

STUDY DESIGNS

IN BIOMEDICAL RESEARCH



COURSE INFORMATION & ORGANIZATION

COURSE OBJECTIVE:

Common weaknesses of most Biostatistics Programs:

- (1) Not enough emphasis on the “**Bio**” part of “Bio-statistics”; and
- (2) Not enough emphasis on the “**Design**” part of “Design **and** Analysis”

The Objective of this course is to fill those gaps, covering “**Study Designs in Biomedical Research**”.

THE CERTAINTY OF UNCERTAINTY

If there is a suspect for a crime, it would involve uncertainty: guilty or not guilty. The Truth is unknown or uncertain because:

(1) Variability: every case is different

(2) Incomplete Information: key evidence


(gun, knife, motive, etc...) may be missing!

(Those are also true in statistical data analysis:

Random variation and Use of samples)

HOW DOES (AMERICAN) SOCIETY DEAL WITH UNCERTAINTIES ?

- ▶ We form a Hypothesis: “Every person is innocent until proven guilty” (written in our Constitution),
- ▶ We gather data: Evidence against Hypothesis- not against the suspect, then
- ▶ We decide whether Hypothesis should be rejected (If it is, the verdict is “Guilty”)



You are familiar with the word “Hypothesis”
which refers to any statement made;
You are familiar with “Data gathering” – or,
“Data collection”;
And the term “Decide” or “Decision” recalls of
“Statistical Inference”


Data collection is a crucial part of the process;
without data, decisions cannot be made.

TRIAL BY JURY

That process to deal with uncertainty is called “Trial by Jury”; “**trial**” as in “**Clinical Trial**”. It is the way our society deals with uncertainties. **Its goal is to minimize errors or mistakes** (not to eliminate them; mistakes are still being made every day)


ELEMENTS OF A SUCCESSFUL TRIAL

- ▶ A probable **CAUSE** (a crime and a suspect)
- ▶ A thorough **INVESTIGATION** (by police)
- ▶ An efficient **PRESENTATION** (by D.A.'s office/attorneys- including the organization and summarization of evidence)
- ▶ A fair & impartial **ASSESSMENT by Jury**, after that a decision is made.



Consider the following “suspects” (some are good, some are bad):

- ▶ **Cigarette smoking (for lung cancer),**
- ▶ **Pesticides (for leukemia),**
- ▶ **Certain defective/mutated gene or genes (for breast cancer),**
- ▶ **Certain new treatment (for some disease)**




In the **trials** of “Smoking”, of “Pesticides”, of a “Defective Gene”, or of certain “Treatment” the Process is not called Trial by Jury; they are referred to as “Research”. And **research involves uncertainties** too. Sometimes we can implicate the suspect, sometimes we fail. The terms Trial/Investigation/Research are sometimes used interchangeably; The “trials” include “clinical trials”.

HOW DOES SCIENCE DEAL WITH UNCERTAINTIES ?

- ▶ We form **Hypothesis**: From experience & observations (The process leads to the so-called research questions)
- ▶ We **gather data**: Experiments & Trials, Surveys, Medical Records Abstractions.
- ▶ We **make decision** based on what we find in the data.

COMPONENTS OF RESEARCH


- ▶ The tool to carry out research is Biostatistics – an analog of trials by jury.
- ▶ The presentation (e.g. by D.A.), the debate, and the verdict form the more visible part of Biostatistics: Data Analysis;
- ▶ The investigation is the other, even more important component: Research Design.



Now, it's “Design and Analysis”.
There are two parts or two steps of data analysis: The first part (by D.A. and prosecutor) is “**Descriptive Statistics**” and the second part (by jurors) is “**Inferential Statistics**”.

ELEMENTS OF GOOD RESEARCH

- ▶ A good Research Question with well-defined objectives & endpoints,
- ▶ A thorough Investigation, lots of data
- ▶ An efficient Presentation: data organization & summarization, and
- ▶ A proper Statistical Inference (the process & methods of drawing conclusions)



We always take the “proper statistical inference” point seriously but sometimes overlook the “thorough investigation” clause. Take study designs more seriously.

