# STUDY DESIGNS

## IN BIOMEDICAL RESEARCH

## CLINICAL RESEARCH ON MICRO SCALE

# THE ANATOMY & PHYSIOLOGY OF CLINICAL RESEARCH

We form or evaluate a research or research project from/on two different angles or parts: the anatomy and the physiology of research; just like the hardware and software to run a computer operation.

# THE ANATOMY PART

- From the anatomy of the research, one can describe/see what it's made of; this includes the tangible elements of the study plan: research question, design, subjects, measurements, sample size calculation, etc…
- The goal is to create these elements in a form that will make the project feasible, efficient, and cost-effective.

In other words, the Anatomy Part describes the structural features of a research project; what it's made of, its components.

# THE PHYSIOLOGY PART

- From the physiology of the research, one can describe/see how it works; first about what happened in the study sample and then about how study findings generalized to people outside the study.
- The goal is to minimize the errors that threaten conclusions based on these inferences.

In other words, the Physiology Part describes the operational features of a research project; how it works , the sequential movements of the research process.

We tackle them in that order: structural features before operational features; just like how or what software are written may depend on what hardware pieces available.
For a computer, hardware are described in a “manual”; that manual for a research project is the Protocol.

# THE PROTOCOL

- The structure of a Research Project, or its anatomy, is described in its protocol; the written part of the study.
- The Protocol have a vital scientific function to help the investigator organize his/her research in a logical, focused, and efficient way.
- Just like a manual for a computer, investigators have to write a Protocol; and that protocol has to be reviewed an approved before a research project is allowed open to enroll patients.

# RESEARCH PROTOCOL

- That proposal, “the protocol”, have to be reviewed and approved by two separate regulatory groups: an Institutional Review Board (IRB) to scrutinize its safety aspects, such as the patients’ written “consent form” and a research-oriented committee of peers to finalize the peer-review, peer-approval process (e.g. Cancer Protocol Review Committee in cancer centers)
- One committee focuses on ethical aspects of the proposal and one on its scientific contents and feasibility.

# PROTOCOL REVIEW

- The protocol review by peers are focused on two different aspects:

(1) Scientific merit of the protocol: scientific relevance, validity of the hypothesis, adequate study design, biostatistics input, adequate patient population, and feasibility of timely completion;

(2) Priority of the proposed study in regards to competing protocols, priority of its scientific merit and its impact on existing studies.

Patients, not money, are the most precious resources.

Cancer – for example, is rare disease and this rare disease has many sub-types; patients are most precious “resource” that have to be “rationed”, only proposals with high priority are approved to be opened.

# COMPONENTS OF THE ANATOMY

- Research Question: What is the objective of the study, the uncertainty the investigator wants to resolve?
- Background and Significance: Why these questions important?
- Design: How is the study structured?
- Subjects: Who are the subjects and how they will be selected and recruited.
- Variables: What measurements will be made: predictors, confounders, and outcomes.
- Statistical Considerations: How large is the study and how will data be analyzed ("Design" is an important statistical component but listed in the Design Section).

**“Statistical Considerations” is an important required section but that section is just about “Data Analysis”. Study Design is in a separate section which includes, among other things, the determination of Sample Size.**

# SUBSEQUENT MONITORING

After getting approved and opened for operation, approved and open protocols are reviewed annually by a different regulatory body (often called Data Safety and Monitoring Board or DSMB) for progress made on accrual rate and progress toward study endpoints. Protocols determined to be unsafe (based on excessive adverse effects), to have inadequate accrual, to be without scientific progress, or have little likelihood of completion may be terminated based on committee vote. And, after a year or two, some studies might no longer be scientifically relevant.

Protocols receiving special attention in the review process are those involved interventions or treatments: Primary therapy of Cancer such as Chemotherapy, Radiotherapy, Surgical therapy, Transplants, Immunotherapy, or Gene therapy.

# RESEARCH QUESTION

- The research question is the objective of the study, the uncertainty the investigator wants to resolve
- A good research question should pass the "so what?" test; setting the answer should contribute usefully to our state of knowledge

Research questions often begin with a “general concern” but that must be narrowed down to concrete, “researchable issues”. For example:

Concern: Should people eat more fish?

Specific issues:

(1) Does eating fish lower the risk of cardiovascular disease?

(2) Is there a risk of mercury toxicity from increasing fish intake in older adults?

The acronym FINER denotes five essential characteristics of a good research question:

Feasible,

Interesting,

Novel,

Ethical, and

Relevant.

