration (with multiple study arms)
  – Titrate to the maximum tolerated dose within a subject
• Dose Escalation Studies (with a control arm that is simultaneously randomized to)
• Not all dose escalation and dose titration studies are randomized
Sequential Trials

• Not a fixed sample size/period
• Terminates when
  – One treatment shows a clear superiority or
  – It is highly unlikely any important difference will be seen
• Special statistical design methods
Group Sequential Trials

• Popular
• Analyze data after certain proportions of results available
• Early stopping
  – If one treatment clearly superior
  – Futility
  – Adverse events
• Careful planning and statistical design
Preventing Mother-Infant HIV Transmission (D.O. Dixon, NIAID)

- Zidovudine able to slow progression of HIV in adults with advanced disease
- AIDS Clinical Trials Group Protocol 076 designed to assess both safety and efficacy of Zidovudine in preventing transmission of HIV from infected (not advanced) women to their babies
Preventing Mother-Infant HIV Transmission

- Powered (80%) to detect a 33% reduction of transmission rate (through 78 weeks of baby’s life) relative to projected rate of 30%
- Target N was 748; began April 1991
- Projected accrual to take at least 5 years and 15% dropouts
Preventing Mother-Infant HIV Transmission

- DSMB met twice a year to monitor safety
- Efficacy reviews planned after each 1/3 of projected infant infections
- 1st efficacy review took place in February 1994, based on mothers enrolled up to December 1993 and their babies
At First Interim Analysis
(1/3 of projected infant infections)

P = 0.00006

DOD/BRB/NIAID
Crossover Trial

- Cross over example
  - 2 treatments
  - 2 period crossover
- Use each patient as own control
- Must eliminate carryover effects
  - Need sufficient washout period
  - Not always known what length of time is needed
  - Not everything can wash out
Women’s Alcohol Study

JNCI 2001

• Three 8-week dietary periods
  – 30 g alcohol/day
  – 15 g alcohol/day
  – 0 g alcohol/day [alcohol free beverage]

• Order of assignment to 3 alcohol levels was random

• Washout periods

• Double blind
Factorial Design

- Each level of a factor (treatment or condition) occurs with every level of every other factor.
- Selenomethionine (Se) and Celecoxib (C).

Gastroenterology 2002; 122:A71

<table>
<thead>
<tr>
<th>SE Placebo</th>
<th>Selenium</th>
</tr>
</thead>
<tbody>
<tr>
<td>C Placebo</td>
<td>C Placebo</td>
</tr>
<tr>
<td>SE Placebo</td>
<td>Selenium</td>
</tr>
<tr>
<td>Celecoxib</td>
<td>Celecoxib</td>
</tr>
</tbody>
</table>
Factorial Design

• Factor 1: Selenium
  – Yes, No
• Factor 2: Celecoxib
  – Yes, No
<table>
<thead>
<tr>
<th></th>
<th>Se (Placebo)</th>
<th>Se (Real)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celecoxib (Placebo)</td>
<td>Se Placebo C Placebo</td>
<td>Selenium C Placebo</td>
</tr>
<tr>
<td>Celecoxib (Real)</td>
<td>Se Placebo Celecoxib</td>
<td>Selenium Celecoxib</td>
</tr>
</tbody>
</table>
## Factorial Design

<table>
<thead>
<tr>
<th></th>
<th>Se (Placebo)</th>
<th>Se (Real)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celecoxib (Placebo)</td>
<td>Se Placebo C Placebo</td>
<td>Selenium C Placebo</td>
</tr>
<tr>
<td>Celecoxib (Real)</td>
<td>Se Placebo Celecoxib</td>
<td>Selenium Celecoxib</td>
</tr>
</tbody>
</table>
Factorial Design

- Power for the interaction or not?
- Is this a 4 arm study?
- 2-2 arm studies?

