Boyi Hu

CONTACT INFORMATION

Address: Columbia University Irving Medical Center, 622 West 168th Street, New York, NY 10032, USA
Email: bh2940@cumc.columbia.edu
Homepage: https://boyihu673.github.io

RESEARCH INTERESTS

Machine learning, high-dimensional multi-omics data integration, functional data analysis, semi-parametric inference, empirical likelihood

TRAINING & EDUCATION

Postdoctoral Research Scientist in Biostatistics 11/2023 - PRESENT
COLUMBIA UNIVERSITY IRVING MEDICAL CENTER
New York, NY 10032, USA

- Advisors: Professor Yuanjia Wang and Professor Annie J. Lee

Ph.D. in Statistics 09/2018 - 08/2023
SIMON FRASER UNIVERSITY
Burnaby, British Columbia, Canada

- Supervisor: Professor Jiguo Cao
- Thesis: Functional regression models

M.Sc. in Statistics 09/2016 - 08/2018
UNIVERSITY OF BRITISH COLUMBIA
Vancouver, British Columbia, Canada

- Supervisor: Professor Jiahua Chen
- Thesis: An R package for monitoring test under density ratio model and its applications

M.Sc. in Mathematical Science 09/2014 - 06/2016
LAKEHEAD UNIVERSITY
Thunder Bay, Ontario, Canada

- Supervisor: Professor Deli Li

B.Sc. in Applied Mathematics 09/2009 - 06/2013
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
Hefei, Anhui, China

HONORS & AWARDS

- **Finalist**, ARISE (Aging Research – Innovations in Statistical Exploration) Program, ASA Statistics and Data Science in Aging Interest Group — 2025
- **AAIC Fellowship**, Alzheimer's Association International Conference (AAIC) — 2025
- **Finalist**, Poster Competition, 14th Annual Taub Institute Research Retreat (Columbia University) — 2024
- **Graduate Fellowship**, Simon Fraser University — Summer 2023
- **Graduate Fellowship**, Simon Fraser University — Spring 2023
- **Graduate Fellowship**, Simon Fraser University — Summer 2022
- **Graduate Fellowship**, Simon Fraser University — Summer 2020
- **Big Data Graduate Scholarship**, Simon Fraser University — Fall 2019
- **Graduate Fellowship**, Simon Fraser University — Summer 2019

- **Mitacs Accelerate Fellowship** 2020–2021
- **International Tuition Award**, University of British Columbia 2016

PUBLICATIONS & PREPRINTS

A. Original, peer-reviewed articles (In chronological order)

- **Hu, B.**, & Cao, J. (2026). Locally sparse estimation for simultaneous functional quantile regression. Accepted by *Journal of Computational and Graphical Statistics*. *arXiv*. https://arxiv.org/abs/2602.01691
- **Hu, B.**, Liu, H., You, J., & Cao, J. (2026). Convolution smoothing–based locally sparse estimation for functional quantile regression. *Accepted by Journal of Agricultural, Biological, and Environmental Statistics*. *arXiv*. https://arxiv.org/abs/2512.01341
- **Hu, B.**, Vardarajan, B., De Jager, P. L., Bennett, D., Wang, Y., & Lee, A. (2025). *TPClust: Temporal Profile–Guided Disease Subtyping Using High-Dimensional Omics Data*. Under review. *bioRxiv*. https://doi.org/10.1101/2025.08.05.668514.
- Touil, H., Luquez, T., Comandante-Lou, N., Lee, A. J., Fujita, M., Habeck, C., Kroshilina, A., Hegewisch-Solloa, E., Pedroza, L. A., McInvale, J., Zuroff, L., Isnard, S., Majumder, B., Walker, E., Zhang, L., Das, A., Routy, J.-P., **Hu, B.**, Zhang, Y., Reyes-Dumeyer, D., Mace, E. M., Klotz, L., Wiendl, H., Orange, J., Xia, Z., Bar-Or, A., Menon, V., Mayeux, R., Stern, Y., & De Jager, P. L. (2025). *Trajectories of immunosenescence highlight age-associated accumulation of PD-L1$^+$ NK cells and association with brain atrophy*. Under review.
- **Hu, B.**, Hu, X., Liu, H., You, J., & Cao, J. (2024). Simultaneous functional quantile regression. *Statistica Sinica*, 34(2).
- Zhuang, W., **Hu, B.**, & Chen, J. (2019). Semiparametric inference for the dominance index under the density ratio model. *Biometrika*, 106(1), 229–241.

B. Conference Abstracts

- **Hu, B.**[1], Vardarajan, B.[1], De Jager, P. L.[1], Bennett, D.[1], Wang, Y.[1], & Lee, A.[1]. *TPClust: Temporal Profile–Guided Disease Subtyping Using High-Dimensional Omics Data*. Alzheimer's Association International Conference (AAIC) 2025, Toronto, Canada. `https://alz.confex.com/alz/2025/meetingapp.cgi/Paper/103996`. [1]Columbia University, New York, NY.
- Lee, A.[1], **Hu, B.**[1], Lu, J.[1], Bennett, D.[2], & Vardarajan, B.[1]. *Novel short tandem repeats on the telomere-to-telomere reference genome are associated with Alzheimer's disease neuropathology*. American Society of Human Genetics (ASHG) 2024, Denver, Colorado. `https://www.ashg.org/wp-content/uploads/2024/10/ASHG2024-PlatformAbstracts.pdf`. [1]Columbia University, New York, NY; [2]Rush University Medical Center, Chicago, IL.