# BACKGROUND & SIGNIFICANCE

- This section of a protocol sets the proposed study in context and gives its rationale.
  - What is known about the topic at hand? (Citing previous research that is relevant – including the investigator's own work)
  - What are problems with the prior/cited research & what uncertainties remain?

# MORE SPECIFICS

- Two basic items are:
  - Why is the research question important?
  - What kind of answers will study provide?
- Often work on the significance section would help/lead investigator to modifications in the issues of the research question.

# STUDY SUBJECTS

- Decisions on two major issues: Who are included and How to recruit them?
- The first is to specify Inclusion and Exclusion criteria that define the kinds of patients best suited to the research questions
- Where/When to recruit enough people in order to answer the research question – including feasibility of recruitment.

# INCLUSION/EXCLUSION

- Inclusion criteria: Patient characteristics required for entry, describing the population of patients that the drug is intended to serve. There are also exclusion criteria as well.
- For eligibility, consideration should be given to patients who are likely to benefit from treatment & to the generalization of the results:
  (i) Effectiveness of the treatment may be masked by the inclusion of patients with little chance of responding; (ii) On the other hand, with narrow criteria, generalization may be compromised.

# VARIABLES

A set of decisions concerning the choice or choices of which variables to measure.

- Predictor or explanatory Variables
- Outcome Variables (primary, secondary)
- Confounders or Confounding Variables (and how to control them)

# CONFOUNDERS

- A confounder is not under investigation, but may be related to the primary "outcome" and/or predictor variables; an effect modifier is a special case – An effect modifier alters effect of some predictor variables.
- We can control for confounders through the design (stratification, stratified randomization) or through the analysis (use of regression).

# BIOSTATISTICS REVIEW

(1) The endpoint or endpoints of the study

(2) Patient characteristics that may affect response (leading to possible stratification)

(3) The design of the study

(4) Accrual goal and statistical power.

(5) A Plan for Data Management

(6) Method or Methods for data analysis

(7) Criteria for stopping the trial

# Trials & Validity

# A BASIC ISSUE IN RESEARCH

Most of the times, inexperienced researchers mistakenly act like there is an identifiable, existent parent population or populations of subjects. We act as if the sample or samples is/are obtained from the parent population or populations according to a carefully defined technical procedure called "random sampling".

This is not true in real-life biomedical studies. The laboratory investigator uses animals in his projects but the animals are not randomly selected from any large population of animals. The clinician, who is attempting to describe the results he has obtained with a particular therapy, cannot say that his patients is a random sample from a parent population of patients.

# ONE-SAMPLE CASE

A surgeon might attempt to convince readers that the results on his 25 patients typify the results expected from his procedure.

On the one hand, he carefully explain/describe his report as “pure description”. On the other hand, he goes to some lengths to assure that that these patients are like a sample – “unselected” . He makes an inference from sample mean to population mean; calculating the standard error which can help in assessing the reliability of the sample mean for this purpose. Then a 95% confidence interval is provided to complete the report.

Many “one-sample studies” are still being conducted because there are no better other choices. A typical case are “Phase II Clinical Trials” for cancers. A group of patients take the same dose of an experimental drug; the result is a “response rate” (the proportion of patients respond to the new drug: size of tumor reduced in half lasting four weeks or longer).

**However, the broad inference to patients operated by other surgeons, in other years or other institutions is still … very dangerous, And the standard error of the mean cannot be trusted to measure all of these uncertainties because "random sampling" has not been done. So, what can investigators like this surgeon do in one-sample cases?**

First, the surgeon can describe his group of patients in some detail so that his readers can see the nature of the patients he operated, their age range, the severity of their disease, and so forth; the logic is that the more similar patients the results are more likely similar.

Second, in measuring the effects of treatment or operation, he can report measurements before and after surgery, so that each patient serves as his/her own control, so to speak.

The focus on the mean/proportion and its standard error might be misleading; single sample studies remain difficult to evaluate, with or without statistics.

# MULTIPLE-SAMPLE CASE

- Because they are not population-based (there is not an identifiable, existent parent population of subjects for sample selection), biomedical studies – designed experiments are "comparative". That is the validity of the conclusions is based on a comparison.
- In a clinical trial, we compare the results from the "treatment group" versus the results from the "placebo group". The validity of the comparison is backed by the "randomization", a method proposed by Fisher in 1923.

Randomization serves two purposes. First, the groups of study units or arms receiving the different treatments tend to be comparable on all variables, known and unknown. Second, such randomization provides a secure foundation on which statistical measures (standard error, p-value) can be justified.

