Shirui (Ray) WANG

Atlanta, GA | 404-490-5889 | wangshirui1021@gmail.com | LinkedIn | GitHub | Blog

Education

Georgia State University	Atlanta, GA, US
Ph.D. in Biostatistics	Jan 2020 – Dec 2025(expected)
University of California, Irvine	Irvine, CA, US
M.S. in Mathematics	Aug 2016 – June 2018
China Agricultural University	Beijing, China
B.S. in Applied Mathematics	Aug 2012 – June 2016

Skills & Knowledge

- **Programming and Computing**: **R** (tidyverse, ggplot2, R Shiny, R Markdown, Parallel Simulation), **SAS**, **Python** (Numpy, Pandas), MATLAB, SQL, C, High-performance computing clusters
- **Statistics: Regression Analysis** (GLM, Mixed effects model, Lasso/Ridge regression), **Survival Analysis** (Cox model, censored data), **Missing Data** (Multiple Imputation, EM, IPW), **Real-World Data** (SEER, Medicaid), **Computational Methods** (MCMC, Bootstrap), **Clinical Trials** (Sample size calculation, Power analysis, Multiplicity adjustments, Experimental Design), **Diagnostic tests and Biomarker Validation** (ROC, AUC), Bayesian Methods, Multivariate Statistics, Longitudinal Data, Meta analysis
- **Machine Learning & Deep Learning**: Random Forest, XGBoost, LightGBM, SVM, K-Means, PCA, Deep Neural Networks, RNNs, GANs, LSTM, Diffusion Models, Large Language Models, Scikit-Learn, PyTorch

Work Experience

American Cancer Society	Atlanta, GA
Research Intern, Cancer Risk Prediction and Health Disparities	May 2025 – Nov 2025

- Developed **imputation** frameworks based on **EM** algorithm and **sequential quadratic programming (SQP)** algorithm, integrated with **flexible parametric survival models**, to address **missing** stage-at-diagnosis data in **SEER** datasets
- Performed a **multinomial logistic regression** analysis to study the association between **Medicaid** coverage continuity and cancer stage at diagnosis, finding that continuous coverage significantly reduced the likelihood of distant cancer

SUNY Upstate University Hospital	Remote
Biostatistician Intern	Dec 2023 – Jan 2024

- Analyzed clinical data obtained from a 2^2 factorial design study to evaluate the impact of Sepsis order sets and notes on key healthcare outcomes (mortality rate, length of stay, ICU utilization, etc.)
- Performed **Chi-square test**, **Friedman test** and **post hoc analysis**, applied **mixed effects model** and provided statistical evidence that the use of order sets and notes significantly lowered the mortality and ICU utilization

Georgia State University	Atlanta, GA
Ph.D. Dissertation Research	Aug 2021 – Dec 2025

- Developed novel statistical approaches for evaluating continuous diagnostic tests (**ROC, AUC, Youden index**) with **partially missing** gold standard verification using **Empirical Likelihood** and **bootstrap** methods
- Conducted extensive **Monte Carlo simulation** studies and performed **biomarker validation** analyses for **real-world cancer and AD datasets** in **R**, achieving accurate and robust inferences when only 36% of the subjects were verified

Georgia State University	Atlanta, GA
Graduate Teaching Assistant	Jan 2020 – May 2025

- Taught elementary statistics as **sole instructor** to non-statistics major undergraduates, effectively conveying statistical concepts and approaches to students from diverse academic backgrounds
- Earned a rating of 4.5/5 on student evaluation and won the **Best Graduate Teaching Assistant Award** (2025) for outstanding teaching performance

China Construction Bank	Shenyang, China
Data Mining Intern	July 2019 – Aug 2019

- Developed predictive models for retail banking operations, training a **Random Forest Classifier** in **Python** to assess customer responsiveness to telemarketing, achieving 93% accuracy on the testing set
- Conducted risk analysis on transaction data to detect potential financial crimes, identifying 500+ gambling accounts based on transaction patterns

Publications & Working Papers

- **Wang S**, Shi S, Qin G. Interval estimation for the Youden index of a continuous diagnostic test with verification biased data. *Statistical Methods in Medical Research* 2025; 34: 796-811. doi:10.1177/09622802251322989
- Shi S, **Wang S**, Qin G. Interval estimation for three-class Youden index with verification bias. *Journal of Biopharmaceutical Statistics* 2025; 1: 1-22. doi:10.1080/10543406.2025.2549361
- Shi S, **Wang S**, Qin G. Interval Estimation for the Sensitivity of a Test to the Early Diseased Stage with Verification Bias. *Revision at Pharmaceutical Statistics*
- Jia S, **Wang S**, Qin G. Methodological Approaches for the Estimation of Confidence Intervals on Partial Youden Index under Verification Bias. *Revision at Pharmaceutical Statistics*
- **Wang S**, Shi S, Qin G. Empirical likelihood inference for the area under the ROC curve with verification biased data. *Submitted*
- **Wang S**, Qin G. Empirical Likelihood-based Methods for Inference of Sensitivity of a Test with Verification Bias. *Submitted*
- Jia S, **Wang S**, Qin G. Empirical Likelihood-Based Confidence Intervals for the Partial AUC with Verification Bias. *Submitted*
- **Wang S**, Islami F, Siegel RL, Jemal A, Choudhury PP. Imputation of missing cancer stage at diagnosis accounting for stage-specific survival. *Manuscript in preparation*
- Choudhury PP, Zhao J, **Wang S**, Jemal A, Yabroff R, Islami F. Are disruptions in Medicaid coverage linked to advanced stage diagnosis of screen-detectable cancers? *Manuscript in preparation*

Awards & Activities

- 2025 Joint Statistical Meeting (JSM), Contributed Paper Presentation, Nashville, TN, Aug 2025
- 2025 Georgia Statistics Day, Student Poster, UGA, Athens, GA, Oct 2025
- Wiley M. Suttles Math Award for outstanding math graduate student (2025)
- Best Graduate Teaching Assistant Award for outstanding teaching performance (2025)
- Mathematics Achievement Award (2022)
- Best undergraduate theses in the department competition (4th in 60 theses) (2016)
- First Prize (Beijing) in China Undergraduate Mathematical Contest in Modeling (2014)
- Second Prize (Beijing) in China Undergraduate Mathematical Contest in Modeling (2013)
- Third-class Scholarship for Outstanding Students of College (2012 & 2014)

More about me...

- **Volunteering**: Blood donation volunteer & China Marrow Donor Program volunteer
- **Movies and Books**: Have watched 2200+ movies and read 200+ books. Interested in history and anthropology (ethnicity, folklore & religion). Write **personal blogs** (in Chinese) to share takeaways, thoughts and insights
- **Basketball**: Play basketball since elementary school. Played for college basketball team. Miami Heat Fan.