Yutong Jin

1518 Clifton Road, GCR 367, Atlanta, GA 30322 | 404-642-2525 | yutong.jin@emory.edu; augparasoljin@gmail.com

EDUCATION

Emory University, Atlanta, GA
Ph.D. in Biostatistics   Aug. 2018 - May 2023 (expected)
M.S. in Public Health, Biostatistics   Aug. 2016 - May 2018
Fudan University, Shanghai, China
B.M. in Clinical Medicine (M.D. Equivalency in the US)   Sept. 2011 - July 2016

TECHNICAL SKILLS

Software: R, Python, SAS, SQL, shell scripting, MS office, LaTeX, git/GitHub, docker, markdown, Jupyter
Selected coursework: causal inference, machine learning, probability theory, statistical inference, generalized linear models, survival analysis, categorical data analysis, stochastic processes, statistical computing, algorithms and data structures, sampling applications

RESEARCH EXPERIENCE

CTMLE for Identifying Antibody-resistant HIV Sequences   Apr. 2020 - present
PhD Researcher, Department of Biostatistics and Bioinformatics, Emory University

- Proposed a new causal method using outcome-adaptive collaborative targeted learning (CTMLE) with Random Forest to (1) efficiently detect meaningful amino acid (AA) positions without manual manipulation from the analysts, (2) stabilize inference for high-dimensional highly correlated data and (3) maintain detection power with better controlled type I error in realistic sample sizes
- Conducted extensive simulations on both small and large samples demonstrating method's performance; showed the advantages of the proposed method in reducing type I error by an average of 50% and decreasing MSE to around 0.5% in both scenarios
- Applied the algorithm to real HIV vaccine data sets from the Compile, Analyze and Tally Nab Panels (CATNAP) database with 200+ Amino Acid positions and 800+ patients
- Identified 7% AA positions that are casually related to the outcome; provided the results for Fred Hutchison Vaccine and Infectious Disease Division to guide a downstream genotypic sieve analysis of phase 2b prevention efficacy trials

Transport of Immunogenicity across Trials with Different Populations and Study Designs   Sept. 2020 - present
PhD Researcher, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center

- Proposed a causal transport method to compare vaccine immunogenicity across different populations using different study designs; our method can handle (1) multiple trials to be standardized, (2) multiple treatment in each trial and (3) different sampling schemes
- Implemented the algorithm from scratch in **R** and validated the standardization with simulated datasets
- Applied the proposed method to four HIV vaccine trials with 1600+ participants, to be extended to evaluating the COVID-19 vaccines developed by Moderna and Pfizer
- Delivered the idea and analytical results in Fred Hutch Vaccine Trial Network Stats meeting to core statisticians

Systematic Evaluation of SNVs on Transcription Factor Binding Affinity   May 2017 – July 2021
Master's Researcher, Department of Biostatistics and Bioinformatics, Emory University

- Proposed a new computational method to quantify the impact of single nucleotide variants (SNVs) on transcription factor (TF) binding based on SVM; experimented on ChIP-Seq data using **R** with 19 TF motifs in GM12878 over the human reference genome GRCh37 with more than 3 million potential binding sites
- Identified a substantial variation in terms of the impact of SNVs on binding affinity; detected about 80% of neutral mutations landed inside canonical motif sites in which all mutations were historically considered detrimental

APPLIED PROJECTS

Investigating whether two anti-tuberculosis drugs induced elongated QT interval   Nov. 2019 - present

- Designed and built a piecewise generalized estimating equation (GEE) model to determine the different effects between two drugs using **R** and **SAS**; independently processed and analyzed medical records including 100 patients with 24-month follow-ups
- Delivered the analytics results to clinical physicians and contributed to the decision on treatments for multidrug-resistant tuberculosis

Validation of Questionnaire Psychometrics from a Cross-sectional Study   June 2021 – Aug. 2021

- Independently preprocessed and analyzed Likert-type questionnaire to provide descriptive statistics and visualizations
- Determined the reliability for each questionnaire and measured the strength of association between similar surveys

Determining the Prevention Effect of Aspirin for Colorectal Adenomas   Dec. 2016 - May 2019

- Constructed the generalized linear mixed model (GLMM) to identify metabolomic signatures that are associated with aspirin treatment in both blood and colon tissue for 1048 patients with over 6000 metabolites from a three-year randomized clinical trial; applied statistical mediation analysis to determine the association between aspirin treatment and colorectal adenoma outcomes for targeted metabolites
- Verified that low-dose aspirin treatment was able to significantly reduce the risk of adenomas; instructed the clinical team to adjust aspirin usage dose in practice

PUBLICATIONS

- Guo, T. A., Wu, Y. C., Tan, C., **Jin, Y. T.,** Sheng, W. Q., Cai, S. J., Liu, F. Q. & Xu, Y. (2019). Clinicopathologic features and prognostic value of KRAS, NRAS, and BRAF mutations and DNA mismatch repair status: a single-center retrospective study of 1834 Chinese patients with stage I-IV colorectal cancer. *International Journal of Cancer*.

- Guo, T., Wu, Y., Huang, D., **Jin, Y.,** Sheng, W., Cai, S., Zhou, X., Zhu, X., Liu, F., & Xu, Y. (2021). Prognostic Value of KRAS Exon 3 and Exon 4 Mutations in Colorectal Cancer Patients. *Journal of Cancer*, 12(17), 5331–5337.

- **Jin, Y.,** Jiang, J., Wang, R. & Qin, Z. S. (2021). Systematic Evaluation of DNA Sequence Variations on in vivo Transcription Factor Binding Affinity. *Frontiers in Genetics* 12:667866. doi: 10.3389/fgene.2021.667866

- **Jin, Y.,** & Benkeser, D. (2021+) Identifying HIV sequences that escape antibody neutralization using random forests and collaborative targeted learning. *Journal of Causal Inference. (under review)*

- Barry, E. L., Fedirko, V., **Jin, Y.,** Liu, K., Mott, L. A., Peacock, J. L., Passarelli, M. N., Baron, J. A. & Jones, D. P. (2021+). Plasma Metabolomics Analysis of Aspirin Treatment and Risk of Colorectal Adenomas. *Cancer Prevention Research (under review)*

- **Jin, Y.,** Moodie, Z., Holly, J., Huang, Y., Luedtke, A., & Benkeser, D. (2021+) Comparing vaccine immunogenicity across trials with different populations and sampling designs. *(in preparation)*

- **Jin, Y.,** Benkeser, D & Kempker, R. (2021+) Investigating the effect of Bedaquiline or Delamanid among Patients with Drug-resistant Tuberculosis on QTc Prolongation. *(in preparation)*

- Wood, K., **Jin, Y.,** & Krafty, R. T. (2021+). The validation of the Patient Perception of Arrhythmia Questionnaire. *(in preparation)*