

# Clinical Research Bootcamp 301

Emory University School of Medicine

April 22, 2022

This course is directed toward faculty in the School of Medicine with limited or no background in basic statistics or statistical software. Participants will be required to bring a laptop and install the necessary software in anticipation of hands on support with sample datasets.

**Learning objectives:** Upon completion of this course, participants should be able to: (1) Describe basic statistical concepts, (2) Importing different types of files in R (.csv, .txt, .xlsx, etc.), (3) Choose and perform basic tests for comparing continuous variables and proportions [e.g., t-test, chi-square, Fisher exact test, etc.], (4) List resources at Emory for biostatistical assistance and prepare datasets for smoother interfacing with those resources, and (5) List resources further study.

Time:	Title:	Presenter(s):	Description:
7:45 AM	Breakfast		
8:00 AM	Welcome	Katie Ross-Driscoll	Introductions and overview of the day
8:15 AM	A Primer on Statistics	Dong Li	Defining about goals of statistics/statistical testing and designing research hypotheses.
9:00 AM	Importance of Data Cleaning and Management	Jessica Harding	Principles of tidy data, introduction to data manipulation and management
9:50 AM	Break (20 min)		
10:10 AM	R-Studio Fundamentals	Ram Jagannathan	RStudio Tour, Installing R Packages, and R notebook
10:40 AM	Importing data	Ram Jagannathan	Importing different types of files in R (.csv, .txt, .xlsx, etc.) which will be used for the subsequent data analysis.
11:15 AM	Data Cleaning and Management in R	Dong Li	Introduction to data manipulation (e.g., age, BMI calculation, averaging BP, convert time into years, etc.) and transformation using Tidyverse. Principles of record data linkage in R.
12:00 PM	Lunch (45 min)		
12:45 PM	Descriptive Statistics	Jessica Harding	Introduction to t-tests (two-sample, paired), Wilcoxon signed-ranks, and Mann-Whitney U
1:30 PM	R Practice	Ram Jagannthan, Dong Li, Katie Ross-Driscoll	Case-based practice of concepts covered in the last module. Creating Table-1 (baseline characteristics of the study population) using tableone package
2:00 PM	Group activity, R-practice, and Troubleshooting		
2:25 PM	Break (20 min)		
2:45 PM	Visualizing the Data in R Using ggplot2	Ram Jagannathan	Data visualization best practices and its implementation in the ggplot2 package.

*\*Subject to change*

3:30 PM	Getting Help	Katie Ross-Driscoll	Free online resources to learn R (Google search and using Stack Overflow). How to access biostatistical resources on campus (about our HSR)
4:00 PM	Overall Wrap-up	Jessica Harding	Best practices for statistical analysis and Time for additional Q&A

***Sponsors:***



**EMORY**  
UNIVERSITY  
SCHOOL OF  
MEDICINE

**Office of Faculty Academic  
Advancement, Leadership  
and Inclusion**



**EMORY**  
UNIVERSITY  
SCHOOL OF  
MEDICINE

**Health Services Research Center  
Department of Medicine  
and Department of Surgery**

*\*Subject to change*