

BIOSTATISTICS SHORT COURE

- **Instructor:** Lili Zhao lili.zhao@corewellhealth.org
- **Nine Lecture** on Thursday 5-6:30pm
 - (45min lecture + 45min R)
- **Course materials:**
<https://github.com/yungcleee/Corewell>
- **Online R compiler:**
<https://www.mycompiler.io/online-r-compiler>

My Experience

- Collaboration. I have collaborated with researchers in various disease areas, including Cancer, Cardiology, Pathology, Cirrhosis, Neurology, Endocrinology, Obstructive Sleep Apnea, Pulmonary Diseases, Infectious Diseases, Diabetes, and Surgery
- Teaching. I have taught the On Job/On Campus (**OJOC**) program at U of M for 3 cohorts (2016-2021). OJOC is a 20-month program offering training in clinical trial design and advanced biostatistics

What is Biostatistics?

- **Statistics** is the science of data
- **Biostatistics** is the application of statistics to a wide range of topics in biology

Topics to Cover I

- **Summary statistics**
 - mean/median/quantiles/standard deviation/frequency (%)
- **Compare two groups (or > 2 groups):**
 - t-test (independent or paired) and non-parametric counterparts, chi-squared/Fisher's exact
- **Concepts** confidence interval and p-value

Plan to finish these topics after 4 lectures. Students can interpret Table I in medical papers after completion of this part

Topics to Cover II

- **Regression methods** when outcome is *continuous, binary, and time-to-event*
 - (multiple) linear regression, logistic regression, Cox regression
- **Concepts** Association (OR and HR)/Confounding/interaction

Plan to finish these topics after 4 lectures. Students can interpret main results in medical papers after completion of this part

Topics to Cover III

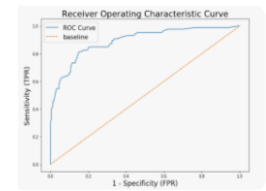
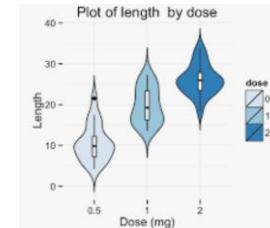
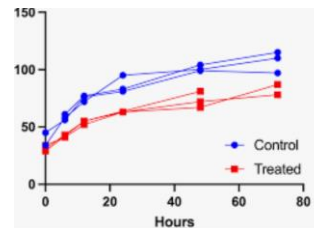
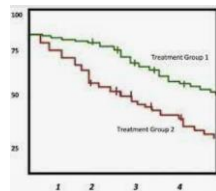
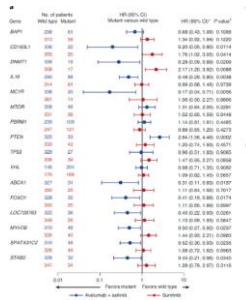
- **Study design:** Clinical trials, Case-control, and Cohort study.
 - **Clinical trials:**
 - (blocked/stratified) randomization, study endpoints, analysis plan, sample size calculation (type I error & power)
 - interim analysis and non-inferiority trials (Optional)

Plan to teach these topics in one lecture. Students learn important statistics to design a simple study after completion of this part

Graphs to Learn

- Graphs

for example: boxplot, spaghetti plot, Kaplan-Meier curve, forest plot, and ROC curve



Students can interpret some figures in clinical research papers and make simple ones on their own

Optional Topics

- **Diagnosis test**
 - Sensitivity/specificity, NPV/PPV, ROC, kappa for agreement
- **Analysis for longitudinal data**
 - (Generalized) linear mixed effects model
- **Select important clinical factors and biomarkers**
 - Variable selection methods (such as LASSO)
- **Prediction model**
 - Linear model and machine learning (trees, cross-validation)
- **Causal inference** (IPW and matching)
- **Meta analysis**
- **Mediation analysis**

After the Course

Students should be able to

- Design a study to answer a simple research question and test the hypothesis
- Understand main results in papers published in medical journals
- Develop quantitative thinking
- Talk to statisticians to conduct complex studies