

CHARLES J. WOLOCK

cwoLOCK@upenn.edu ♦ cwoLOCK.github.io

Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania
417 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104

EDUCATION

University of Washington, Seattle, WA

September 2018 - August 2023

Ph.D., Biostatistics

Advisors: Marco Carone, Ph.D. and Noah Simon, Ph.D.

Harvard University

September 2011 - May 2015

B.A., Organismic and Evolutionary Biology

Language citation, Spanish

Summa cum laude, Highest Honors

PROFESSIONAL POSITIONS

Postdoctoral Researcher

Department of Biostatistics, Epidemiology and Informatics

University of Pennsylvania

Supervisor: Yong Chen, Ph.D.

September 2023 - present

Research Staff Associate

Institute for Genomic Medicine

Columbia University

Supervisors: Andrew S. Allen, Ph.D. and David B. Goldstein, Ph.D.

November 2016 - July 2018

TEACHING EXPERIENCE

Instructor of Record

University of Washington

BIOST311: Regression Methods in the Health Sciences

March 2022 - June 2022

Teaching Assistant

University of Washington

BIOST 511: Medical Biometry I (Lead TA)

September 2020 - December 2020

BIOST 310: Biostatistics for the Health Sciences (Lead TA)

September 2021 - December 2021

Research Mentor

Fred Hutchinson Cancer Center

Pathways Undergraduate Research Program

June 2021 - August 2021

University of Washington

Directed Reading Program: Introduction to Prediction

January 2023 - March 2023

Guest Lecturer

University of Washington

BIOST 561: Computational Skills for Biostatistics

May 2023

AWARDS, HONORS, FELLOWSHIPS

University of Washington

NSF Graduate Research Fellowship	September 2020 - September 2023
Thomas R. Fleming Excellence in Biostatistics Award	June 2023
UW Biostatistics highest honor for Ph.D. students	
Donovan J. Thompson Award	October 2020
Best combined performance on the Ph.D. Applied and Theory qualifying exams	
NIH T32 Statistical Genetics Training Grant	September 2018 - September 2020

Harvard University

Phi Beta Kappa	May 2015
Herchel Smith Research Fellowship	June 2014 - August 2014
Microbial Sciences Initiative Research Fellowship	June 2014 - August 2014
John Harvard Scholar	May 2013
National Merit Scholarship	September 2011

PROFESSIONAL SERVICE

Manuscript Reviewer

Bayesian Analysis
Biometrics
Biostatistics
Journal of Machine Learning Research
Scientific Reports
Statistics and Computing

Conference Session Chair

Next Generation Methods for Evidence Synthesis and Reproducibility (*JSM, Portland, August 2024*)
Heterogeneous Treatment Effects (*ACIC, Seattle, May 2024*)

UNIVERSITY SERVICE

University of Washington, Department of Biostatistics

Faculty Search Committee	September 2022 - March 2023
Student Seminar Coordinator	September 2020 - June 2022
Peer mentor	June 2019 - June 2022
Admissions Committee	September 2021 - March 2022
Education Policy and Teaching Evaluation Committee	September 2020 - September 2021
Equity, Diversity, and Inclusion Committee	September 2019 - June 2021
Student-Faculty-Staff Relations Committee	September 2019 - September 2020

REFEREED JOURNAL PUBLICATIONS

(The * symbol denotes equal contribution.)

1. Sperling EA, **Wolock CJ**, Morgan AS, Gill BC, Kunzmann M, Halverson GP, Macdonald FA, Knoll AH, Johnston DT. (2015). Statistical analysis of iron geochemical data suggests limited late Proterozoic oxygenation. *Nature* 523: 451–454.

