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, large-scale multi-omics data, precision medicine, functional data analysis, semi-parametric models, empirical likelihood

TRAINING & EDUCATION

Postdoctoral Research Scientist in Biostatistics 10/2023 - 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; selected for presentation in the ARISE Webinar Series (ASA Statistics and Data Science in Aging Interest Group)	2025
Alzheimer's Association International Conference 2025 (AAIC) Fellowship	2025
Finalist, Poster competition, *14th Annual Taub Institute Research Retreat*	2024
Simon Fraser University Graduate Fellowship	2022 - 2023
Mitacs Accelerate Fellowship	2020 - 2021
Simon Fraser University Big Data Graduate Scholarships	2019
University of British Columbia International Tuition Award	2016

PUBLICATIONS & PREPRINTS

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

Statistical Methodology Publications:

1. **Boyi Hu** and Jiguo Cao (2024). "Locally Sparse Estimation for Simultaneous Functional Quantile Regression". Revision invited by *Journal of Computational and Graphical Statistics*, resubmitted.

2. **Boyi Hu**, Hua Liu, Jinhong You and Jiguo Cao (2024). "Convolution Smoothing based Locally Sparse Estimation for Functional Quantile Regression". Revision invited by *Journal of Agricultural, Biological, and Environmental Statistics*, resubmitted.

3. **Boyi Hu**, Xixi Hu, Hua Liu, Jinhong You and Jiguo Cao (2022). "Simultaneous Functional Quantile Regression". Accepted by *Statistica Sinica*. https://doi.org/10.5705/ss.202021.0248

4. Weiwei Zhuang, **Boyi Hu** and Jiahua Chen (2019). "Semiparametric Inference for the Dominance Index under the Density Ratio Model". *Biometrika*. 106(1), 229-241.

B. Preprints

Substantive Area Publications:

1. **Boyi Hu**, Badri N. Vardarajan, Philip L. De Jager, David A. Bennett, Yuanjia Wang and Annie J. Lee (2025). TPClust: Temporal Profile-Guided Disease Subtyping Using High-Dimensional Omics Data. Submitted to *Nature Methods*. https://www.biorxiv.org/content/early/2025/08/07/2025.08.05.668514

C. Working papers

1. **Boyi Hu**, Jinfeng Lu, David A. Bennett, Annie J. Lee and Badri N. Vardarajan (2025). Novel Short Tandem Repeats on the Telomere-to-Telomere Reference Genome are Associated with Alzheimer's Disease Neuropathology.

2. **Boyi Hu**, and Jiahua Chen (2025). Non-Parametric and Data-Driven Basis Function Estimation Under the Density Ratio Model.

3. **Boyi Hu**, and Jiguo Cao (2025). A Semi-Parametric Functional Generalized Linear Model with Density Ratio Structure.

D. Posters

1. "TPClust: Temporal Profile-Guided Disease Subtyping Using High-Dimensional Omics Data". *2025 Joint Statistical Meetings (JSM)*, Nashville, Tennessee, USA, 08/2025.

2. "TPClust: Temporal Profile-Guided Disease Subtyping Using High-Dimensional Omics Data". *Alzheimer's Association International Conference (AAIC) 2025*, Toronto, Ontario, Canada, 07/2025.

3. "Alzheimer's Disease Subtyping Using Longitudinal Clinical Data and Brain Transcriptome". *14th Annual Taub Institute Research Retreat*, New York, NY, USA, 09/2024.

4. "A Semi-Parametric Approach for Longitudinal Outcome-Guided Disease Subtyping Using High-Dimensional Omics Data". *2024 Columbia University Postdoctoral Research Symposium*, New York, NY, USA, 05/2024.

TALKS & PRESENTATIONS

Invited Talks:

- 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 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

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

- 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
Held weekly labs and office hours, created and marked assignments and exams, delivered lectures in place of the instructor during their absence

- 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