AREAS OF BIOSTATISTICS

Research/Biostatistics is a three-step process:

- ❖ Sampling/design: Find a way or ways to collect data (going from population to sample).
- ❖ Descriptive statistics: Learn to organize, summarize and present data which can shed light on the research question (investigating sample).
- ❖ Inferential statistics: Generalize what we learn from the sample or samples to the target population and answer the research question (going from sample to population).

THE IMPORTANT PHASE


Just as in the case of “Trial by Jury”, the most important stage of the “Research Process” is the Design: How & How Much data are collected! Also, It dictates how data should be analyzed. May be it’s less on the question of “how” to collect your data but **the decision on “when to do what”**. Studies may be inconclusive because they were poorly planned, not enough data were collected to accomplished the goals and support the hypotheses.

WHAT IS A RESEARCH DESIGN?


Very briefly, designed experiments are conducted to demonstrate a cause-and-effect relationship between one or more explanatory factors (or predictors) and a response variable. The **demonstration of a cause-and-effect relationship** is accomplished by altering the level or levels of the explanatory factors and observing the **effect of the changes** (i.e. designed values of predictors X's) on the response variable Y.

A Simple Example:


An experiment on the effect of Vitamin C on the prevention of colds could be simply conducted as follows. A number of n children (the sample size) are randomized; half were each give a 1,000-mg tablet of Vitamin C daily during the test period and form the “experimental group”. The remaining half , who made up the “control group” received “placebo” – an identical tablet containing no Vitamin C – also on a daily basis. At the end, the “Number of colds per child” could be chosen as the outcome/response variable, and the means of the two groups are compared.



The “t-test”, at the end, is part of “**Data Analysis**” and “randomization”, by which a child is assigned to the “Vitamin C group” or “Placebo group”, is a part of “**Study Design**”.



The setup is a **Regression Application**, designed experiments are conducted to “demonstrate” a cause-and-effect relation between one or more explanatory factors (or predictors) and a response variable. **Different ways to show case the relationship form different “designs”**. And this course is aimed to introduce some of these designs.

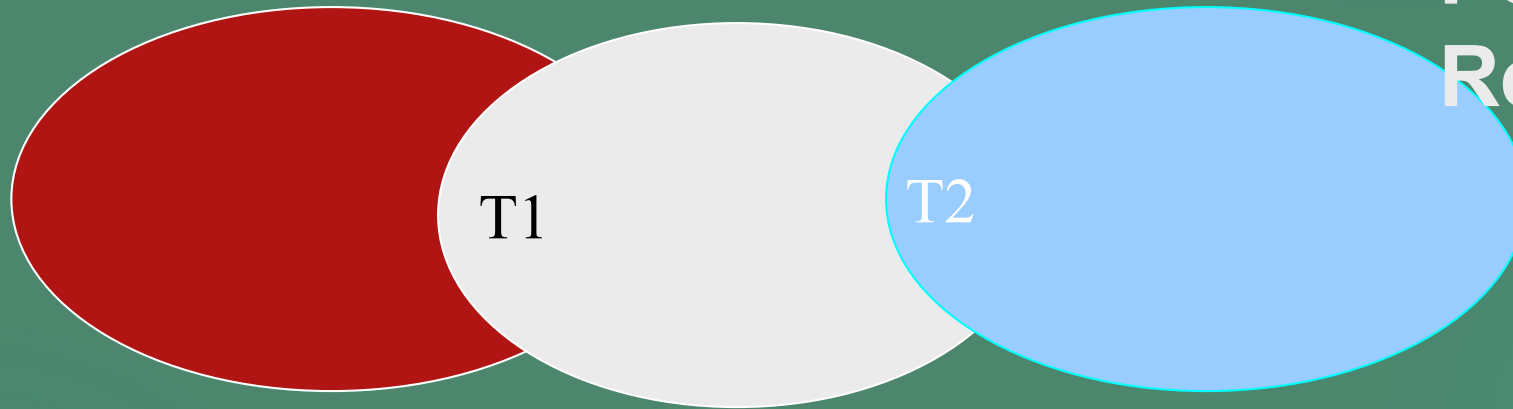


We cover not only “how” to collect data but **we also cover some methods to analyze some special forms of data coming out of these designs.** As previously mentioned, many designs dictate how data should be analyzed; some of these methods are not included in conventional data analysis courses.