<table>
<thead>
<tr>
<th></th>
<th>selenium</th>
<th>C placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE placebo</td>
<td>selenium</td>
<td>C placebo</td>
</tr>
<tr>
<td>C placebo</td>
<td>selenium</td>
<td>C placebo</td>
</tr>
<tr>
<td>SE placebo</td>
<td>celecoxib</td>
<td>Celecoxib</td>
</tr>
<tr>
<td>Celecoxib</td>
<td>selenium</td>
<td>Celecoxib</td>
</tr>
</tbody>
</table>
## MsFLASH Factorial Design

<table>
<thead>
<tr>
<th></th>
<th>Yoga once a week, daily home practice</th>
<th>Aerobic exercise 3 times a week</th>
<th>Usual activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omega 3 supplement 615 mg gel capsule daily</td>
<td>15%</td>
<td>15%</td>
<td>20%</td>
</tr>
<tr>
<td>Placebo gel capsule</td>
<td>15%</td>
<td>15%</td>
<td>20%</td>
</tr>
</tbody>
</table>

*Menopause, April 2014*
Incomplete/Partial/Fractional Factorial Trial

- Nutritional Intervention Trial (NIT)
- 4x4 incomplete factorial
- A,B,C,D
- Did not look at all possible interactions
  - Not of interest (at the time)
  - Sample size prohibitive
Adaptive Designs

- Gaining popularity
- 2-8+ arms
- Dose ranging (perhaps)
- Smaller overall sample size (potentially)
- Run-in then analyze data continuously or at fixed points
Adaptive Designs

- Some act like group sequential designs
- Close an arm early
- Re-estimate sample size based on a nuisance parameter (variance)

- Big negative: any time a decision to continue is made, information is provided to the study investigators, public, investors....
Enriched Enrollment Designs

- Sometimes variant of crossover or n-of-1 study design
- Identify potential ‘responders’
- Enter ‘responders’ on to 2\textsuperscript{nd} prospective comparison study
- Results not generalizable to entire patient population
- Regression to the mean
  - Average 10 hot flashes a day for 14 days, and 3 on study in the placebo arm
Group or Cluster Randomized Studies

• Unit of randomization is not the individual
  – School
  – Community
  – Clinic
• Change eating patterns to impact obesity and other health outcomes
  – Altering supermarket environments
  – School based nutrition programs
• Charges for bed nets and impact on infant malaria cases
Outline

✔ Where to start
✔ Taxonomy and examples
  ✔ Observational and interventional studies
  ➢ General good study design
    ➢ What is the question
    ➢ Comparisons
• Conclusions
• Confounding and effect modification
How to Start Designing

- Study aims, background, rationale
- Endpoints or outcome variables, other assessments
  - Specific variables, how measured, specific & sensitive to changes expected, reliability & validity of measure
- Inclusion/Exclusion criteria
How to Start Designing

• Accrual plan and preparatory tasks
  – Timeline for overall study
  – Timeline for individual study participant
• Treatments? Participant implications?
  – Product, dose, quality, administration, and reproducibility of interventions (including training)
What is a Dose?

- Number of sessions/pills/treatments/social media attempts
- Frequency
- Length of sessions (each, total treatment time)
- Amount of practice
- Leader
  - Contact time
  - Who that person is
- Many different combinations
Intervention Definition

- What is it; can someone replicate it?
- Dose; choices based on what?
- Does it interfere with patient management
- Generalizability often lost in quest for specificity
- Specify criteria for withdrawal from study or deviation from protocol or definitions
- List concurrent medications, procedures, etc. that are prohibited, permitted, must be recorded in certain way
Practitioner Impact

- False negatives and positives
  - Details of protocol
- Examples
  - Massage for low back pain
  - Surgery
  - Prevention course
Study Analysis Population

• Mechanistic, proof of concept
  – Throw best at it
  – Per protocol
• General use
  – Like intent-to-treat analysis

• What is the question?
Intent-to-Treat versus Completers “How Many in the Data Analysis”

• ITT = Intent-To-Treat analysis
  – Include all randomized (if randomized study)
  – Assume all study participants
    • Adhered to study regime assigned
    • Completed the study
• MITT = Modified ITT analysis
  – ITT, but only include people who start intervention they are assigned to
• Completers or Adherers analysis
  – Only the well behaved
Vocabulary

