

# Dewei Lin

☎ (+1) 984-234-8126 | ✉ dewei.lin@gwu.edu | 🏠 Homepage | 🐙 GitHub | 🔗 LinkedIn

## Education

### George Washington University

DOCTOR OF PHILOSOPHY IN STATISTICS

Washington, D.C., US

Aug. 2026 – May. 2031 (Expected)

### University of North Carolina at Chapel Hill

MASTER OF SCIENCE IN BIOSTATISTICS

Chapel Hill, US

Aug. 2024 – May. 2026

- GPA: Not assigned
- **Research Advisors:** Dr. Tanya P. Garcia and Dr. Baiming Zou
- **Master Thesis:** *An Enhanced Deep Mendelian Randomization for Adjusting Complex Confounding Structures*

### University of Toronto

HONOURS BACHELOR OF SCIENCE (WITH HIGH DISTINCTION)

Toronto, Canada

Sept. 2020 – May. 2024

- Major in Statistics (GPA: 3.7) and Economics (GPA: 3.9)
- **Research Advisors:** Dr. Sotirios Damouras and Dr. Sohee Kang

## Publications

1. Akbar, N.; **Lin, D.**; Cheung, Y.H.; *et al.*; Low body mass index is associated with higher rates of mild cognitive impairment and worse memory performance among older Canadian adults: An analysis of the Canadian Longitudinal Study on Aging (CLSA) database. *Journal of Alzheimer's Disease* 0, 0. (2026)
2. Vazquez, J.E.; **Lin, D.**; Mendizabal, A.; *et al.*; Higher Educational Attainment in Huntington Disease Families: Evidence from the Enroll-HD Study. *Orphanet Journal of Rare Diseases* 21, 142. (2026)

## Papers in Progress

1. Raman, M.; Vazquez, J.E.; **Lin, D.**; *et al.*; **Predicting Huntington Disease Integrated Staging System Stage Transitions to Inform Clinical Trial Enrichment.**
2. **Lin, D.**; Garcia, T.P.; Vazquez, J.E.; **Identifiability of Regression Models with a Right-Censored Covariate: Extending Canonical Causal Diagrams.**

## Research Experience

### Statistical Consultant

DEPARTMENT OF BIOSTATISTICS, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Chapel Hill, US

Jan. 2026 – May. 2026

- Queried and joined multi-modal EHR data across multiple MIMIC-IV database modules via SQL on Google BigQuery, extracting and integrating records for 33,000+ hospital admissions for downstream NLP/LLM analysis.
- Full fine-tuned Gemma 3 4B on clinical discharge notes for infection classification, with a custom data pipeline for long-document truncation and class-balanced training
- Benchmarked LLM performance against traditional ML baselines (logistic regression, LASSO, random forest).

### Student Research Assistant

METHODS FOR INCOMPLETE DATA (MIND) LAB, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Chapel Hill, US

Aug. 2024 – present

- Conducted literature reviews in Huntington's disease and right-censored covariate research, identified present research gaps.
- Developed a RShiny App to compare different existing survival models for Huntington's disease.
- Collaborated with team members on processing data and implemented multi-state survival model on Huntington's Disease Integrated Staging System (HD-ISS) framework, interpreted and communicated results.
- Prepared data and method sections of manuscripts for submission to a peer-reviewed journal.
- Derived Proofs for (non)identifiability of regression model with right censored covariates.

## Student Research Assistant

ZOU'S LAB, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Chapel Hill, US

Oct. 2024 – present

- Conducted literature review and identified existing gap in current Mendelian Randomization research.
- Developed methodologies for semi-parametric instrumental variable analysis under dual challenge of complex confounding structures and missing data.
- Adapted 5 Machine Learning models to optimize non-parametric estimation, proposed revised estimator and cross-fitting for efficient yet robust estimations.

## Presentations

### The Governor's Challenge

Ottawa, Canada

BANK OF CANADA

Nov. 2023

- Presented forecasts of key economics indicators and predicted future trends of inflation and GDP growth, provided advice on monetary policy.
- Answered technical questions as lead statisticians in the team.

### CMS Research Symposium

Toronto, Canada

DEPARTMENT OF COMPUTER AND MATHEMATICAL SCIENCES, UNIVERSITY OF TORONTO

Oct. 2023

- Hosted a poster presentation on my project '*Risk factors in survival of patients with malignant melanoma*'.

## Honors & Awards

- 2026 **Honorable Mention**, STAI-X Challenge — Statistical Agents Track
- 2024 **Undergraduate Researcher Scholarship**, University of Toronto
- 2023 **Finalist**, Rotman School of Management Datathon
- 2020-23 **Dean's List**, Awarded by University of Toronto for high academic standing

## Programming Languages and Softwares

**Skills** R, Python, SAS, SQL, LaTeX, Bash, Docker, GitHub, RShiny, Parallel computing

**Certificates** SAS Certified Specialist: Base and Advanced Programming Using SAS 9.4

## Leadership & Service Roles

- 2025-26 **Mentor**, UNC-CH Biostatistics Student Association
- 2024-25 **Student representative**, Computing Committee
- 2021-22 **Human Resources Chair**, Green Path Student Association
- 2021-22 **Organizer & Director of Undergraduate mentorship program**, Green Path Student Association
- 2021-24 **Mentor**, University of Toronto

## Teaching Experience

GR (Graduate level), UG (Undergraduate level)

### BIOS 664: Sample Survey Methodology (GR)

GRADER

Chapel Hill, US

2026 Spring

### BIOS 665: Analysis of Categorical Data (GR)

GRADER

Chapel Hill, US

2025 Fall

### STAC33H3: Introduction to Applied Statistics (UG)

TEACHING ASSISTANT

Toronto, Canada

2024 Winter

### STAA57H3: Introduction to Data Science (UG)

TEACHING ASSISTANT

Toronto, Canada

2023 Winter

### STAB22H3 & STAB23H3: Introduction to Statistics for the Social Sciences (UG)

TEACHING ASSISTANT

Toronto, Canada

2023 Fall, 2023 Winter

### STAB53H3: Introduction to Probability Theory (UG)

TEACHING ASSISTANT

Toronto, Canada

2023 Fall