# Zhengyang Fei

## **Education**

## University of Toronto

Toronto, Canada

M.Sc. Biostatistics, Dalla Lana School of Public Health

2025

Relevant Coursework: Categorical Data Analysis, Survival Analysis, Statistical Programming, Machine Learning, Applied Bayesian Methods, Causal Inference

### University of Toronto

Toronto, Canada

B.Sc. (Honours), Applied Statistics and Mathematics

2024

Relevant Coursework: Experimental Design, Statistical Learning, Data Analysis, Tensor-on-tensor Regression

## Technical Skills

Programming Tools and Productivity Softwares: R, Python, SAS, JAGS, GitHub, Microsoft Office Suite, Visual Studio, MiKTeX, Adobe Acrobat, Adobe Premiere Pro

# Experience

#### Institute for Better Health - Trillium Health Partners

May - Aug 2025

Data Analyst - Student Researcher (Supervisor: Dr. Simona Minotti)

Mississauga, Ontario

- Contributed to a hospital-wide equity analysis by working with high-dimensional complex inpatient experience survey data to support statistical modeling efforts.
- Researched methods for extracting themes from free-text using text mining and natural language processing.
- Used statistical software (R, SAS) for analysis, including the I-MAIHDA multilevel model for identifying disparities in patient-reported experience across intersecting social strata, informing hospital equity strategies.
- Collaborated with a team of multidisciplinary researchers to ensure data quality and consistency.
- Delivered actionable insights to hospital leadership through formal reports and presentations, influencing quality improvement priorities; contributed to a manuscript under preparation.

#### University of Toronto - MiDATA Lab

Nov 2024 - Present

Data Analyst - Researcher

Toronto, Ontario

- Student researcher at the MiDATA Lab, contributing to a range of ongoing projects including clinical research on juvenile arthritis and MR/PET radiomics features accuracy.
- Used statistical software (R, SAS) to develop models and create data visualizations for exploratory and inferential analysis across both small and large datasets.
- Interpreted analytical results, drafted summary reports, and contributed to manuscript writing for journal submission.
- Collaborated with a team of multidisciplinary researchers to ensure data quality and consistency.
- Conducted literature review for appropriate methodologies.

## University of Toronto

Sep 2022 - Present

Teaching Assistant

Toronto, Ontario

- Fulfilled core TA responsibilities across multiple undergraduate courses—including differential and integral calculus, linear algebra, probability and statistics, statistical modeling, and data analysis—by leading tutorials, grading assessments, and hosting office hours.
- Served as Head TA for the Statistical Modelling course, creating lecture videos, co-developing course materials with the coordinator, and designing marking schemes to support consistent evaluation.

## Personal Projects and Competitions

# Atrial Fibrillation Prediction Modeling Using ECG and EHR Data (Link)

2025

- Performed missing data imputation using the MICE package with predictive mean matching (PMM).
- Conducted variable selection and model evaluation to optimize predictive performance using features derived from ECG measurements and electronic health records.
- Trained and tuned survival gradient boosting and random survival forest models to predict new-onset atrial fibrillation.
- Presented project findings as a poster at the 2025 Statistical Society of Canada (SSC) Student Research Competition.

#### Multiple Imputation Deletion R Package (Link)

2024

• Developed an R package (DIMP) that addresses the issue of Multiple Imputation then Deletion by removing imputed outcomes, enabling users to perform analyses and build models using observed outcomes alongside imputed covariates.