C. Manuscripts in preparation

- **Boyi Hu**, Yuanjia Wang and Annie J. Lee (2025). MV-TPClust: Multivariate Temporal Profile–Guided Clustering with High-dimensional Multi-omics Data.
- Annie J. Lee, **Boyi Hu**, Jinfeng Lu, David A. Bennett and Badri N. Vardarajan (2025). Novel Short Tandem Repeats on the Telomere-to-Telomere Reference Genome are Associated with Alzheimer's Disease Neuropathology.

TALKS & PRESENTATIONS

Invited Talks:

- Invited talk at **ARISE (Aging Research – Innovations in Statistical Exploration) Fall 2025 Webinar Series**: Temporal Profile-Guided Subtyping Using High-Dimensional Omics Data. 09/2025
- Invited talk at **2025 Lifetime Data Science (LiDS) Conference**: Novel Machine Learning Method Identifies Alzheimer's Disease Subtypes Using Longitudinal Clinical Data and High-dimensional Omics Data. 05/2025
- Invited seminar talk at **Department of Mathematics and Statistics, York University**: Functional Regression Models. (remote via Zoom) 10/2023
- Invited seminar talk at **School of Mathematics and Statistics, Carleton University**: Functional Quantile Regression. 04/2023
- **International Workshop on Complex Functional Data Analysis**: Simultaneous Functional Quantile Regression. (remote via Zoom) 06/2022

Contributed Talks:

- **2025 International Conference on Alzheimer's and Parkinson's Diseases and Related Neurological Disorders (AD/PD)**: Novel Machine Learning Method Identified Alzheimer's Disease Subtypes Using Longitudinal Clinical Data and High-dimensional Omics Data. (Virtual) 04/2025
- **2024 Joint Statistical Meetings (JSM)**: A Semi-Parametric Approach for Longitudinal Outcome-Guided Disease Subtyping Using High-Dimensional Omics Data. 08/2024
- **2023 Statistical Society of Canada (SSC) Annual Meeting**: Simultaneous Functional Quantile Regression. 05/2023
- **2022 Statistical Society of Canada (SSC) Annual Meeting**: Convolution Smoothed Semi-parametric Quantile Functional Linear Regressions with Locally Sparse Adaptation. 06/2022
- **2021 Canadian Statistical Sciences Institute (CANSSI) Showcase**: Simultaneous Functional Quantile Regression. 11/2021

Posters:

- **2025 Alzheimer's Association International Conference (AAIC)**: TPClust: Temporal Profile–Guided Disease Subtyping Using High-Dimensional Omics Data. Toronto, Canada. 07/2025
- **2025 Joint Statistical Meetings (JSM)**: TPClust: Temporal Profile-Guided Disease Subtyping Using High-Dimensional Omics Data. 08/2025
- **14th Annual Taub Institute Research Retreat**: Alzheimer's Disease Subtyping Using Longitudinal Clinical Data and Brain Transcriptome. 09/2024
- **2024 Columbia University Postdoctoral Research Symposium**: A Semi-Parametric Approach for Longitudinal Outcome-Guided Disease Subtyping Using High-Dimensional Omics Data. 05/2024
- **Case Study Poster Session at 2018 Joint Statistical Meeting (JSM)**: Uncertainty Quantification of Weather Forecasts. 07/2018

MENTORING & TEACHING EXPERIENCE

Mentoring Experience for Master's Students' Practicum
Masters' students, Department of Biostatistics, Columbia University

- Yaduo Wang 09/2024 - 05/2025
- Xuesen Zhao 01/2024 - 05/2024
- Zuoqiao Cui 01/2024 - 05/2024
- Zhengwei Song 01/2024 - 05/2024

Guest Lecturer, Columbia University

- Seminar Series: Statistical and Computational Methods in Neurodegenerative Disease Research 11/2025
- BIST P9186-Statistical Practice and Research for Interdisciplinary Sciences (SPRIS) 04/2025
- BIST P9186-Statistical Practice and Research for Interdisciplinary Sciences (SPRIS) 04/2024

Teaching Assistant, Simon Fraser University
Delivered tutorial lectures; led weekly labs and office hours; prepared and graded assignments and exams.

- STAT 380-Intro to Stochastic Process 01/2022 - 04/2022
- STAT 330-Introduction to Mathematical Statistics 09/2021 - 12/2021
- STAT 831-Statistical Theory 01/2020 - 04/2020
- Statistics Workshop 09/2019 - 12/2019
- Statistics Workshop 05/2019 - 08/2019
- STAT 380-Intro to Stochastic Process 01/2019 - 04/2019
- Statistics Workshop 09/2018 - 12/2018

Teaching Assistant, University of British Columbia
Held weekly labs and office hours, marked assignments and exams

- STAT 344-Sample Surveys 09/2017 - 12/2017
- STAT 251-Introductory Probability and Statistics 01/2017 - 04/2017
- STAT 300-Intermediate Statistics for Applications 09/2016 - 12/2016

PROFESSIONAL EXPERIENCE

Mitacs Accelerate Internship 11/2020 - 06/2021
SIMON FRASER UNIVERSITY
Burnaby, British Columbia, Canada

- Developed a machine learning-based wine recommendation engine for Quini, a startup company, as part of the Mitacs Accelerate program
- Advisor: Professor Jiguo Cao

PROFESSIONAL ACTIVITIES

Poster Judge

- 2024 Columbia University Postdoctoral Research Symposium

Reviewer

- Biometrics
- Statistics in Medicine
- Journal of Agricultural, Biological, and Environmental Statistics
- Stat

Volunteer

- Volunteer at 2018 Joint Statistical Meetings (JSM), Vancouver, British Columbia, Canada
- Volunteer at International Chinese Statistical Association (ICSA)-Canada Chapter 2017 Symposium, Vancouver, British Columbia, Canada