Sherry Wu

sherryshwu.github.io sherry.wu.gr@dartmouth.edu sherryshwu sherryshwu Google Scholar

Summary

I am a fourth-year Ph.D. candidate in Quantitative Biomedical Sciences at Dartmouth College. Trained in economics and biostatistics before my PhD, I develop statistical and machine learning methods to study complex biomedical systems through a causal and quantitative lens — spanning computational immunology and population health, from high-performance bioinformatics pipelines to novel modeling frameworks for longitudinal biological data.

Education

Dartmouth College — **Hanover, NH**
Ph.D. in Quantitative Biomedical Sciences — *Sep 2023 – June 2028 (expected)*

New York University — **New York, NY**
Master of Science in Biostatistics — *Sep 2021 – May 2023*

Chongqing University — **Chongqing, China**
Bachelor of Economics — *Sep 2016 – Jun 2020*

University of Queensland — **Brisbane, Australia**
International Student Exchange Program — *Feb 2018 – Jun 2018*

Publications

[1] Tae-Wook Chun, Xiaoran Shang, Andrey Galkin, Xiaozhen Zhang, **Sherry Wu**, et al. "Autologous Antibodies Targeting the Silent Face of HIV Envelope Exert Potent Virologic Control". In: *Nature Immunology* (2026).
 # Computational Immunology # Phylogenetics # BCR Repertoire Analysis

[2] Jessie J. Fielding, **Sherry Wu**, Hunter J. Melton, Nic Fisk, Louis Du Plessis, and Kenneth B. Hoehn. "TyCHE Enables Time-Resolved Lineage Tracing of Heterogeneously-Evolving Populations". In: *Molecular Biology and Evolution* (2026). In Revision. DOI: 10.1101/2025.10.21.683591.
 # Computational Biology # Phylogenetics # Longitudinal Modeling

[3] Wenbo Wu, **Sherry Wu**, Sim Berlene Mariano, Richard E. Burney, and Jonathan P. Kuriakose. "Racial Disparities in Extended Venous Thromboembolism Prophylaxis after Hysterectomy". In: *PLOS ONE* 20.1 (2025), e0318433. DOI: 10.1371/journal.pone.0318433.
 # Population Health # Generalized Linear Models

[4] Ray Congrui Yu, **Sherry Wu**, and Jiang Gui. Residual Kolmogorov-Arnold Network for Enhanced Deep Learning. *Preprint*. 2024. arXiv: 2410.05500 [cs.CV]. https://arxiv.org/abs/2410.05500.
 # Deep Learning # Neural Networks # Function Approximation

Conference Abstracts and Presentations

[1] Chenjuan Ma and **Sherry Wu**. "Medicare Advantage and Home Health Outcomes in Persons Living with Dementia". In: *Academy Health Annual Research Meeting*. Poster presentation. Academy Health. Minneapolis, MN, June 2025.

[2] **Sherry Wu**, Ray Yu, Siting Li, Elijah Stommel, Mark Garret, et al. "ALS Genetic Risk Prediction Using Kolmogorov-Arnold Networks". In: *Annual NEALS Meeting*. Poster presentation. NEALS. Clearwater, FL, Oct. 2025.

Awards

- Dartmouth Geisel School of Medicine Biomedical Data Science Travel Award ($1200) *2025*

Research Experience

Dartmouth College Geisel School of Medicine **Hanover, NH**

Doctoral Researcher *2023 – Present*

[1] Long-Term B-Cell Molecular Clock Dynamics *Nov 2025 – Present*

PhD thesis · Prof. Kenneth Hoehn

- Characterized time-dependent rate decay in B cell molecular clocks by engineering a computational framework to model affinity maturation in longitudinal BCR-seq data over extended timescales
- Developing a mechanistic model to delineate the impact of purifying selection on long-term mutation rates

[2] TyCHE: Time-Resolved Lineage Tracing for Heterogeneous Evolution *Jun 2024 – Oct 2025*

PhD thesis · Prof. Kenneth Hoehn · Molecular Evolution & Biology (in revision)

- Developed a computational pipeline for evaluating the performance of TyCHE that infers time-resolved phylogenies of distinctively evolving cell populations
- Benchmarked tree topology and ancestral cell type inference against simulated ground truth

[3] BCR Repertoire Analysis of Antibodies Targeting the HIV Envelope Silent Face *Sep 2025 – Mar 2026*

Collaboration · Prof. Kenneth Hoehn · Nature Immunology (in revision)

