Statistics

- A mathematical science pertaining to the collection, analysis, interpretation or explanation, and presentation of data. It is applicable to a wide variety of academic disciplines, from the natural and social sciences to the humanities.

  - The word statistics is also the plural of statistic, which refers to the result of applying a statistical algorithm to a set of data.

  - Mathematical statistics is concerned with the theoretical basis of the subject.
Statistics cont’d

- Applied statistics is both descriptive and inferential statistics.
  - Descriptive Statistics allow a researcher to describe or summarize their data.
  - Inferential statistics allow a researcher to draw conclusions about the data.

▷ The goal of clinical research is to draw inferences – first about what happened to the people in the study and then how these study findings generalize to people outside the study.

▷ **Objective of this lecture:** To discuss the role of a Biostatistician in research.

Biostatistician’s Role

▷ The role of a Biostatistician is to work with the research team or researcher.

The most important time to involve a biostatistician is at the beginning of the study.
How can he/she help you.

- Help with the design of the experiment/study.
- Help with the formulation of specific aims/hypotheses.
- Determine sample size needed.
- Work with the team preparing grants.
  - Ensure the methods and statistical analysis sections address each specific aim.
  - Calculate appropriate sample size to address the specific aims.

- Discuss variables to be measured and how they relate to the study objectives.
- Develop data collection process.
- Provide methods to analyze data.
- Interpret the results.
Biostatistician’s Role cont’d

▷ Design of the Study

- Choose most efficient design.
  - Ensure there are no flaws in the design that will affect the stats and the ability to find anything or make inferences.
- Consider all aims of the study.
- Selecting the experimental units.
- Assure that every variable measured has a bearing on at least one aim.
- Interim monitoring plan.

▷ Assist with endpoint selection

- Multiple endpoints (e.g. safety and efficacy).
- Measurement issues.
  - Subjective versus objective.
  - Pain measurement.
  - Quality of Life.
- Patient benefit.
- Primary outcome vs secondary outcome.
- Continuous outcome vs categorical outcome.
Analysis Plan

- Statistical method for each specific aim.
- Stratifications are included if necessary.
- Account for loss to follow-up.
- Account for Type I and Type II error.

Sample Size and Power

The most common reason statisticians are contacted.

- Sample size is dependent on:
  - The study design.
  - The analysis plan.
  - The outcome of interest.
Problems with having the wrong sample size:
- The study to be ‘underpowered’, not allowing you to make conclusions.
- Waste time and money because the sample size was larger than it needed to be to answer the research question.

Problems interpreting the results due to wrong sample size:
- If a result is not significant, is there no treatment effect or is the sample size too small.
- Did the treatment really work, or is the effect too small to warrant further research.
- CLINICAL significance versus STATISTICAL significance

Sample size ALWAYS requires the investigator to make assumptions
- How much better do you expect the experimental treatment group to perform (improve) compared to the standard treatment
- How much variability might be expected in measurements
- What would be a clinically relevant improvement/difference.

The clinical investigator is responsible for defining these parameters.

Unless you provide data, the statistician cannot tell you what these numbers should be.
Sample Size and Power cont’d

- Power is the probability of concluding the new treatment is effective if it really is effective.

- Type I error is the probability of concluding the new treatment is effective when it really is NOT effective.
  - Type I error = the alpha level of the test.
  - Type II error = 1 - power

- It is hard to conclude a treatment is effective if the sample size is too small.

Data Collection

- Help with designing a database for data collection and storage. (This will facilitate data analysis.)

  - Recommend using REDCap for clinical data
    - REDCap allows for data checks at the point of entry.
    - REDCap allows the data to be exported in several formats for analysis: Excel, SPSS, SAS, R

  - Recommend using REDCap Survey if your research involves a survey.
One of the final stages in research is manuscript preparation. The statistician will:

- Verify the accuracy of any data presented.
- Explain the analysis approach used.
- Ensure the results are correctly interpreted in the context of the research questions.

Biostatistics Resources

- Biostatics Clinic held daily at noon.
  Clinic website:
  http://biostat.mc.vanderbilt.edu/Clinics

- Collaboration Plan
  - This plan allocates biostatistician percent efforts to work long-term with research groups, divisions, or departments.
  - If your research group is already covered by a Biostatistics Collaboration Plan, contact the lead statistician who is responsible for your group. Assignments are listed at: http://biostat.mc.vanderbilt.edu/CollaborationAssignments.
Biostatistics Resources

- **StarBRITE**
  - StarBRITE is an interactive system that provides one stop shopping for research needs.
  - [https://www.mc.vanderbilt.edu/starbrite/](https://www.mc.vanderbilt.edu/starbrite/)

- **CTSA**
  - If you are applying for assistance through VICTR (CTSA), apply for CTSA resources by going to [https://www.mc.vanderbilt.edu/starbrite/funding.](https://www.mc.vanderbilt.edu/starbrite/funding)

- **CRC Research Skills Workshop**
  - Offers basic instruction and practical advice on commonly encountered clinical research topics.
  - Workshops are held each Friday from 9:00 to 10:00 am in the CRC Conference Room (A-3210 MCN).
  - [http://www.mc.vanderbilt.edu/gcrc/workshop.html](http://www.mc.vanderbilt.edu/gcrc/workshop.html)