Naishu Kui

Contact Information

Department of Biostatistics
University of Texas M.D. Anderson Cancer Center
1400 Pressler Street
Houston, Texas 77030

Tel: (347) 413-0863
E-mail: nkui@mdanderson.org

Education

University of Texas Health Science Center at Houston , Houston, Texas

Ph.D., Biostatistics and Data Science, September 2022 - Present

- Dissertation Title: "Large Impact of Genetic Data Processing Steps on Reproducibility of Set-Based Analyses in Genome-Wide Association Studies"
- Advisor: Professor Ryan Sun

New York University, New York, New York

M.S., Biostatistics, May 2018

Southeast University, Nanjing, China

B.E., Pharmaceutical Engineering, May 2016

Honors and Awards

Robert H. Bigelow Endowed Scholarship, 2025

UTHealth School of Public Health Travel Award, 2024

UTHealth Outstanding New Student Doctoral Scholarship Award, 2022

NYU School of Global Public Health New Student Scholarship, 2016 - 2018

Academic Experience

M.D. Anderson Cancer Center, Houston, Texas

Graduate Research Assistant/Ph.D. Trainee **September 2022 - Present**

- Integrated high-dimensional functional annotation data into GWAS pipelines for whole-exome sequencing (WGS) studies to identify rare germline genetic variants associated with pancreatic cancer risk
- Analyzed large-scale human genetics data from biobanks using parallel computing on high-performance computing (HPC) clusters
- Improved the theoretical power calculation of set-based association tests by deriving the analytical power for the Sequence Kernel Association Test (SKAT)
- Developed an open-source R package to demonstrate the impact of different data processing choices through sensitivity analysis for study design
- Analyzed cancer-specific transcriptomic profiles from clinical liquid biopsy samples using next-generation sequencing (NGS), integrating deconvolution techniques and machine learning for classification

Professional Experience

Icahn School of Medicine at Mount Sinai, New York, New York

Biostatistician II **November 2021 - March 2022**

- Developed statistical analysis plans (SAPs) and timelines for clinical research projects
- Performed complex statistical analyses, including survival analysis, logistic regression, and data imputation, using SAS and R

- Collaborated with principal investigators and external partners to ensure accurate project delivery and reporting

Baim Institute for Clinical Research , Boston, Massachusetts

Biostatistician II **January 2021 - November 2021**

- Led statistical analysis for clinical research projects, specializing in cardiovascular clinical trials
- Applied machine learning techniques and survival models to longitudinal datasets for clinical outcome assessments
- Conducted data cleaning, validation, and visualization to generate reports for FDA submissions
- Reviewed and contributed to case report forms (CRFs) and study protocols

Statistical Programmer **March 2019 - January 2021**

- Lead programmer for multi-center clinical trials, performing data analysis and quality control in compliance with CDISC SDTM and ADaM standards
- Generated Tables, Listings, and Graphs (TLGs) for clinical study reports and FDA submissions using SAS macros
- Developed efficient SAS programming scripts for large-scale data integration and statistical outputs

Software Packages

SetDesign: An R package for performing analytical power calculation and bias estimation under model misspecification for sequence kernel association test in genetic association studies. It can be used at study design stages to study the impact of potential data processing choices.

Relevant Coursework

- Statistical Inference
- Generalized Linear Models
- Multivariate Analysis
- Missing Data
- Bayesian Data Analysis
- Categorical Analysis
- Survival Analysis
- Linear Models
- Statistical Computing
- Supervised and Unsupervised Machine Learning
- Stochastic Process

Skills

- SAS® Certified Advanced Programmer for SAS 9 (MACRO/SQL/STAT)
- R
- Python
- SQL
- Linux
- Bash
- High Performance Computing
- Git/GitHub

Publications

1. **Kui, N.**, Zhu, H., Sun, R. (2025). Formal statistical replication analysis of pancreatic cancer genome-wide association studies identifies high-quality risk variants. Poster presented at the *American Society of Human Genetics (ASHG) Annual Meeting.*

2. **Kui, N.**, Yu, Y.,Li, X., Huff, C., Sun, R. (2025). Large Impact of Genetic Data Processing Steps on Reproducibility of Set-Based Analyses in Genome-Wide Association Studies. Under review at *Genetics.*

3. **Kui, N.**, Yu, Y., Scheet, P., Li, D., Huff, C., Sun, R. (2024). Integration of high-dimensional functional annotation data to identify rare germline genetic variants contributing to pancreatic cancer risk. Poster presented at the *Joint Statistical Meetings (JSM).*

4. Zhu, H., Choi, J., **Kui, N.**, Yang, T., Wei, P., Li, D., Sun, R. (2024). Identification of pancreatic cancer germline risk variants with effects that are modified by smoking. *JCO Precision Oncology.*

5. Bahrambeigi, V., Lee, J. J., Branchi, V., Rajapakshe, K. I., Xu, Z., **Kui, N.**, Dhebat, S. (2024). Transcriptomic profiling of plasma extracellular vesicles enables reliable annotation of the cancer-specific transcriptome and molecular subtype. *Cancer Research*.

6. Jallouk, A. P., **Kui, N.**, Sun, R., Westin, J. R., Steiner, R. E., Nair, R., Nastoupil, L. J., Fayad, L. E., Al Zaki, A., Hawkins, M., Noorani, M., Das, K., Henderson, J., Shpall, E. J., Kebriaei, P., Ramdial, J., Flowers, C. R., Ahmed, S., Strati, P. (2024). Effect of delayed cell infusion in patients with large B-cell lymphoma treated with chimeric antigen receptor T-cell therapy. *Haematologica*.

7. Antonoff, M. B., **Kui, N.**, Sun, R., et al. (2023). Factors associated with receipt of pulmonary metastasectomy in patients with lung-limited metastatic colorectal cancer: Disparities in care and impact on overall survival. *Journal of Cardiac Failure*.

8. Abboud, A., **Kui, N.**, Gaggin, H. K., Ibrahim, N. E., et al. (2021). Multiple cardiac biomarker testing among patients with acute dyspnea from the ICON-RELOADED study. *Journal of Heart Failure*.

9. Ballantyne, C., Bhatt, D., de Lemos, J., Gao, Q., **Kui, N.**, Rosenson, R., et al. (2022). GOULD EDU: Primary results of a cluster-randomized trial of an educational intervention to improve guideline adherence and intensification of lipid-lowering therapy. *Journal of Clinical Lipidology*.

10. Ducrocq, G., Bhatt, D., Lee, J., **Kui, N.**, et al. (2022). Balance of benefit and risk of Ticagrelor in patients with diabetes and stable coronary artery disease according to bleeding risk assessment with the CRUSADE score: Data from THEMIS and THEMIS PCI. *American Heart Journal*.

11. Peterson, B., Bhatt, D., Ballantyne, C., de Lemos, J., Exter, J., Alam, S., **Kui, N.**, et al. (2021). TCT-250 Differential intensity of lipid-lowering therapy among patients with and without previous coronary revascularization—Insights from GOULD. *Journal of the American College of Cardiology*.

12. **Kui, N.**, Goldmann, E., Parikh, N. S., Boden-Albala, B. (2018). Risk perception in a multi-ethniccohort of stroke survivors. Poster presented at the *International Stroke Conference*.