2. Raghavan NS, Brickman AM, Andrew H, Manly JJ, Schupf N, Lantigua R, The Alzheimer's Disease Sequencing Project, **Wollock CJ**, Kamalakaran S, Petrovski S, Tosto G, Vardarajan BN, Goldstein DB, Mayeux R. (2018). Whole exome sequencing in 20,197 individuals identifies ultra-rare SORL1 loss-of-function variants in late-onset Alzheimer's disease. *Annals of Clinical and Translational Neurology* 5(7): 832–842.
3. Bittleston LS, **Wollock CJ**, Bakhtiar EY, Chan XY, Chan KG, Pierce NE, Pringle A. (2018). Convergence between the microcosms of Southeast Asian and North American pitcher plants. *eLife* 7.
4. Hayeck TJ, Stong N, **Wollock CJ**, Copeland B, Kamalakaran S, Goldstein DB, Allen AS. (2019). Improved Pathogenic Variant Localization using a Hierarchical Model of Sub-regional Intolerance. *American Journal of Human Genetics* 104(2): 299–309.
5. **Wollock CJ**, Stong N, Ma F, Nagasaki T, Lee W, Tsang SH, Kamalakaran S, Goldstein DB, Allikmets R. (2019). A case-control collapsing analysis identifies retinal dystrophy genes associated with ophthalmic disease in patients with no pathogenic *ABCA4* mutations. *Genetics in Medicine* 21: 2336–2344.
6. Gelfman S, Dugger SA, Moreno CAM, Ren Z, **Wollock CJ**, Shneider NA, Phatnani H, Cirulli ET, Lasseigne BN, Harris T, Maniatis T, Rouleau GA, Brown RH, Gitler AD, Myers RM, Petrovski S, Allen AS, Harms MB, Goldstein DB. (2019). A new approach for rare variation collapsing on functional protein domains implicates specific genic regions in ALS. *Genome Research* 29(5): 809–818.
7. Cameron-Christie S, **Wollock CJ**, Groopman E, Petrovski S, Kamalakaran S, Povysil G, Zhang M, Fleckner J, March RE, Gelfman S, Marasa M, Li Y, Sanna-Cherchi S, Kiryluk K, Allen AS, Fellström B, Haeffliger C, Platt A, Goldstein DB, Gharavi A. (2019). Exome-based rare-variant analyses in chronic kidney disease. *Journal of the American Society of Nephrology* 30(6): 1109–1122.
8. Ma CJ, **Wollock CJ**, Stong N, Nagasaki T, Lee W, Goldstein DB, Allikmets R. (2019). Case-control collapsing analysis identifies genes mimicking Stargardt/*ABCA4* disease. *Investigative Ophthalmology & Visual Science* 60(9): 2935–2935.
9. Eade K, Gantner ML, Hostyk JA, Nagasaki T, Giles S, Harkins-Perry S, Fallon R, Baldini M, Schepke L, Dorrell MI, Cai C, Baugh EH, **Wollock CJ**, Wallace M, Berlow RB, Goldstein DB, Metallo CM, Friedlander M, Allikmets R. (2021). Serine biosynthesis defect due to haploinsufficiency of phosphoglycerate dehydrogenase (PHDGH) causes retinal disease. *Nature Metabolism* 3(3): 366–377.
10. Bansal A, Heagerty PJ, Inoue LYT, Veenstra DL, **Wollock CJ**, Basu A. (2021). A Value of Information Framework for Personalizing the Timing of Surveillance Testing. *Medical Decision Making* 42(4): 474–486.
11. Wasser SK, **Wollock CJ**, Kuhner MK, Brown JE, Morris C, Horowitz R, Wong A, Fernandez CJ, Otiende MY, Hoareau Y, Kaliszewska ZA, Jeon E, Han KL, Weir BS. (2022). Elephant genotypes reveal the size and connectivity of transnational ivory traffickers. *Nature Human Behaviour* 6(3): 371–382.
12. Heil J, **Wollock CJ**, Pierce NE, Pringle A, Bittleston LS. (2022). Pitcher plant-associated microbial communities differ primarily by host species across a longitudinal gradient. *Environmental Microbiology* 24(8): 3500–3516.
13. Alfieri J, Coble MD, Conroy C, Dahl A, Hares D, Weir BS, **Wollock CJ**, Zhao E, Kingston H, Zolandz T. (2022). A new implementation of a semi-continuous method for DNA mixture interpretation. *Forensic Science International: Reports* 6.
14. Chen YT, Williamson BD, Okonek T, **Wollock CJ**, Spieker AJ, Hee Wai TY, Hughes JP, Emerson SS, Willis AD. (2022). *rigr*: Regression, Inference, and General Data Analysis Tools in R. *The Journal of*

Open Source Software 7(80): 4847.