Designed experiments are conducted to “**demonstrate**” a cause-and-effect relation between an explanatory factor (or predictor) and a response variable. The demonstration of a cause-and-effect relationship is accomplished, to put it in a simple way, by altering the level of the explanatory factor (i.e. “designed”) and observing the effect of the changes (i.e. designed values of predictor X) on the response variable Y. Designed experiments are often used as “**comparative**” in nature; that is comparing responses from different levels of the predictor.

Clinical Research

Population
Research



Laboratory
Research


Studies can be grouped into these areas: Population, Laboratory, and Clinical; plus **Translational Research**, the component of basic science that interacts with clinical science (T1) or with population science (T2). T1 is a huge area; a lot of studies we do belong there and **clinical trials form only a small part of Clinical Research.**

SOME TERMINOLOGIES

- ▶ **Clinical Studies**: Class of all scientific approaches to evaluate Disease Prevention, Diagnostics, and Treatments; the term ‘clinical’ aims at people/human.
- ▶ **Clinical Trials**: Subset of clinical studies that evaluates Investigational Drugs; they are in prospective and longitudinal form (the very basic nature of “trials” is prospective). Clinical trials are clinical studies but not the other way around. **Clinical trials form only a small part of Clinical Research.**

In addition, what you would learn in a typical clinical trial course are all about “randomized, controlled, phase III clinical trials”. It’s about “serious, large-scale efforts” in **later stage/phase** of clinical research.

How often, or how likely, you’re going to see these trials in your practice as a consulting statistician or practitioner?



The scope of this course is expanded beyond “clinical trials” because not all biomedical research studies are clinical trials; there are studies of other forms. There are more studies of other forms than there are clinical trials.

ABOUT THE COURSE

- ▶ 7470: Study Designs in Biomedical Research
- ▶ Prerequisites: A year of applied statistics – including ANOVA & Regression; some statistical theory would help but not required.
- ▶ Contents: Course covers topics in 4 major categories
 - ❖ Basic Design Principles
 - ❖ Designs in Diagnostic Medicine,
 - ❖ Some modern popular study designs,
 - ❖ Early Phase Clinical Trials,
- ▶ I blend in these topics a few lectures on Applications of Regression – including Bioassays, Seasonal Diseases or Demand Curves, Combination Chemotherapy , and Personalized Medicine – whenever appropriate.

COURSE INFORMATION

- ▶ Meet here: 11:15-12:30 T & Th (MoosT 1-435)
- ▶ Information are at Web address:
www.biostat.umn.edu/~chap/pubh7470B
- ▶ You can read, print, or down load the files to keep; there is a separate file for each class;
- ▶ You could always go back to this web site, no password required.
- ▶ I recommend that you read the file before each class – then read again after class.

ABOUT THE INSTRUCTOR

- ▶ **Chap T. Le, Ph.D.**
- ▶ **Distinguished Professor of Biostatistics.**
- ▶ **Teach “7405: Biostatistics Regression” & “7470: Study Designs in Biomedical Research”, and “6432: Selected Topics in Translational Research and Clinical Research”.**
- ▶ **Serve as Director of Biostatistics for the Masonic Cancer Center & several Clinical Cancer Research and Tobacco Research Programs (large P- & U-grants)**
- ▶ **“I do practice what I teach”.**



<https://directory.sph.umn.edu/bio/sph-a-z/chap-le>

Full Resume: Research Grants, Books, Research publications/papers (in real time through Google Scholar – citations included), courses and lectures (both PubH 7405 and **PubH 7470**, near the bottom of the page – alternative access to course's home page).

CONTACT INFORMATION

- ▶ Office: A441 Mayo Building
- ▶ Phone: (612) 624-9963
- ▶ e-mail: chap@umn.edu
- ▶ Home Page: www.biostat.umn.edu/~chap; you can access the course's page from here.

(No appointments needed; preferred option: e-mail, answered mail the same or next day)


A FEW GOOD BOOKS

There is **no textbook**; hard to find one that covers all areas. The following are a few good references; read a few chapters if you can find copies of books:

- ❖ **“The Statistical Evaluation of Medical Tests for Classification and Prediction”** by M. Pepe.
- ❖ **“Bioassay”** by John Hubert, or **“Statistical Method in Biological Assay”** by David Finney
- ❖ **“Clinical Trials in Oncology”** by S. Green.

