



BB 101
BIOSTATISTICS AND BIOINFORMATICS 101:
Best Practices for Cancer Data
at
The University of Florida Health Cancer Center

Promoting Rigorous and Reproducible Statistical and Bioinformatics Approach for Your Research

Fall 2023

Director
Ji-Hyun Lee
Professor and Director
Division of Quantitative Sciences & Biostatistics Shared Resources
jihyun.Lee@ufl.edu
352-273-9079

Administrative Assistant:
Ken Thompson
Kenneththompson@ufl.edu
352-273-8822

Background

Understanding biostatistics and computational biology is paramount for cancer research. BIOSTATISTICS AND BIOINFORMATICS 101, *BB 101: Best Practices for Cancer Data*, at the University of Florida Health Cancer Center (UFHCC) is a **diverse program presented in a 12-lecture** series. The lectures introduce the basic principles of statistics, bioinformatics, and computational biology, including AI. They are intended for those who are in the **process of learning biostatistical and bioinformatics applications** or for those who **desire a refresher course**. Topics of data, descriptive statistics, estimation, hypothesis testing, correlation/regression, survival analysis, sample size calculation/power analysis, and introduction to Bioinformatics and AI will be taught in this course series. Some classes will introduce basic theory and application of computing as a tool. Over the course, the **reproducibility and replicability** of scientific research will be emphasized considerably.

Course Goal	This class aims to introduce the basic statistical and computational biology concepts and methods for cancer research and promote rigorous and reproducible statistical approaches for cancer research.
Course Prerequisites	None
Course Format	The course will consist of online lectures. The lecture materials (slides) will be posted before the class.
Registration Policy	There is <i>no fee</i> for this lecture series. However, faculty and students should be registered before Sep 3 rd via the website * Note that some students may want to take only selected lectures , and this will be acceptable as long as they registered.
Who will take this course?	Clinicians, Fellows, Cancer Researchers, and cancer research-focused students. No prerequisites on graduate school-level statistical courses are required. Only UF students/employees are eligible, and exceptional cases should be discussed with the coordinator before the registration deadline.
Where	Via Zoom
Time	Sept 6 – Nov 29 (there will not be a lecture on Wed, Nov 22) 12 weeks: Every Wednesday, 2:00 pm – 3:00 pm
Credit policy	No credit for this course.
Homework	Every week there will be a homework assignment (5 questions), considered a test that engages if the student understands the material.
Grade	50% Homework, 50% Attendance → Pass or Fail
Certificate	The class certificate will be given to those who completed and passed the course (10 attendances of 12 lectures are required as the minimum attendance).
Course evaluation	At the end of each class, students must complete an evaluation form for the lecture.

Faculty Profile for BB 101

Name	Academic Rank	Education	Research Focus
Dr. Ji-Hyun Lee	Professor	DrPH. Biostatistics, University of North Carolina at Chapel Hill, NC	Correlated data analysis, Group randomized trials, Generalized linear mixed models, and Best statistical practices.
Dr. Jason Orr Brant	Research Assistant Professor	Ph.D. Biomedical Sciences, University of Florida, FL	Role of epigenetics in disease initiation and progression, Bioinformatics
Dr. Tuo Lin	Research Assistant Professor	Ph.D. Biostatistics, University of California, San Diego	Regression analysis, Longitudinal data analysis, machine learning in clinical research, and statistical consultation

Fall 2023 BB 101
Syllabus

#	Date	Contents	Goals
1	Wed Sep 6	Introduction	<ul style="list-style-type: none"> • Understand data types • Descriptive statistics/graphical summaries • When/how to use and interpret different measures of mean, variance, and %. • Basic statistical distributions
2	Wed Sep 13	Estimation/ Hypothesis Testing and Statistical Inferences	<ul style="list-style-type: none"> • Understand how to lay out scientific questions with hypotheses • Understand basic statistical inference: effect size, p-value, 95% CI, type I error, and statistical power
3	Wed Sep 20	Common Statistical Tests I	<ul style="list-style-type: none"> • Association/Group Comparisons • Basic assumptions and usages for common statistical tests, including the t-, paired t-tests, Chi-Square, Fisher's exact test, and the corresponding non-parametric tests
4	Wed Sep 27	Common Statistical Tests II	<ul style="list-style-type: none"> • Association/Modeling • Correlation • Linear Regression and Logistic Regression Models,
5	Wed Oct 4	Introduction to Clinical Trials	<ul style="list-style-type: none"> • Concepts of Phase I, II, III, and IV clinical trials in oncology • Translational study → Confirmation study • New Trend for Phase I/II, including the Bayesian approach
6	Wed Oct 11	Survival Analysis	<ul style="list-style-type: none"> • Kaplan-Meier Curve with Log Rank Test • Cox Proportional Hazard Regression Model
7	Wed Oct 18	Sample Size Calculation I	<ul style="list-style-type: none"> • Understand the role of sample size, effect size, variability, and power in study design
8	Wed Oct 25	Sample Size Calculation II	<ul style="list-style-type: none"> • For Survival data • For <i>in vitro</i> experiments
9	Wed Nov 1	Introduction to Bioinformatics and Computational Biology	<ul style="list-style-type: none"> • Introduction to bioinformatics and computational biology concepts and tools. • Overview of genomic sequencing technology: historical to cutting-edge
10	Wed Nov 8	RNA-Seq Analysis: Best Practices	<ul style="list-style-type: none"> • Introduction to RNA-Seq assays and concepts • Best practices for RNA-Seq analysis • Visualization of RNA-Seq data
11	Wed Nov 15	Single-cell RNA-Seq Analysis: Best Practices	<ul style="list-style-type: none"> • Introduction to scRNA-Seq assays and concepts • Prior knowledge regarding RNA-Seq • Best practices for scRNA-Seq • Visualization of scRNA-Seq data
12	Wed Nov 29	DNA Methylation Analysis: Best Practices	<ul style="list-style-type: none"> • Introduction to Methyl-Seq assays and concepts • Best practices for Methyl-Seq analysis

Lecture Evaluation Form

This evaluation is anonymous.
DO NOT PUT YOUR NAME ON THIS SHEET.

COURSE: BB 101 at UFHCC, Fall 2023

Course Sequence #:

Please Score

(Lowest) 1 – 2 – 3 – 4 – 5 (Highest) for each box below

Lecture Content	Organization & Presentation	Overall Rating

COMMENTS:

-

-

-

Question for you: How long has been taken to touch your cell-phone since the class started?

Answer: _____ min.

- Please note that the data collected from this question will be used to illustrate the **Survival Data Analysis Method (Lecture 6)**.