15. Kohn M, **Wolock CJ**, Poulson I, Fernando N. (2023). A meta-analysis of outcomes of patients with chronic hepatitis C vs. patients without chronic hepatitis C undergoing total hip or total knee arthroplasty. *EFORT Open Reviews* 8(4): 180–188.
16. Bittleston LS*, **Wolock CJ***, Maeda J, Infante V, Ané JM, Pierce NE, Pringle, A. (2023). Carnivorous *Nepenthes* pitchers with less acidic fluid house nitrogen-fixing bacteria. *Applied and Environmental Microbiology* e00812-23.
17. **Wolock CJ**, Gilbert PB, Simon N, Carone M. (2024). A framework for leveraging machine learning tools to estimate personalized survival curves. *Journal of Computational and Graphical Statistics* 33(3): 1098–1108.
18. Kennedy CJ, Kearns JC, Geraci JC, Gildea SM, Hwang IH, King AJ, Liu H, Luedtke A, Marx BP, Panini S, Petukhova MV, Sampson NA, Smoller JW, **Wolock CJ**, Zainal HN, Stein MB, Ursano RJ, Kessler RC. (2024+). Predicting post-transition suicides among U.S. Army soldiers transitioning out of service: Results from the Study to Assess Risk and Resilience in Servicemembers-Longitudinal Study (STARRS-LS). *JAMA Psychiatry*, in press.

PREPRINTS/UNDER REVIEW

(The * symbol denotes equal contribution.)

1. **Wolock CJ**, Gilbert PB, Simon N, Carone M. (2023+). Assessing variable importance in survival analysis using machine learning. [arXiv.2311.12726](https://arxiv.org/abs/2311.12726). Invited revision, *Biometrika*.
2. **Wolock CJ**, Williamson BD, Shortreed SM, Simon GE; Coleman KJ, Yeargans R, Ahmedani BK, Daida Y, Lynch FL, Rossom RC, Ziebell RA, Cruz M, Wellman RD, Coley RY. (2024+). Importance of variables from different time frames for predicting self-harm using health system data. [medRxiv.2024.306260](https://medrxiv.org/lookup/doi/10.1101/2024.306260). Invited revision, *Journal of Biomedical Informatics*.
3. Wu Q, Pajor NM, Lu Y, **Wolock CJ**, Tong J, Lorman V, Johnson KB, Moore JH, Forrest CB, Asch DA, Chen Y. (2024+). Post-acute healthcare demands following SARS-CoV-2 infection in children and adolescents: a hospital-specific investigation through heterogeneous transfer learning. Invited revision, *Patterns*.
4. **Wolock CJ***, Jacob S*, Weil AA, Bennett JC, Elias-Warren A, O'Hanlon J, Kenny A, Jewell NP, Rotnitzky A, Chu HY, Carone M. (2024+). Investigating symptom duration using current status data: a case study of post-acute COVID-19 syndrome. [arXiv.2407.04214](https://arxiv.org/abs/2407.04214). Under review.
5. Kenny A, **Wolock CJ**. (2024+). SimEngine: A Modular Framework for Statistical Simulations in R. [arXiv:2403.05698](https://arxiv.org/abs/2403.05698). Under review.

SOFTWARE

survML	Tools for flexible survival analysis using machine learning (cre, aut)
SimEngine	Framework for reproducible statistical simulations in R (aut)
rigr	Regression, inference, and general data analysis tools for R (aut)
vimp	Inference on algorithm-agnostic variable importance (ctb)

PRESENTATIONS

1. Estimating causal effects from electronic health records data with underreported exposure.
JSM, Portland, August 2024. (contributed)
2. Nonparametric variable importance for time-to-event outcomes with application to prediction of HIV infection.
JSM, Toronto, August 2023. (invited)
ENAR, Baltimore, March 2024. (contributed)
3. Flexible estimation of the conditional survival function via observable regression models.
ENAR, Nashville, March 2023. (contributed)
JSM, Washington, D.C., August 2022. (contributed)
WNAR, virtual, June 2022. (contributed)

PROFESSIONAL AFFILIATIONS

American Statistical Association
International Biometric Society (ENAR)

SKILLS

Programming	Python, R, Bash, SQL
Other applications	\LaTeX , Git
Languages	English (native), Spanish (proficient)