

# Yiming (Emmett) Peng

 [emmettpeng.github.io](https://github.com/emmettpeng) |  [emmett.peng@mail.utoronto.ca](mailto:emmett.peng@mail.utoronto.ca) |  [Emmett Peng](#) | +1 437-669-4396

## EDUCATION

- **University of Toronto, Dalla Lana School of Public Health** Sept 2024 – Aug 2026 (Expected)  
*Master of Science in Biostatistics– Data Science and Artificial Intelligence Option; GPA: 4.00/4.00*  
Thesis: Bayesian Latent Class Models for Estimating the Causal Effects on Multiple Cognitive Trajectories in Aging Populations  
◦ Supervised by Dr. Kuan Liu and Dr. Zihang Lu
- **Beijing Institute of Technology, Zhuhai** Sept 2020 – Jun 2024  
*Bachelor of Science in Applied Statistics; GPA: 3.96/4.00 (WES)*  
Thesis: Deep Learning and Stochastic Modeling of Human Activity Patterns through Smartwatch Sensors for Health Interventions (*Honored as Outstanding Thesis*)  
◦ Supervised by Dr. Han C.W. Hsiao

## PUBLICATIONS AND PRESENTATIONS

C=CONFERENCE, S=IN SUBMISSION, W=WORKING PAPERS

- [C] **Yiming (Emmett) Peng**, Victoria Truong, Aoqi Xie, Yu Shi, Pingzhao Hu (2025) **Enhancing Breast Cancer Treatment Response Prediction with Single-Cell RNA Sequencing and Large Language Models**. 2025 *Statistical Society of Canada (SSC) Annual Meeting*, p58. May 25 - 28, Saskatoon, Saskatchewan, Canada.
- [C] Jiahui Zhang, Yu Shi, Aoqi Xie, **Yiming (Emmett) Peng**, Pingzhao Hu (2025) **Uncertainty-Calibrated Interpretable Tabular Transformer Model for Atrial Fibrillation Prediction with Competing Risk**. 2025 *Statistical Society of Canada (SSC) Annual Meeting*, p42. May 25 - 28, Saskatoon, Saskatchewan, Canada.
- [C] **Yiming (Emmett) Peng**, Victoria Truong, Yu Shi, Pingzhao Hu (2025) **Enhancing Breast Cancer Treatment Response Prediction with Single-Cell RNA Sequencing and Large Language Models**. 2025 SORA-TABA Annual Workshop & DLSPH Biostatistics Research Day, May 2, Toronto, Ontario, Canada.
- [S] **Yiming (Emmett) Peng\***, Victoria Truong\*, Aoqi Xie, Yu Shi, Pingzhao Hu (2025) **PRECISE: A Framework for the Prediction of REsponse using Cell-type Inference and Single-cell Embedding in Breast Cancer Patients Receiving Anti-PD-1 Treatment**.
- [W] **Yiming (Emmett) Peng**, Zihang Lu, Kuan Liu (2025) **Bayesian Latent Class Causal Inference with Nonparametric Covariate Modeling for Baseline Cognitive Outcomes in Aging Populations**.
- [W] Daniela Denier, **Yiming (Emmett) Peng**, Kuan Liu, Amanda Ricciuto (2025) **Longitudinal Analysis of Repeated Serum Markers to Identify Clinically Relevant Patient Subgroups in Pediatric Primary Sclerosing Cholangitis: Data from the Pediatric PSC Consortium**.
- [W] Yuxin Jing, **Yiming (Emmett) Peng**, Ashley Danguedan, Andrea Knight (2025) **Examining the Relationship Between Socioeconomic Factors and Mental Health in Childhood-Onset Systemic Lupus Erythematosus Using a Multilevel Analysis**.
- [W] Jiahui Zhang, Yu Shi, Aoqi Xie, **Yiming (Emmett) Peng**, Pingzhao Hu (2025) **Predicting New-Onset Atrial Fibrillation with Competing Risks Using an Uncertainty-Calibrated Interpretable Tabular Transformer: Evidence from the Cardiovascular Imaging Registry of Calgary (CIROC)**.