# ETHICS OF CLINICAL TRIALS

- Clinical trials are mostly confirmatory; or so believed by investigators.
- According to the principle of "Good Medicine", physicians are obligated to work for better treatments for disease; then why putting a patient in a trial where that patient has 50% chance of receiving a treatment/placebo which is believed to be second best?
- We could argue that there are a number of other factors that counterbalance this ethical dilemma, that counterbalance any disadvantage to the patient and accrue to his/her net benefit.

Here is a short list of ethical supports for the modern physician-scientists:

(i) An informed consent is required;

(ii) Many patients cannot get the new treatment unless he/she participates in a clinical trial – in fact, results of both treatment often turn out substantially better than anticipated;

(iii) In many clinical trials, patients do get better for a number of reasons (closer, nursing attention, more frequent lab tests, more frequent visits and care by study physicians);

(iv) Placebo effects;

(v) Patients are often promised the new treatment later if it turns out more effective and he/she got assigned to placebo arm; and

(vi) The trial is terminated as evidence emerged that the new treatment is more superior

Biomedical studies are often conducted to "demonstrate" or confirm or establish a relationship between an exposure or explanatory factor and an outcome or response variable. The demonstration is accomplished by comparing the outcomes or responses from different levels of the explanatory factor or exposure. Different ways to show case the relationship form different "designs".

# COMPARISON OF TWO MEANS

- In many cohort studies, the endpoint is on a continuous scale. For example, a researcher is studying a drug which is to be used to reduce the cholesterol level in adult males aged 30 and over. Subjects are to be randomized into two groups, one receiving the new drug (group 1), and one a look-alike placebo (group 2). The response variable considered is the change in cholesterol level before and after the intervention.
- The hypothesis to be tested is $H_0$: $\mu_2 - \mu_1 = 0$?

# COMPARISON OF 2 PROPORTIONS

- In many cohort studies, the endpoint may be on a binary scale. For example, a new vaccine will be tested in which subjects are to be randomized into two groups of equal size: a control (not immunized) group (group 1), and an experimental (immunized) group (group 2). Subjects, in both control and experimental groups, will be challenged by a certain type of bacteria and we wish to compare the infection rates.
- The hypothesis to be tested is $H_0: \pi_2 - \pi_1 = 0$?

# THE TASKS IN THE "STATISTICAL TESTING" PROCESS

To proceed through the Testing Process - a successful one, We need the following items:

(1) A Null and an Alternative Hypotheses

(2) The Research Design & Data

(3) Key Statistic (called "Test Statistic")

(4) (Statistical Guidelines) & The Conclusion

[A Hypothesis is just a statement, usually by an investigator- but could be by anyone, true or false, with or without supports.]

# NULL HYPOTHESIS

- Among the numerous possible hypotheses involved in a problem, there is a very special one- called the "Null hypothesis" and is denoted by $H_0$.
- The Null Hypothesis $H_0$ is the counterpart of the Constitution statement stipulating "Innocence". For example, when a researcher is concerned about the relationship between Oral Contraceptive (the "Pill") and SBP; it is about the Means of two Populations: the populations of OC users and of OC non-users.
- The underlying Null Hypothesis is "$H_0$: $\mu_1 = \mu_2$".

# HYPOTHESIS TESTS

- A "Hypothesis Test" is a Decision-making Process that examines a set or sets of data and, on the basis of expectation under $H_0$, leads to a decision to "reject" or not to reject $H_0$. $H_0$ is "rejected" (Guilty!) if the data show overwhelmingly that it is almost impossible to have the data that we already collected if $H_0$ is true.
- Hypothesis Tests are also called "Tests of Significance"; the term "significant" only means "real"; the conclusion that, say, an observed difference is real- not happened "by chance".

