Haoran Hu

Phone:646-479-8979   *Email: hah112@pitt.edu*

QUALIFICATION SUMMARY

- Proficient in statistical programming with R and Python; familiar with SAS and STATA
- Experienced in effectively collaborating with non-statisticians in the medical field
- Strong background in statistical machine learning and statistical modeling in single-cell multi-omics data
- Theoretical foundation and practical experience in longitudinal data analysis, survival analysis, and clinical trial design

EDUCATION

University of Pittsburgh, School of Public Health, Pittsburgh, PA   Sep 2021 – present
Ph.D. student in Biostatistics

- **Advisor**: Prof. Wei Chen, PhD
- **Research interest**: Statistical machine learning, Statistical modeling, Multi-omics data analysis
- **Cross-registered courses at Carnegie Mellon University**: Introduction to Machine Learning (CMU-10601), Advanced Deep Learning (CMU-10707)

Columbia University, Mailman School of Public Health, New York, NY   Sep 2018 – May 2020
Master of Science in Biostatistics;

- **GPA**: 4.00/4.00
- **Selected courses taken**: Introduction to Randomized Clinical Trial (BIOST-8140), Clinical Trial Methodology (BIOST-8142), Statistical Methods for Causal Inference (BIOST-8122)

South China University of Technology (SCUT), Guangzhou, China   Sep 2014 – Jun 2018
Bachelor of Science in Mathematics and Applied Mathematics;

- **GPA**: 3.78/4.00, Top 4 among 50 students

WORK EXPERIENCE

University of Pittsburgh Medical Center (UPMC) Division of Pulmonary, Pittsburgh, PA
          Sep 2022 – present
Graduate Student Researcher
Supervisor: Prof. Sally E. Wenzel

- Coordinated single-cell data collection from 8 severe asthma patients and 5 healthy controls
- Conducted statistical analyses on gene expression across disease statuses and authored comprehensive reports to support data-driven insights
- Delivered monthly presentations to interdisciplinary teams, contributing to research strategy

UPMC Division of Gastroenterology, Pittsburgh, PA   Jan 2022 – present
Graduate Student Researcher
Supervisor: Prof. Richard H. Duerr

- Established statistical analysis pipelines for cutting-edge single-cell multi-omics data, such as DOGMA-seq data and RNA isoform data from long-read sequencing
- Investigated factors promoting or inhibiting CD4 T-cell activation in Crohn's disease patients

UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA   Sep 2021 – present
Graduate Student Researcher
Supervisor: Prof. Wei Chen
- Developed novel statistical methods for analyzing single-cell multi-omics data

Columbia University, Department of Biostatistics, New York, NY   Aug 2019 – Apr 2020
Research Assistant
Supervisor: Prof. Zhezhen Jin
- Investigated the relationship of education level and Apoe4 gene status with the rate of decline in cognitive level

PEER-REVIEWED PAPERS
^: co-first author
Methodology
- **Hu H**, Wang X, Feng S, Xu Z, Liu J, O'Hare E, Chen Y, Yue M, Zeng L, Rong Z, Chen T, Billiar T, Ding Y, Huang H, Duerr R, Chen W. A unified model-based framework for doublet or multiplet detection in single-cell multiomics data. *Nature Communications* 15, 5562 (2024). https://doi.org/10.1038/s41467-024-49448-x
  Selected as a featured article in "Biotechnology and methods" section
  https://www.nature.com/collections/idhhgedgig

- Chen Y, Wu X, Ni K, **Hu H**, Yue M, Chen W, Huang H. Robust and Accurate Doublet Detection of Single-Cell Sequencing Data via Maximizing Area Under Precision-Recall Curve. bioRxiv; **doi:** https://doi.org/10.1101/2023.10.30.564840

- Wang X, Xu Z, **Hu H**, Zhou X, Zhang Y, Lafyatis R, Chen K, Huang H, Ding Y, Duerr R, Chen W. SECANT: a biology-guided semi-supervised method for clustering, classification, and annotation of single-cell multi-omics. *PNAS Nexus*. 2022 Sep; 1(4): pgab165

- Rong Z^, **Hu H**^, Duan J, Liu T, Feng S, Zhao C, Duerr, R, Zhou X, Chen W. scDitu: a cluster-free method for identifying isoform-level differential transcript usage in sparse single-cell long-read sequencing data. Under review at *RECOMB* conference

- Yue M, Cai M, Zhao C, Tao S, Liu J, **Hu H**, Chen Y, Celedon J, Wang J, Huang H, Chen W. Multi-Omics intermediate Fusion Enable Digital White Blood Cells Count Prediction. Submitted to *Genome Medicine*

- Yuan R, Rong Z, **Hu H**, Liu T, Tao S, Chen W. Harmony-Based Data Integration for Distributed Single-Cell Multi-Omics Data. Submitted to *Nature Communications*

Collaborative work

- Xu Z, Tao S, Feng S, Wang T, **Hu H**, O'Hare E, Chen W, Duerr R. Dissecting PGE2-Driven inhibition of T Cell Activation Using Single-Cell Multi-Omic and Inflammatory Bowel Disease Genetic Association Analysis. *Nature Genetics (In revision)*, 2024

BOOK CHAPTER

- Wang X, **Hu H**, Chen W. Model-based Clustering of Single-cell Data. Book Chapter In: Handbook of Statistical Bioinformatics (2nd Edition) 2022 (pp. 85-108). Chapman & Hall/CRC.

PRESENTATIONS

- (Invited talk) "A unified model-based framework for doublet or multiplet detection in single-cell multiomics data", Joint Statistical Meetings, Portland, OR, August 2024
- (Contributed talk) "COMPOSITE: A Novel Compound Poisson Model for Enhanced Multiplet Detection in Single-Cell Multiomics Data", Biostatistics Research Day, University of Pittsburgh, Pittsburgh, PA, February 2024
- (Contributed talk) Exploration: Can SECANT clustering uncertainty benefit downstream trajectory inference? Biostatistics Research Day, University of Pittsburgh, Pittsburgh, PA, February 2022

HONORS AND AWARDS

University of Pittsburgh Biostatistics Research Day Honorable Mentions	2022
Outstanding graduate award of school of mathematics (SCUT)	2018
"Hongping Changqing" Science and Technology Innovation Award	2018
Merit Student in South China University of Technology	2015-2017
"Meritorious Winner" in the COMAP's Mathematical Contest in Modeling (MCM)	2016
Macao alumni scholarship (SCUT)	2016
Excellent leader of the student organizations (SCUT)	2015

PROGRAMMING SKILLS

R, Python, MATLAB - Frequently used in research projects
Stata, SAS, SQL, Mathematica, C++, C# - Frequently applied in coursework

SERVICES

- Organizer, University of Pittsburgh Biostatistics Journal Club, Spring 2024
- Teaching Assistant, University of Pittsburgh Next Generation Sequencing Workshops, Fall 2023
- Teaching Assistant, University of Pittsburgh Next Generation Sequencing Workshops, Spring 2023

PROFESSIONAL AFFILIATIONS

- Member of American Statistical Association (ASA)
- Member of Eastern North American Region International Biometric Society (ENAR)
- Member of International Chinese Statistical Association (ICSA)