## PROFESSIONAL EXPERIENCE

- **Knight Lab, The Hospital for Sick Children** Jun 2025 - Present  
*Clinical Research Project Assistant, supervised by Dr. Andrea Knight* Toronto, Canada
  - Apply causal inference methods in R to investigate socioeconomic and mental health outcomes in childhood-onset systemic lupus erythematosus (cSLE).
  - Collaborate on designing the causal framework and lead statistical analyses, including data pre-processing, data imputation, and modeling.
  - Provide statistical consulting for interdisciplinary teams and contribute to manuscript drafting for publication.
- **The Hospital for Sick Children** Jun 2025 - Present  
*Research Volunteer, supervised by Dr. Amanda Ricciuto and Dr. Anne Griffiths* Toronto, Canada
  - Apply latent class mixed models in R to analyze longitudinal serum biomarker trajectories in primary sclerosing cholangitis (PSC) patients.
  - Implement survival analysis incorporating latent class membership, showing its statistical and clinical significance in predicting transplant outcomes alongside other covariates.
  - Collaborate with clinicians and biostatisticians to interpret findings and contribute to manuscript preparation.

- **Hu Lab, Dalla Lana School of Public Health, University of Toronto** Oct 2024 - Jun 2025  
Toronto, Canada  
Data Science Research Student, supervised by Dr. Pingzhao Hu
  - Developed an integrated multi-stage pipeline in Python combining single-cell RNA-seq data with embeddings from large language models (LLMs) and foundation models to identify cell type-specific biomarkers for breast cancer treatment response.
  - Applied supervised classification using cell-type-specific biomarkers and cluster-specific bulk gene-expression profiles to predict treatment response, achieving improved predictive performance over baseline methods.
  - Drafted manuscripts and presentations for conferences and peer-reviewed journals.
- **CITIC Securities Company Limited** Aug 2023 - Sep 2023  
Zhengzhou, China  
Summer Internship
  - Engaged in hands-on learning about investment, financing, and capital intermediary practices.
  - Assisted experienced customers in formulating shareholding reduction strategies.

## HONORS AND AWARDS

- **Winner – Case Studies in Data Analysis Competition, SSC 2025** May 2025  
Statistical Society of Canada (SSC)
  - Ranked 1st out of 16 teams in the 2025 SSC Case Study Competition (Case Study #1): Prediction of New Onset Atrial Fibrillation Using Routinely Reported 12-Lead ECG Variables and Electronic Health Data.
  - Poster Title: Uncertainty-Calibrated Interpretable Tabular Transformer Model for Atrial Fibrillation Prediction with Competing Risk.
  - Team: Jiahui Zhang, Yu Shi, Aoqi Xie, and Yiming (Emmett) Peng. Mentor: Dr. Pingzhao Hu. Award valued at \$2,500 CAD per team.
- **CSSC 2025 Travel Award** May 2025  
Canadian Statistics Student Conference (CSSC), sponsored by CANSSI
  - Awarded to select student presenters at CSSC 2025 for academic excellence and research contributions. Sponsored by CANSSI and valued at up to \$150 CAD for travel support.
- **BITZH President's Scholarship** May 2023  
Beijing Institute of Technology, Zhuhai
  - Awarded to 112 students in the Beijing Institute of Technology, Zhuhai for outstanding performance across all areas of academic and extracurricular achievement. Valued at \$2,000 CAD.
- **College Dean's Honor Scholarship** 2021 - 2024  
Beijing Institute of Technology, Zhuhai
  - 3x recipient (2021-24). Awarded to the top 10% of students in the department. Valued at \$1,000, \$2,000, and \$4,000 CAD, respectively.
- **BITZH Freshman Entrance Scholarship and Continuing Scholarship** 2020 - 2024  
Beijing Institute of Technology, Zhuhai
  - 4x recipient (2020-24). Awarded to the top 10% of students in the program based on outstanding performance in the National College Entrance Examination and continued academic excellence. Valued at \$4,800 CAD per year.
- **BITZH Outstanding Student Scholarship** 2020 - 2024  
Beijing Institute of Technology, Zhuhai
  - 8x recipient (each semester from 2020 to 2024). Awarded to the top 10% of students in each major.

## SKILLS

- **Programming Languages:** R, Python, MATLAB, C++
- **Tools & Technologies:** Git, LaTeX, Eviews, Maple, Mathematica, RMarkdown, Jupyter Notebook, SQL, Excel
- **Statistics Concepts:** Causal Inference, Latent Class Analysis, Survival Analysis, Longitudinal Data
- **Machine Learning Models:** Transformers, RNNs, Regression (GLM, Penalization), CNNs, Neural Networks
- **Languages:** Native proficiency in Chinese; professional working proficiency in English

## COMMUNITY INVOLVEMENT AND LEADERSHIP EXPERIENCE

- **Biostatistics Union of Graduate Students (BUGS)** Sep 2024 - Present  
PHSA MSc Representative and Seminar Committee member 
- **Health Data Working Group** Sep 2024 - Present  
Student Member 
- **Student Union, College of Global Talent, Beijing Institute of Technology, Zhuhai** Sep 2020 – Jun 2022  
Officer