# Penghui Huang

130 De Soto St, Pittsburgh, PA 15261, USA
peh65@pitt.edu | Personal Website | Google Scholar | GitHub

## EDUCATION

**University of Pittsburgh**, Pittsburgh, PA — Aug 2022 – Nov 2025 (Expected)

Ph.D. in Biostatistics (Advisor: Dr. Chris McKennan)
Research focus: Statistical genomics, machine learning, Bayesian methods, multi-omics integration

**University of Science and Technology of China**, Hefei, China — Sep 2018 – Jun 2022

B.S. in Statistics

## EMPLOYMENT

**Graduate Student Researcher, University of Pittsburgh** — Aug 2022 – Present

- Research in statistical genomics and bioinformatics, focusing on multi-omics integration, cfDNA, single-cell and spatial transcriptomics.
- Developed novel methods and software: HiDecon, BLEND, cf-TREBLE, EMixed.
- Collaborated with clinicians and biologists in projects on Alzheimer's, asthma, and reproductive health.

**Teaching Assistant, University of Pittsburgh** — Aug 2023 – Present

- Assisted instruction in graduate-level courses: Causal Inference, Omics Data Analysis, Mixed Models, Meta-Analysis, Consulting Practicum, and Computer Methods in Clinical Research.
- Instructed MS/PhD students, MDs, and medical students; supervised consulting sessions.

## RESEARCH INTERESTS

Statistical genomics, machine learning, Bayesian modeling, multi-omics integration, single-cell and spatial transcriptomics, cfDNA, causal inference

## PUBLICATIONS

*Note: '_' indicates first author, '†' indicates co-corresponding author.*

[1] **Huang, Penghui**, Cai, M., McKennan, C., Wang, J. "Accurate estimation of rare cell-type fractions from tissue omics data via hierarchical deconvolution." *Annals of Applied Statistics*, 18(2), 1178–1194 (2024).

[2] **Huang, Penghui**†, Cai, M., McKennan, C., Wang, J. "BLEND: Probabilistic Cellular Deconvolution with Individualized Single-Cell Reference Integration." *Genome Biology*, (2025).

[3] Cai, M., Zhao, K., **Huang, P.**, et al. "EMixed: Probabilistic Multi-Omics Cellular Deconvolution of Bulk Omics Data." *Journal of Data Science*, (2025).

[4] **Huang, Penghui**, et al. "cf-TREBLE: Tree-based Bayesian framework for cfDNA deconvolution and biomarker discovery." (Under review).

## SOFTWARE

- HiDecon: R package for rare cell-type fraction estimation GitHub
- BLEND: R package for bulk RNA-seq deconvolution GitHub

## PRESENTATIONS

**TALKS**

- ICSA Applied Statistics Symposium — 2023
- New England Statistics Symposium — 2023

**POSTERS**

- ICSA Applied Statistics Symposium — 2023
- ENAR Spring Meeting — 2023
- Pitt Biostatistics Research Day — 2023
- ASA Pittsburgh Spring Banquet — 2023

## AWARDS

Student Research Award, New England Statistics Symposium (2023)
Research Poster Award, ICSA Symposium (2023)
Best PhD Poster Award, University of Pittsburgh (2023)

## ACADEMIC AFFILIATIONS

American Statistical Association (ASA), ENAR, ICSA