- Built a reproducible four-stage SLURM pipeline integrating 100k+ bulk BCR, 10x single-cell, and sorted B-cell sequences from longitudinal HIV data into a unified AIRR-format dataset using IgBLAST and Immcantation
- Performed clonal clustering via optimal threshold detection and hierarchical clustering with light-chain refinement, and inferred B-cell lineage phylogenies for key antibody lineages using IgPhyML and Dowser

[4] RKAN: Residual Kolmogorov-Arnold Network *Sep 2024 – Oct 2025*

Collaboration · Prof. Jiang Gui · Preprint

- Designed and implemented the RKAN module, a compact plug-in that integrates learnable polynomial feature transformations into existing CNN architectures
- Evaluated performance on multiple benchmarks demonstrating consistent gains over baseline CNNs

[5] ALS Genetic Risk Prediction using Kolmogorov-Arnold Network *Jun 2024 – Aug 2024*

Collaboration · Prof. Jiang Gui · NEALS poster presentation

- Evaluated Kolmogorov-Arnold Networks (KAN) against MLPs and Gradient Boosting to capture non-linear genetic risk using dbGaP GWAS data from $22,418$ ALS individuals and 10 genetic principal components
- Engineered biological simulation frameworks ($300,000$ samples, 372 SNPs) to benchmark architectural performance across epistatic, sequential, and inhibitory gene-gene interaction patterns in genetic risk modeling

[6] Decomposition Analysis of Racial Disparities in Regional Mortality *Mar 2024 – May 2024*

Rotation · Prof. James OMalley

- Decomposed regional variance in mortality rates into contributions from racial composition heterogeneity and race-specific regional effects using fixed-effects and mixed-effect linear models

[7] Single-Cell BCR and Transcriptomic Profiling of Adaptive Immunity *Dec 2023 – Feb 2024*

Rotation · Prof. Jiwon Lee

- Built a bioinformatics workflow integrating single-cell V(D)J and gene expression data using Immcantation and Seurat to characterize B cell repertoires

[8] Automated BEAST 2 Pipeline for BCR Lineage Phylogenetics *Sep 2023 – Nov 2023*

Rotation · Prof. Kenneth Hoehn

- Developed an XML writer in R to automate BEAST 2 configuration file generation for seamless phylogenetic analysis of BCR-seq via Dowser

New York University **New York, NY**
Graduate Research Assistant *2022 – 2023*

[1] Neural Network for Healthcare Provider Profiling *Sep 2022 – May 2023*
Master's thesis · Prof. Wenbo Wu, Grossman School of Medicine
- Developed a feed-forward neural network model that captures non-linear effects in complex longitudinal data using fixed effect models to evaluate the performance of healthcare providers implementing the back-propagation algorithm and gradient-descent technique

[2] Racial Disparities in VTE Prophylaxis after Hysterectomy *May 2022 – Aug 2023*
Collab · Prof. Wenbo Wu, Grossman School of Medicine
- Led the statistical analysis for a PLOS ONE study on healthcare equity, utilizing mixed-effects logistic regression and incorporating Pannucci VTE risk scores to evaluate racial disparities in post-hysterectomy prophylaxis
- Co-authored the methodology section and generated all publication-ready statistical tables and figures

[3] Hospital Quality and Postsurgical Prognosis of Pancreatic Cancer Patients *May 2022 – Oct 2022*
Collab · Prof. Wenbo Wu, Grossman School of Medicine
- Estimated hospital-specific random effects for provider profiling leveraging generalized linear models
- Fit time-varying coefficient Cox models for vital status using provider effects and plotted the trajectory of hazard ratios for various hospital groups over the post-surgical time to death

[4] Medicare Advantage and Home Health Outcomes in Persons Living with Dementia *May 2022 – May 2023*
Collab · Prof. Chenjuan Ma, Rory Meyer School of Nursing
- Engineered a multi-level dataset in Stata by linking national patient-level clinical registries (OASIS, MedPAR, MBSF) with agency metrics and county-level characteristics (AHRF, OES) for 743k+ ADRD patients via ZIP-to-FIPS crosswalk mapping
- Fit multilevel logistic regression models to evaluate 30-day HHC hospitalization risks by Medicare Advantage status and race/ethnicity across 6,410 agencies, adjusting for patient, agency, and neighborhood-level covariates
- Drafted methods and results for the manuscript detailing data linkage protocols and hierarchical modeling