• Arm = Sample = Group
• Demonstrate superiority
  – Detect *difference* between treatments, study arms, groups
• Demonstrate equally effective
  – Equivalence trial
  – Sample size required to demonstrate equivalence larger than required to demonstrate a difference
• Demonstrate non-inferiority
  – Lots of issues
Superiority vs. Equivalence
Non-Inferiority
Comparison Groups

- Experimental intervention vs control
- Case vs control
- Exposed vs unexposed
- Various levels of exposure
- Men vs women
- Old vs young
- BMI $>25$ vs $\leq 25$
- Usual or standard care or practice
- History; pre-post
Placebo, Standard of Care, and Attention Controls

- Experimental treatment
  - Supportive care
  - Current treatment
- Yoga
  - Exercise or stretching
  - Cooking classes
  - Book club
  - Nothing
Control Usually Costs Money = Larger Sample Sizes

• Control everything except smallest element of intervention that you want to test
  – Be careful it is not too small a difference
• Consequences if study has
  – More control imposed
  – Less sensitive or precise outcome measures
• Plan accordingly
What is the control group?

- Placebo
- Most widely accepted treatment
- Standard treatment
- Most accepted prevention intervention
  - Condoms and HIV?
- Usual care
- Accepted means of detection or diagnostic test
- Non-diseased population
Randomized Study Control Groups

- Ethical
  - Control intervention itself
  - Assigning ANYONE meeting study criteria to ANY study group
- If “standard of care” (SOC) evidence: is really standard practiced usual care in that format?
- Good controls can always be masked?
  - No, sometimes IV versus pill of same medication
  - Try to mask interventions
No Placebo/Control = Problems

• People tend to do ‘better’ receiving some treatment, even placebo or standard of care
  – Care matters
  – Comparing a patient on treatment to baseline does not take this into account

• Comparing population incidence rate to beginning of program does not take into account many factors (may look worse before better)
Additional Problems

• No blinding
  – Researchers and participants tend to interpret findings in favor of new treatment
  – Investigator/participant bias

• No randomization
  – Impossible to distinguish effect of time from treatment effects
  – Confounding
What is the right control group for a randomized study?

- Waitlist control
- Placebo control
  - Remove ‘active’ component; may still have an impact on outcomes
- Active controls
  - Standard of care
  - Alternative intervention
  - Fewer parts of multi-part intervention
- Attention control
- No universally-appropriate control group
Controls Also Present in Many Non-Randomized Studies such as Case-Control Studies

- Siblings
- Community controls
- Hospital controls
- Worked in same area but not present for workplace exposure

- Match on many different variables or not
- Multiple control groups
Consider

• Do the effects plateau?
• Time
  – Long term differences
  – Attenuation
  – Delayed response
Time and Other Elements

• Time is our favorite confounder in uncontrolled studies
  – Differential time participating is an issue
  – Differential drop-outs
  – Time in an environment, age, season
• Social support
  – Meeting in a group may have an impact
  – Talking to someone, empathy, may impact
• Exercise
  – Exercise helps cardiovascular risk factors
  – Exercise helps stress
Masking/Blinding

- Less common in non-randomized studies, but can mask outcome assessors as to hypothesis
- Specify whom to be masked, why, how, and to what
- Assess effectiveness of masking
- Specify criteria for unmasking, whom to be unmasked
- Mask determination of outcome so that reviewers are unaware of treatment assignment; provide information on "need to know" basis
Reproducible Measurements: Regardless of Study Design

- Well defined cohort
- Exclusion and inclusion criteria
- Study conduct
- Outcomes
- Study data, analyses
Studies ≠ Gold

• True for randomized and non-randomized studies
  – Volunteer bias
  – Inclusion/exclusion criteria
  – Measures
  – Artificial interventions/treatment definitions
Outline

✓ Where to start
✓ Taxonomy and examples
  ✓ Observational and interventional studies
✓ General good study design
  ✓ Intervention and dose
  ✓ Comparisons
➢ Conclusions
  • Confounding and effect modification
Your Question Comes First