COURSE EVALUATION

- ▶ **Emphasis is on learning**, not testing; conducted as one of “seminar courses”; your input are appreciated.
- ▶ There are lectures and recommended readings
- ▶ Attendance required – **some homework are only suggested**, not due and not graded (but you are strongly recommended to try some or all).
- ▶ There will be **no in-class exams**; major course feature consists of class participation, two take-home exams (30% each), and 2 group presentations (in one class period, at the end of semester, 40%).



Class will be divided into 3-4 groups of 2-3 students each. Each group will receive:

- (a) Copy of 1 research paper in category A (focusing more on the Design);**
- (b) Copy of 1 research paper in category B (focusing more on statistics/methods, analysis of data coming out of some special designs);**

Problems for Biostatistics Majors:

- (1) A few lectures maybe overlapped with others
- (2) Too many new applications
- (3) Not enough time to enjoy statistical details;

Goal: getting yourself into **applied statistics**

Problems for Non-majors:

- (1) Level of statistics maybe overwhelmed
- (2) Many applications are not in your field or what you think you might need; hard to motivate
- (3) Rather fast pace with not enough details.

Goal: Only to familiarize with the topics; be practical on how much you can learn – and put up with what you feel you might not need

Problems for Everybody:

(1) No books; reading articles from different sources

(2) Different learning model; out of “cram & exam taking” mode – and into seminar environment.

Goals: Participation + Discussion + Self learning

TRIAL #1 for GROUP PRESENTATION

Thomas D.B. et al.

Randomized Trial of Breast self-examination in Shanghai: Methodology and Preliminary Results

JNCI: Journal of the National Cancer Institute,
89(5), 1997

[if interested, could supplement with:

Thomas D.B. et. al.

Randomized trial of breast self-examination in Shanghai: final results.

JNCI: Journal of the National Cancer Institute,
94(19):1445-57, 2002]

TRIAL #2 for GROUP PRESENTATION

Donny E.C. et. al.

Randomized trial of reduced-nicotine standards for cigarettes

NEJM: New England Journal of Medicine.
373(14):1340-1349, 2015.

[If interested, could link to a “sister” trial:
Hatsukami et al.

Effect of immediate vs gradual reduction in nicotine content of cigarettes on biomarkers of smoke exposure: a randomized clinical trial
JAMA 320 (9), 880-891, 2018]

TRIAL #3 for GROUP PRESENTATION

Samavat H. et. Al.

The Minnesota Green Tea Trial (MGTT), a randomized controlled trial of the efficacy of green tea extract on biomarkers of breast cancer risk: study rationale, design, methods, and participant characteristics.

Cancer Causes and Control 26 (10): 1405-1419, 2015.

[if interested, could supplement with:

Dostal AM. et. al.

The safety of green tea extract supplementation in postmenopausal women at risk for breast cancer: Results of the Minnesota Green Tea Trial.

Food and Chemical Toxicology 83: 26-35, 2015]

TRIAL #4 for GROUP PRESENTATION

Yuan J.M. et.al.

Clinical Trial of 2-Phenethyl Isothiocyanate as an Inhibitor of Metabolic Activation of a Tobacco-Specific Lung Carcinogen in Cigarette Smokers.

Cancer Prevention Research 9(5): 396-406, 2016.

[If interested could link to a smaller trial in the same area:

Fujioka N. et al.

Urinary 3,3'-diindolylmethane: a biomarker of glucobrassicin exposure and indole-3-carbinol uptake in humans.

Cancer Epidemiol Biomarkers Prev. 23(2):282-287, 2014.]

METHOD PAPER #1

Sargent and Goldberg.

A flexible design for multiple armed screening trials.

***Statistics in Medicine* 20: 1051-1060, 2001**

METHOD PAPER #2

Murphy S.A.

An experimental design for the development of adaptive treatment strategies.

Statistics in Medicine 24:1455-1481, 2005

METHOD PAPER #3

Le C. T.

Statistical comparison of two hand washing protocols.

***Statistics in Medicine* 5: 593-596, 1986.**

METHOD PAPER #4

Koffanus M.N. et.al.

**A Modified Exponential Behavioral Economic
Demand Model to Better Describe Consumption
Data**

**Experimental and Clinical Psychopharmacology
23(6): 504–512, 2015**