[5] Oral Health Pilot Study of Filipino Immigrants in the United States *May 2022 – May 2023*
Collab · Prof. Emerson Ea, Rory Meyer School of Nursing
- Processed and analyzed a 100+ item Likert-scale survey on oral health and dental care utilization among elderly Filipino Americans using descriptive statistics and regression models, and drafted internal reports
- Assessed psychometric measurement scale reliability and executed regression diagnostics

[6] Promoting Equity via Change In Practice for Respiratory Failure *Sep 2022 – May 2023*
Collab · Prof. Mari Armstrong-Hough, Global School of Public Health
- Contributed to random intercept modeling via generalized estimating equations for statistical analysis plan

[7] Racial and Ethnic Disparities in the Functional Recovery Post Respiratory Failure *May 2022 – May 2023*
Collab · Prof. Mari Armstrong-Hough, Global School of Public Health
- Performed data cleaning and exploratory analyses on $\sim$ 10k longitudinal data with R
- Fit ordinal logistic regression models to compare the Likert scale functional outcomes across racial groups
- Generated code snippets for descriptive statistics, independence tests and odds ratios to optimize the workflow

[8] Identifying Adverse TB Treatment Outcomes by Glycemic Trajectory in HIV-TB Setting *May 2022 – Aug 2022*
Collab · Prof. Mari Armstrong-Hough, Global School of Public Health
- Prepared statistical analysis plan for a prospective cohort study to evaluate the effect of glycemic exposure on adverse TB treatment outcomes in a high HIV-TB-burden setting

Teaching experience

Foundations of Data Science (QBS 103), Dartmouth College *Summer 2025*
Teaching Assistant · Profs. Noelle Kosarek and Monica Espinoza

Intermediate Programming in Python (QBS 101.1), Dartmouth College *Winter 2025*
Teaching Assistant · Prof. Christian Darabos

Data Visualization (QBS 180), Dartmouth College *Winter 2025*
Teaching Assistant · Prof. Ramesh Yapalparvi

Foundations of Biostatistics I (QBS 120), Dartmouth College *Fall 2024*
Teaching Assistant · Prof. Rob Frost

Professional Experience

IQVIA Holdings **Shanghai, China**
Intern, Management Consulting *Mar 2021 – May 2021*

- Synthesized 18 expert interviews and 1,000+ patient cards to characterize HCP risk stratification, lipid management mindsets, and Amgen Repatha prescribing patterns across patient segments
- Drafted findings for cross-region, and cross-hospital analyses on PCSK9i awareness, prescribing behavior, and volume-based procurement (VBP) policy impact in the Repatha Awareness Tracking report slide deck

Jones Lang Lasalle (JLL) **Shanghai, China**
Analyst Intern, Valuation & Advisory Services *Oct 2020 – Feb 2021*

- Assessed the market value of 30+ institutional real estate properties including commercial and industrial buildings by adopting a combination of DCF analysis and Income Capitalization Method with Argus Enterprise
- Conducted investment yield analysis by researching major transactions for a brief industry market overview
- Examined project characteristics and revised development schemes to facilitate investment decision-making

Nielsen Holdings **Beijing, China**
CPG Analyst Intern, Customer Insight *Jul 2020 – Sep 2020*

- Conducted desk research for Baby, Kids & Maternity market overview to uncover the growth of offline channels
- Discovered characteristics of online shopping platforms by mapping user profiles to identify key sales triggers
- Evaluated the performance of e-tailer's brand image & category penetration with funnel analysis by segmentation
- Forecast innovation trends based on the hotspots analysis of questionnaire with Excel pivot tables and vlookup

HuaTai Securities Co., Ltd. **Nanjing, China**
Remote Data Analyst Intern *Sep 2018 – Oct 2018*

- Performed regression analyses of historical performance (net worth, change rate, etc.) of the yields of stocks
- Built a multi-factor forecasting model combining industrial characteristics with Python and presented the proposal implementing quantitative methods for basic stock selection in the food and beverage industry

Skills and interests

Programming & Database: R, Python, SQL, Stata, SAS, SLURM, Git
Statistical Methods: Regularization (Lasso, Ridge Regression), Mixed-Effects Models, Survival Analysis, SVM, Random Forest, Naïve Bayes, KNN
Bioinformatics: IgBLAST, Immcantation, Dowser, IgPhyML, Seurat, BEAST 2
Data Science: PySpark, Pandas, Numpy, Scikit-learn, Tidyverse, Stan
Visualization: Matplotlib, Seaborn, Tableau, ggplot
Language: Proficient in English, fluent in Mandarin, basic in Cantonese
Interests: Jogging, hiking, musicals & operas