

Yutong Jin

1518 Clifton Road, GCR 367, Atlanta, GA 30322 | 404-642-2525 | yutong.jin@emory.edu | [Personal Website](#) | [Google Scholar](#)

EDUCATION

Emory University, Atlanta, GA

Ph.D. in Biostatistics

2018 - Present

M.S. in Public Health, Biostatistics

2016 - 2018

Fudan University, Shanghai, China

B.M. in Clinical Medicine (M.D. Equivalency in the US)

2011 - 2016

PROFESSIONAL SKILLS

Research Interest: Causal Inference, machine learning, infectious disease, HIV/AIDS prevention, vaccine development

Software: R, Python, SAS, SQL, shell scripting, LaTeX, Git/GitHub, docker, markdown, Jupyter, MS office

Selected coursework: causal inference, machine learning, probability theory, statistical inference, generalized linear models, artificial intelligence, survival analysis, categorical data analysis, stochastic processes, algorithms and data structures, sampling applications

PUBLICATIONS

Statistical Methodology:

[3] **Jin, Y.**, Luedtke, A., & Benkeser, D. (2022+). Comparing Vaccine Immunogenicity across Trials with Different Populations and Study Designs. (To be submitted)

[2] **Jin, Y.** & Benkeser, D. (2022). Identifying HIV sequences that escape antibody neutralization using random forests and collaborative targeted learning. *Journal of Causal Inference*. (Accepted)

[1] **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*. doi.org/10.3389/fgene.2021.667866

Applied:

[10] **Jin, Y.**, Li, J. & Du, Y. (2022+). Improved Inference for Binary Outcome Dichotomized from Continuous Variable with Super Learner and Targeted Learning. (In preparation)

[9] **Jin, Y.**, Benkeser, D., Kipiani, M., Mikiashvili, L., Barbakadze, K., Avaliani, Z., Alghamdi, W., Al-Shaer, M.H., Peloquin, C.A., Blumberg, H.M., & Kempker, R.R. (2022+). The impact of QTc prolonging drugs and anti-tuberculosis pharmacokinetics including bedaquiline and delamanid on QTc prolongation. (To be submitted)

[8] Wood, K., **Jin, Y.**, & Krafty, R. T. (2022+). The validation of the Patient Perception of Arrhythmia Questionnaire. (In preparation)

[7] Zhao, Y., **Jin, Y.**, Kempker R, Kipiani M, & Benkeser, D. (2022). Evaluating Doubly Robust, Machine Learning-based Approaches for Estimating Treatment Effects in Small Observational Studies. *Epidemiology*. (Under revision)

[6] Shi, X., Cai, Y., Zhou, M., **Jin, Y. T.**, Sun, G. S., Zhao, L., Han, F., Qu, J. F. & Zhao, M. W. (2022). Diurnal changes of retinal microvascular circulation and RNFL thickness measured by optical coherence tomography angiography in patients with obstructive sleep apnea-hypopnea. *Frontiers in Endocrinology*. <https://doi.org/10.3389/fendo.2022.947586>

[5] Moodie, Z., Dintwe, O., Sawant, S., Grove, D., Huang, Y., Janes, H., Heptinstall, J., Laher Omar, F., Cohen, K., De Rosa, S. C., Zhang, L., Yates, N. L., Sarzotti-Kelsoe, M., Seaton, K. E., Laher, F., Bekker, L. G., Malahleha, M., Innes, C., Kassim, S., Naicker, N., Govender, V., Sebe, M., Singh, N., Kotze, P., Lazarus, E., Nchabeleng, M., Ward, A. M., Brumskine, W., Dubula, T., Randhawa, A. K., Grunenberg, N., Hural, J., Kee, J. J., Benkeser, D., **Jin, Y.**, Carpp, L. N., Allen, M., D'Souza, P., Tartaglia, J., DiazGranados, C., Koutsoukos, M., Gilbert, P., Kublin, J., Corey, L., Andersen-Nissen, E., Gray, G., Tomaras, G. & McElrath, M. J. (2022). Analysis of the HVTN 702 phase 2B-3 HIV-1 vaccine trial in South Africa assessing RV144 antibody and T-cell correlates of HIV-1 Acquisition

Risk. *The Journal of Infectious Diseases*. <https://doi.org/10.1093/infdis/jiac260>

[4] Barry, E.L., Fedirko, V., **Jin, Y.**, Liu, K., Mott, L.A., Peacock, J.L., Passarelli, M.N., Baron, J.A. & Jones, D.P. (2022). Plasma Metabolomics Analysis of Aspirin Treatment and Risk of Colorectal Adenomas. *Cancer Prevention Research*. doi.org/10.1158/1940-6207.CAPR-21-0555

[3] Guo, T. A.*, Lai, H. Y.*, Li, C., Li, Y., Li, Y. C., **Jin, Y. T.**, Zhang, Z.Z., Huang, H.B., Huang, S.L. & Xu, Y. (2022). Plasma Extracellular Vesicle Long RNAs Have Potential as Biomarkers in Early Detection of Colorectal Cancer. *Frontiers in Oncology*, 1251. doi.org/10.3389/fonc.2022.829230

[2] 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), 5 331 5337. doi.org/10.7150/jca.59193

[1] 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 single-center retrospective study of 1834 Chinese patients with stage I I-IV colorectal cancer. *International Journal of Cancer*. doi.org/10.1002/ijc.32489

Research Experience

Eli Lilly and Company

2022

Research Scientist Summer Intern

- Proposed a framework that fully utilizes continuous measurements to improve the performance of estimation with valid statistical inference for dichotomous outcomes.

Fred Hutchinson

2020 - Present

Research Assistant

- Proposed a framework to identify appropriate causal estimands and estimators that can be used to provide standardized comparisons of vaccine immunogenicity across trials with different populations and study designs.

Fudan University

2018 - Present

Statistical Consultant

- Performed statistical data analyses in the biomarker assessment of patients with colorectal cancer for several studies with focus on but not limited to survival data.
- Held regular meetings with collaborative scientists to discuss analysis plans and results; prepared written results for publication in peer-reviewed journals.

Emory University

2016 - Present

Research Assistant

- Proposed a new causal method using collaborative targeted learning (CTMLE) with random forests to identify mutations that potentially lead to a resistance to particular antibodies among high-dimensional strong correlated HIV genetic sequence data.
- Proposed a new computational method to quantify the impact of single nucleotide variants (SNVs) on transcription factor (TF) binding using SVM based on 19 TF motifs in GM12878 over the human reference genome GRCh37.
- Conducted data analysis on Likert-type questionnaires to determine the reliability for each questionnaire and the strength of association between similar surveys.
- Designed and built a piecewise generalized estimating equation (GEE) model to determine the relationship between anti-tuberculosis drugs and QTc prolongation.
- Constructed the generalized linear mixed model (GLMM) to identify metabolomic signatures that are associated with aspirin treatment in both blood and colon tissue from a three-year randomized clinical trial.

TEACHING AND MENTORING

Emory University

2018 - Present

Teaching Assistant

- Responsibilities include hosting review sessions, preparing homework keys, holding office hours, and grading assignments.
 - BIOS 510 Probability Theory (Master core course): 36 students.

- BIOS 511 Introduction to Statistical Inference (Master core course): 23 students.
- BIOS 526 Modern Regression Analysis (Master core course): 32 students.
- BIOS 745R Biostatistical Consulting (PhD core course): 10 students.

Mentoring

- Mentored an undergraduate student on summer research at Emory, 2019.
- Mentored a Master student on thesis project at Emory, 2019 – 2021.