- May need to rewrite
- If you change your question later
  - May not have the power
  - May not have the data
- Need to know something about the population
Consider

• Questions you want to ask
• Hypotheses you want to test
• Key factors you wish to control
• Ethical issues and constraints
• What can be said with each control group?
  – More than one control group?
Outline

✓ Where to start
✓ Taxonomy and examples
  ✓ Observational and interventional studies
✓ General good study design
  ✓ Intervention and dose
  ✓ Comparisons
✓ Conclusions

➢ Effect Modification and Confounding
Different Variables May Be

- Effect Modifier(s)
- Potential Confounder(s)
- Other things

- If measured these are usually “covariates” in the statistical model
Effect Modification

• Interaction
• Synergy
  – Could be larger or smaller
• Association between outcome and another variable (e.g. intervention) is modified by different levels of a third variable
Smoking, Asbestos, and Lung Cancer

- Smoking (alone) \( \uparrow \) risk of lung cancer by A
- Asbestos exposure (alone) \( \uparrow \) risk of lung cancer by B
- Smoking AND having asbestos exposure \( \uparrow \) risk of lung cancer by MORE/LESS than A+B
Effect Modification

- *A Short Introduction to Epidemiology*
  - Neil Pearce chapter (2005)
- The phrase effect modification, defined for different professions
  - Biostatisticians, public health workers, physicians, lawyers, biologists, epidemiologists,....
Confounding

• Two or more variables
• Known or *unknown* to the researchers
• Confounded when their effects on a common response variable or outcome are mixed together

• Association between an exposure and outcome is misestimated due to the failure to account for a third factor (the confounder)
Consider

- Association observed between carrying matches in your pocket and lung cancer
  - Carrying matches causes lung cancer
  OR
  - Association between carrying matches and lung cancer is result of confounding by another unmeasured variable associated with both

(Pam Shaw, CTR Course 2013)
Coffee and Pancreatic Cancer

Coffee Drinking

Pancreatic Cancer
Confounding Example

• Relationship between coffee and pancreatic cancer, **BUT**
• Smoking is a known risk factor for pancreatic cancer
• Smoking is associated with coffee drinking
  – Coffee drinking is associated with smoking
• Smoking is not a result of coffee drinking
Coffee and Pancreatic Cancer

- Coffee Drinking
- Smoking

Pancreatic Cancer
What is Confounding

• If an association is observed between coffee drinking and pancreatic cancer
  – *Coffee actually causes pancreatic cancer*, or
Coffee Causes Pancreatic Cancer

Coffee Drinking → Pancreatic Cancer
What is Confounding

• If an association is observed between coffee drinking and pancreatic cancer
  – Coffee actually causes pancreatic cancer, or
  – The coffee drinking and pancreatic cancer association is the result of confounding by cigarette smoking
Smoking is a Confounder: Coffee does NOT cause Pancreatic CA
How to Handle Confounding

• Identify potential confounders
  – MEASURE THEM!
  – In the data analysis use
    • Stratification, or
    • Adjustment (add the variable to the model)

• Fear the unknown!
More to Confounding? Yes!

• Residual confounding
  – Poor measure of the confounder
    • Physical activity
  – Even when we put the confounder as measured in the model, not really explaining the effect of real physical activity in the model

• Example
  – Ever Smoked yes/no; pack years
Randomization = No Confounders! Wrong!

- Side note
- Randomization helps protect against confounding
- Does not prevent confounding
- Non-random drop-out or attrition
- Patients testing substance
  - And then dropping out, or taking more of item
Confounding and Effect Modification

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Baseline variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>No confounding</td>
<td>test, control</td>
</tr>
<tr>
<td>No effect modification</td>
<td>effect size</td>
</tr>
</tbody>
</table>

| No confounding           | Confounding       |
| Effect modification      | Effect modification |

John Powers 35 March 2014 IPPCR
Thanks!

• Please submit questions and comments electronically so all 4000+ of us can share in the dialog