Peiyu Liu

Biostatistician | Ph.D. Student in Biostatistics, University of Florida

Gainesville, FL • peiyu.liu.stats@gmail.com • +01 (352) 562 3237 • https://peiyuliu-biostats.github.io/

Summary

• 3rd year Biostatistics Ph.D student at University of Florida with 8+ years' experience in statistical modelling.

• Proficient in multiple coding languages including R/Rshiny, Matlab, SAS, Python, C++, etc.

• Specialized in statistical models, differential equations, and Bayesian estimation.

Skills

Mathematical modelling • Biostatistics computing • Machine learning • Clinical trial analysis • R • Matlab • SAS

Positions and Experiences

Biostatistician (Consulting Core&CSQUID)

University of Florida | Emerging Pathogens Institute

• Develop new biostatistical methodology for the design and analysis of EPI studies as needed.

• Conduct rapid-response statistical analyses within Center for Statistics and Quantitative Infectious Diseases (CSQUID), delivering timely evidence to support outbreak evaluation and control strategies.

• Provide pre-proposal study design guidance: perform mathematical modelling, power analyses, calculate appropriate sample size for testing intervention effects, and develop statistical analysis plans.

• Assist with data presentation, manuscript completion for research dissemination.

Research Assistant

University of Florida | Biostatistics Department | Aug 2023 - Current

• Innovated a unified trial evaluation framework by integrating a stochastic antibody kinetics model (ANPP) with temperature-driven SEIR-SIR model, comparing the novel serological vs traditional clinical endpoints and informing trial design strategies.

• Translated biological mechanisms into statistical power by deriving closed-form antibody moment equations, validating them with large-scale Monte Carlo simulations, and providing standardized sample size formulas for reproducibility.

• Synthesized quantitative evidence demonstrating that serological endpoints can reduce trial sample sizes by several orders of magnitude, and provided design insights by linking different interventions to endpoint choices and follow-up timing.

• Led the statistical analysis for a clinical-genomic RSV study on a pediatric cohort, applying Cox PH and regression models to identify key drivers of disease severity, inform patient risk stratification, and explore genotype-phenotype associations.

• Developed 4 web-pages via R shiny predicting sample size, vaccine coverage rate, and infectious diseases.

• Delivered audit-ready statistical outputs ensuring rigor and supporting grant and policy decisions, while conveying findings in clear terms for media and public audiences.

Research Fellow and Lecturer

Hunan Institute of Science and Technology | Math Department | Aug 2021 - Aug 2023

• Modeled tumor–immune dynamics to prove stability of tumor-free states and assess how radioimmunotherapy parameters influence treatment strategies.

• Analyzed meteorological and air quality data from 10 countries using statistical modeling and spatial visualization to quantify their impact on COVID-19 transmission dynamics.

• Directed a ¥50,000 municipal-funded project to build a COVID-19 visualization platform integrating statistical and mathematical models for epidemic monitoring.

• Mentored undergraduate team to a second prize in the National Mathematical Modeling Contest.

• Taught Probability Theory and Mathematical Statistics to around 200 students of sophomore, 243 class hours each year for two years.

Researcher

Shaanxi Normal University | School of Mathematics and Statistics | Aug 2017 - Aug 2021

• Built continuous and discrete SEIAR compartmental model, applying Nonlinear Least Squares, Bootstrap, and MCMC with sensitivity and threshold analyses to quantitatively evaluate and inform vaccination and NPI strategies.

• Supported a data-driven COVID-19 forecasting and visualization platform by designing and maintaining compartmental modeling components.

• Applied data imputation (interpolation, EM), variable selection (Lasso regression, Spearman analysis), Generalized additive model, and spatial econometric models (Moran's I, Spatial Lag Models) to quantify meteorological impacts on air pollution across Xi'an districts.

- Developed and analyzed a nonlinear dynamical system (Riker model) with saturation effects, proving equilibrium coexistence, conducting bifurcation analysis under stochastic disturbances, and validating dynamics through Poincaré map simulations.

Projects

Public Health Outbreak Risk Modeling & vaxineR Package

Lead Statistician

Center of Statistics and Quantative Infectious Dieases | Sep 2025 - Present

- Led rapid-response analyses of kindergarten vaccine coverage under new government policies, coordinating a 12-member cross-functional team to align methodologies and deliver reproducible findings within days.
- Communicated epidemiological insights to high-profile media (e.g., Science, CNN), with analytical visuals cited.
- Directed the full development lifecycle of the open-source vaxineR R package and its companion RShiny dashboard

AI-Powered Identification of Pediatric Crohn's Disease from Medical Imaging

Lead Statistician

Department of Pediatrics | Ongoing

- Defining the statistical and machine learning strategy for classifying active vs. inactive Crohn's disease, including data preprocessing to validation.
- Establishing the evaluation framework for performance, bias assessment, and interpretability.

Large-Scale Application and Inferential Extension of the ANPP Model

Lead Methodologist & Analyst

Department of Biostatistics | Aug 2025 - Present

- Developing a moment-based estimation framework as a computationally efficient alternative to likelihood-based methods for fitting complex stochastic antibody models to cross-sectional data.
- Leading the application across 20+ serological studies, conducting inverse inference to estimate parameters such as mosquito biting intensity and antibody boosting/waning rates.
- Characterizing geographic variation in vector exposure and immune responses.

Statistical Modeling of Arbovirus Transmission Drivers in Puerto Rico

Lead Methodologist & Analyst

Centers for Disease Control and Prevention | Sep 2025 - Present

- Building a household-level transmission model (chain-binomial) for dozens of families, estimating secondary attack rates and within- vs. between-household spread. .
- Applying multilevel regression with geospatial random effects and lagged climate inputs to test drivers (location, demographics, rainfall/temperature); quantify effect sizes with CIs..
- Characterizing policy simulations to compare targeted interventions (e.g., focused vector control or prophylaxis when available) and prioritize high-risk households/areas.

Education

Ph.D. in Biostatistics

College of Public Health & Health Professions, University of Florida

Aug 2023 - Current

M.S. in Computational Mathematics

School of Mathematics and Statistics , Shaaxi Normal University

Aug 2018 - May 2021

B.S. in Statistics

School of Mathematics and Statistics , Shaaxi Normal University

Aug 2013 - May 2017