

Welcome to the workshop

This site presents materials for the short course **Tidy Survival Analysis: Applying R's Tidyverse to Survival Data** to be taught at the 2025 Joint Statistical Meetings (JSM) in Nashville, TN.

This course aims to equip participants with the skills to apply tidy principles to survival analysis, fostering a more organized and reproducible approach to data analysis in R.

Target Audience: Statisticians, data analysts, researchers, and students interested in survival analysis who are familiar with R and the Tidyverse.

Time and Place

- Sunday, Aug 3: 8:30 AM - 12:30 PM
- Music City Center | Room: CC-110A

Instructor Profile

Lu Mao, PhD

- Associate Professor of Biostatistics at UW-Madison
- Methodologic research
 - [R01HL149875](#): *Novel Statistical Methods for Complex Time-to-Event Data in Cardiovascular Clinical Trials* (12/01/2019 – 07/31/2028)
 - [DMS2015526](#): *Randomized Trials with Non-Compliance* (07/01/2020 – 06/30/2024)
- Collaborative research
 - Cardiovascular disease, cancer, radiology, behavioral health interventions
- Teaching
 - [Survival Analysis: Theory and Methods](#) (UW; 2020 - 2025)
- Editorial service
 - [Statistical Editor](#), *JACC Journals*
 - [Associate Editor](#), *Statistics for Biopharmaceutical Research*

Learning Outcomes

1. Understand the fundamentals of survival analysis, including key concepts such as censored data, survival functions, and hazard functions.
2. Utilize R's Tidyverse packages to manipulate, visualize, and analyze survival data.
3. Fit and interpret survival models using the `survival` and `survminer` packages in conjunction with Tidyverse functions.
4. Create clear and informative visualizations of survival data, including Kaplan-Meier curves and survival distributions.
5. Communicate survival analysis results effectively using tidy principles.

Outline

- **1. Introduction to Survival Analysis** (30 min)
 - 1.1 Key concepts: survival functions, hazard functions, and censoring
 - 1.2 German breast cancer study: a working example
 - 1.3 Overview of survival analysis with `survival` package
- **2. Data Manipulation with Tidyverse** (45 min)
 - 2.1 Importing and cleaning survival data
 - 2.2 Using `dplyr` for data manipulation
 - 2.3 Simple visualization of follow-up data with `ggplot2`
- **3. Nonparametric Survival Analysis** (50 min)
 - 3.1 Tabulation of summary statistics with `gtsummary`
 - 3.2 Visualizing Kaplan-Meier curves with `ggsurvfit` (or `survminer`)
 - 3.3 Tidy analysis of competing risks using `tidycmprsk`
- **4. Cox proportional hazards regression** (40 min)
 - 4.1 Tabulation of regression results with `gtsummary`
 - 4.2 Model diagnostics and residual plots with `survminer`
 - 4.3 Fine-Gray model for competing risks with `tidycmprsk`
- **5. Machine learning using `tidymodels`** (50 min)
 - 5.1 Modeling basics
 - 5.2 Using the `censored` package for survival models