

Catherine Wang

 catherinewang1 |  website |  catheri2@andrew.cmu.edu

EDUCATION

2020 - present PhD candidate at **Carnegie Mellon University** (GPA: 4.0/4.0)
Department of Statistics and Data Science
Advised by Kathryn Roeder and Larry Wasserman

2016 - 2020 Bachelor's Degree at **University of California, Berkeley** (GPA: 3.96/4.0)
Computer Science, BA
Statistics, BA

RESEARCH INTERESTS

Statistical topics causal inference, empirical Bayes/shrinkage, semi-parametric theory and efficiency, survival analysis, network analysis, mixed effects modeling, statistical machine learning

Application areas genetics (e.g. single-cell CRISPR screens), psychiatry (e.g. structural magnetic resonance imaging [SMRI], diffusion-weighted imaging [DWI], functional MRI [fMRI]), nephrology, electronic health records

SKILLS

Data Analysis & Programming R, python, command line, git, SQL
experience: Java, C, RISC-V, MATLAB

Computer Science algorithms (time and space complexity analysis), complexity theory

Mathematics linear algebra, calculus, optimization, analysis

WORK EXPERIENCE

Teaching Assistant Aug 2020 - June 2021
TA at Carnegie Mellon University for Statistics and Data Science classes 36-401: Modern Regression and 36-402: Advanced Data Analysis. Received the Outstanding TA Award 2020-2021.

Evidation: Data Science Intern Summer 2019
Interned at Evidation and worked with electronic health records, especially in conjunction with data derived from activity trackers such as Fitbit devices, Garmin devices, and the Apple Watch (e.g. heart-rate, steps, exercise, and sleep).

Tutor and Grader for Stats 134 at UC Berkeley 2018 - 2019

PROJECTS

Proximal Causal Inference for scCRISPR screens

We propose a methodology for analyzing causal effects using proximal inference. Single-cell CRISPR screen experiments allow for the tests and estimation of causal effects by perturbing a genomic region of interest and measuring changes in gene expression; however, insufficient adjustment for confounding may result in bias, producing large false positive rates. We explore the proximal inference framework in scCRISPR screens by constructing valid and strong treatment-inducing confounding proxies and outcome-inducing confounding proxies. We consider various estimation procedures such as two-stage least squares approaches, generalized method of moments approaches, plug-in and one-step corrected method, and subspace estimation methods.

Multi-level Treatments with multivariate outcomes with applications to scCRISPR screens

In the setting where we are interested in the effect of a treatment with multiple levels on many outcomes, we present two analysis procedures with different goals: (1) estimation and (2) hypothesis testing. In an effort to take advantage of structure in the matrix of effects, we consider an empirical Bayes approach that will additionally decrease overall estimation error and provide robust confidence intervals.

Researchers often lack the resources to test every hypothesis of interest directly, but often possess auxiliary data from which we can compute an estimate. We introduce a novel approach for selecting which hypotheses to query by leveraging estimates to compute proxy statistics. Through simulations and real data analysis of scCRISPR screen experiments, we empirically demonstrate that our proxy framework has both high power and low resource usage.

PUBLICATIONS

Wang, Catherine, Kathryn Roeder, and Larry Wasserman (2026+). “Multi-level treatments with multivariate outcomes: with applications to single-cell CRISPR screens”. (Working paper).

Wang, Catherine, Larry Wasserman, and Kathryn Roeder (2026+). “Proximal Causal Inference for single-cell CRISPR screens”. (Working paper).

Catherine Wang, Armin Ahmadi, Rasha Hussein, Hanjie Zhang, Len Usvyat, Yuedong Wang, and Peter Kotanko (2026). *Longitudinal albumin and blood pressure trajectories identify mortality risk phenotypes in hemodialysis*. (Submitted to BMC Nephrology).

Xu, Ziyu, **Catherine Wang**, Larry Wasserman, Kathryn Roeder, and Aaditya Ramdas (2025). *Active multiple testing with proxy p-values and e-values*. (Major Revisions at JASA). arXiv: [2502.05715](https://arxiv.org/abs/2502.05715) [stat.ME]. URL: <https://arxiv.org/abs/2502.05715>.

Glazer, Amanda, Hubert Luo, Shivin Devgon, **Catherine Wang**, Xintong Yao, Steven Siwei Ye, Frances McQuarrie, Zelin Li, Adalie Palma, Qinqin Wan, Warren Gu, Avi Sen, Zihui Wang, Grace O’Connell, and Philip Stark (Jan. 2023). “Look who’s talking: gender differences in academic job talks”. In: *ScienceOpen Research*. DOI: [10.14293/S2199-1006.1.SOR.2023.0003.v1](https://doi.org/10.14293/S2199-1006.1.SOR.2023.0003.v1).

Kaufman, Harvey W., **Catherine Wang**, Yuedong Wang, Hao Han, Sheetal Chaudhuri, Len Usvyat, Carly Hahn Contino, Robert Kossman, and Michael A. Kraus (2023). “Machine Learning Case Study: Patterns of Kidney Function Decline and Their Association With Clinical Outcomes Within 90 Days After the Initiation of Renal Dialysis”. In: *Advances in Kidney Disease and Health* 30.1. AI in Kidney Disease: What Will the Future Bring? Part II, pp. 33–39. ISSN: 2949-8139. DOI: <https://doi.org/10.1053/j.akdh.2022.11.006>. URL: <https://www.sciencedirect.com/science/article/pii/S2949813922000076>.

Wang, Catherine, Rebecca Hayes, Kathryn Roeder, and Maria Jalbrzikowski (2023). “Neurobiological Clusters Are Associated With Trajectories of Overall Psychopathology in Youth.” In: *Biological psychiatry. Cognitive neuroscience and neuroimaging* 8.8, pp. 852–863. ISSN: 2451-9030 (Electronic). DOI: [10.1016/j.bpsc.2023.04.007](https://doi.org/10.1016/j.bpsc.2023.04.007).

Thadhani, Ravi, Joanna Willetts, **Catherine Wang**, John Larkin, Hanjie Zhang, Lemuel Rivera Fuentes, Len Usvyat, Kathleen Belmonte, Yuedong Wang, Robert Kossman, Jeffrey Hymes, Peter Kotanko, and Franklin Maddux (2021). “Transmission of SARS-CoV-2 considering shared chairs in outpatient dialysis: a real-world case-control study.” In: *BMC nephrology* 22.1, p. 313. ISSN: 1471-2369 (Electronic). DOI: [10.1186/s12882-021-02518-4](https://doi.org/10.1186/s12882-021-02518-4).

PEER-REVIEWED ABSTRACTS

[Black, Non-Hispanic Individuals Are More Likely to Have Slower eGFR Decline Before Dialysis](#). Michael A. Kraus, Yuedong Wang, **Catherine Y. Wang**, Len A. Usvyat Carly Hahn Contino, Peter Kotanko, Robert J. Kossman, and Harvey Kaufman. *Journal of the American Society of Nephrology* 33: 2022.

[eGFR Trajectory Pre-Dialysis Initiation Predicts 90-Day Hospitalization and Mortality Risk](#). Michael A. Kraus, Yuedong Wang, **Catherine Y. Wang**, Carly Hahn Contino, Len A. Usvyat, Peter Kotanko, Robert J. Kossman, and Harvey W. Kaufman. *Journal of the American Society of Nephrology* 33: 2022