# ALTERNATIVE HYPOTHESIS

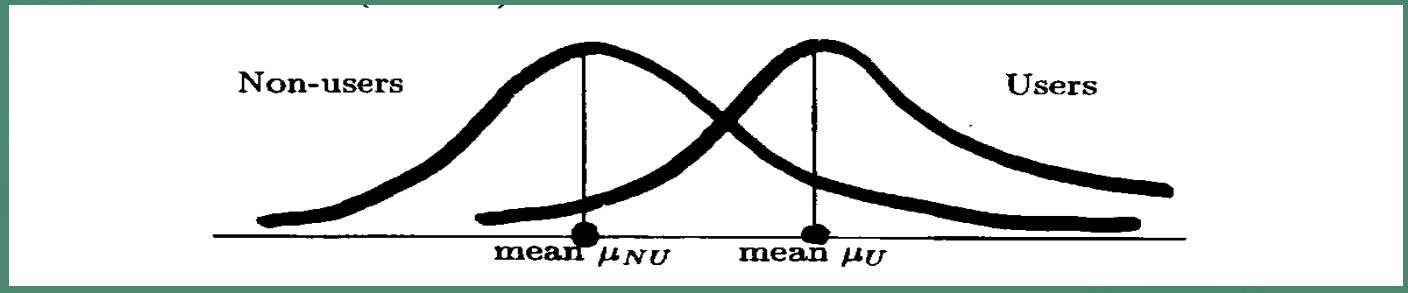
- The "Alternative Hypothesis" $H_A$ is the counterpart of the "Charge" in a Trial by Jury (e.g. first-degree murder). It is important affecting the decision; the jury may see that the suspect is "kind of guilty" but the charge is more wrong, too severe that they may vote to acquit him/her.
- In the context of a research project, say, about the relationship between Oral Contraceptive and SBP with $H_0$: $\mu_1 = \mu_2$; a possibility is $H_A$: $\mu_1 > \mu_2$.
- To a researcher, $H_A$ is his/her (primary) Hypothesis.

# THE CHOICE

- In deciding whether to reject or not to reject a Null Hypothesis, the choice is not between $H_0$ and "The Truth"; because the Truth may not be relevant.
- The choice is between $H_0$ and $H_A$; if there are enough data to support $H_A$ then $H_0$ is rejected.
- There are two forms/types of Alternatives: (1) $H_A: \mu_1 > \mu_2$ is "one-sided" Alternative, (2) $H_A: \mu_1 \neq \mu_2$ is a "two-sided" Alternative; For example, from the data $\overline{x_1} < \overline{x_2}$ showing that $H_0$ may be wrong but the Alternative $H_A: \mu_1 > \mu_2$ is even more wrong; $H_0$ is not rejected (data support "the other side"!).

# Variability & Errors

In some medical cases such as infections, the presence or absence of bacteria and viruses are easier to confirm correctly. In other cases, it's not clear-cut. One possible model for these situations would be to think of the blood pressure X is distributed with differences means for users (U) and nonusers (NU). It can be seen from the figure below that errors are unavoidable when the two means $\mu_{NU}$ and $\mu_U$ may be close



# Errors

In making a decision concerning the Null Hypothesis to compare $\mu_U$ versus $\mu_{NU}$, errors are unavoidable. Since a null hypothesis $H_0$ may be true or false and our possible decisions are whether to reject or not to reject it, there are four possible outcomes combinations. Two of the four outcomes are correct decisions:

(i) not rejecting a true $H_0$

(ii) rejecting a false $H_0$

but there are also two possible ways to commit an error:

Type I: a true $H_0$ is rejected

Type II: a false $H_0$ is not rejected

# Types Of Errors

| Truth | $H_0$ not rejected | $H_0$ is rejected |
|---|---|---|
| $H_0$ is true | Correct Decision | Type I Error |
| $H_0$ is false | Type II Error | Correct Decision |

$\alpha$ = Pr(Type I Error)

$\beta$ = Pr(Type II Error)

**Aim is to keep $\alpha$ and $\beta$ as small as possible. If resources are limited, this goal requires a compromise; these actions are contradictory: We fix $\alpha$ at some specific level - say .05 or .01 and $\beta$ is controlled through the use of sample size; (1-$\beta$) is called the "Statistical Power".**

# INFERENCES & VALIDITIES

- Two major levels of inferences are involved in interpreting a study, a clinical trial
- The first level concerns Internal validity; the degree to which the investigator draws the correct conclusions about what actually happened in the study.
- The second level concerns External Validity (also referred to as generalizability or inference); the degree to which these conclusions could be appropriately applied to people and events outside the study.
- Statistical contributions and assessment involve both Internal Validity and External Validity

# STATISTICAL ISSUES

- Statistics is a way of thinking, thinking about ways to gather and analyze data.
- The gathering part (i.e. data collection)comes before the analyzing part; the first thing a statistician or a learner of statistics does when faced with a biomedical project is data collection (followed by data management and data analysis).
- Studies may be inconclusive because they were poorly planned, not enough data were collected to accomplished the goals and support the hypotheses.
- To assure external validity, we have to assure of adequate sample size

Of course, it is always an issue of possible trade-offs: On the one side are the issues of internal and external validities (say, you need a study with large enough sample size); on the other feasibility (dictated by your ability to recruit patients). Therefore, once the study plan has been formulated, it's still a final decision: whether or not to go for it.

## Suggested Readings:

Search and learn about the structure and functions of the IRB, and the subjects' consent process.