

CHARLES J. WOLOCK

cwoLOCK@upenn.edu cwoLOCK.github.io

Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania
417 Blockley Hall, 432 Guardian Drive, Philadelphia, PA 19104

EDUCATION

University of Washington, Seattle, WA 2018 – 2023
Ph.D., Biostatistics

Harvard University, Cambridge, MA 2011 – 2015
B.A., Organismic and Evolutionary Biology
Language citation, Spanish
Summa cum laude, Highest Honors

PROFESSIONAL POSITIONS

Assistant Professor
Department of Biostatistics and Computational Biology
University of Rochester *as of July 2025*

Postdoctoral Researcher
Department of Biostatistics, Epidemiology and Informatics
University of Pennsylvania 2023 – 2025

Research Staff Associate
Institute for Genomic Medicine
Columbia University 2016 – 2018

TEACHING AND MENTORING

Instructor of Record
University of Washington
BIOST 311: Regression Methods in the Health Sciences 2022

Teaching Assistant
University of Washington
BIOST 511: Medical Biometry I (Lead TA) 2020
BIOST 310: Biostatistics for the Health Sciences (Lead TA) 2021

Research Mentor
Fred Hutchinson Cancer Center
Pathways Undergraduate Research Program 2021
University of Washington
Directed Reading Program: Introduction to Prediction 2023

Guest Lecturer*University of Washington*

BIOST 561: Computational Skills for Biostatistics

2023

AWARDS, HONORS, FELLOWSHIPS

NSF Graduate Research Fellowship	2020 – 2023
Thomas R. Fleming Excellence in Biostatistics Award (University of Washington)	2023
Donovan J. Thompson Award (University of Washington)	2020
NIH T32 Statistical Genetics Training Grant (University of Washington)	2018 – 2020
Phi Beta Kappa (Harvard University)	2015
Herchel Smith Research Fellowship (Harvard University)	2014
Microbial Sciences Initiative Research Fellowship (Harvard University)	2014
John Harvard Scholar (Harvard University)	2013

PROFESSIONAL SERVICE

Manuscript Reviewer

Bayesian Analysis

Biometrics

Biostatistics

Journal of Machine Learning Research

Journal of the Royal Statistical Society: Series B

Scientific Reports

Statistics and Computing

Statistics in Biopharmaceutical Research

Conference Session ChairExploring Frontiers in Functional Data Analysis and Applications (*ENAR, New Orleans, 2025*)Next Generation Methods for Evidence Synthesis and Reproducibility (*JSM, Portland, 2024*)Heterogeneous Treatment Effects (*ACIC, Seattle, 2024*)**Paper Competition Reviewer**

ICSA Applied Statistics Symposium Student Paper Competition (2025)

UNIVERSITY SERVICE

University of Washington, Department of Biostatistics

Faculty Search Committee

2022 – 2023

Student Seminar Coordinator

2020 – 2022

Peer mentor

2019 – 2022

Admissions Committee

2021 – 2022

Education Policy and Teaching Evaluation Committee

2020 – 2021

Equity, Diversity, and Inclusion Committee

2019 – 2021

BIBLIOGRAPHY

(* equal contribution)

Refereed Journal Publications

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. doi:[10.1038/nature14589](https://doi.org/10.1038/nature14589)
2. Raghavan NS, Brickman AM, Andrew H, Manly JJ, Schupf N, Lantigua R, The Alzheimer's Disease Sequencing Project, **Wolock 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. doi:[10.1002/acn3.582](https://doi.org/10.1002/acn3.582)
3. Bittleston LS, **Wolock 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. doi:[10.7554/eLife.36741](https://doi.org/10.7554/eLife.36741)
4. Hayeck TJ, Stong N, **Wolock 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. doi:[10.1016/j.ajhg.2018.12.020](https://doi.org/10.1016/j.ajhg.2018.12.020)
5. **Wolock 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. doi:[10.1038/s41436-019-0495-0](https://doi.org/10.1038/s41436-019-0495-0)
6. Gelfman S, Dugger SA, Moreno CAM, Ren Z, **Wolock 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. doi:[10.1101/gr.243592.118](https://doi.org/10.1101/gr.243592.118)
7. Cameron-Christie S, **Wolock CJ**, Groopman E, Petrovski S, Kamalakaran S, Povysil G, Zhang M, Fleckner J, Mar. RE, Gelfman S, Marasa M, Li Y, Sanna-Cherchi S, Kiryluk K, Allen AS, Fellström B, Haefliger 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. doi:[10.1681/ASN.2018090909](https://doi.org/10.1681/ASN.2018090909)
8. Ma CJ, **Wolock 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, **Wolock 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. doi:[10.1038/s42255-021-00361-3](https://doi.org/10.1038/s42255-021-00361-3)
10. Bansal A, Heagerty PJ, Inoue LYT, Veenstra DL, **Wolock CJ**, Basu A. (2021). A Value of Information Framework for Personalizing the Timing of Surveillance Testing. *Medical Decision Making* 42(4): 474–486. doi:[10.1177/0272989x211049213](https://doi.org/10.1177/0272989x211049213)
11. Wasser SK, **Wolock 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. doi:[10.1038/s41562-021-01267-6](https://doi.org/10.1038/s41562-021-01267-6)
12. Heil J, **Wolock 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. doi:[10.1111/1462-2920.15993](https://doi.org/10.1111/1462-2920.15993)

13. Alfieri J, Coble MD, Conroy C, Dahl A, Hares D, Weir BS, **Wolock CJ**, Zhao E, Kingston H, Zoladz T. (2022). A new implementation of a semi-continuous method for DNA mixture interpretation. *Forensic Science International: Reports* 6: 100281. doi:[10.1016/j.fsir.2022.100281](https://doi.org/10.1016/j.fsir.2022.100281)
14. Chen YT, Williamson BD, Okonek T, **Wolock 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. doi:[10.21105/joss.04847](https://doi.org/10.21105/joss.04847)
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. doi:[10.1530/eor-22-0117](https://doi.org/10.1530/eor-22-0117)
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* 89(7): e00812-23. doi:[10.1128/aem.00812-23](https://doi.org/10.1128/aem.00812-23)
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. doi:[10.1080/10618600.2024.2304070](https://doi.org/10.1080/10618600.2024.2304070)
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 Suicides Among US Army Soldiers After Leaving Active Service. *JAMA Psychiatry* 81(12): 1215-1224. doi:[10.1001/jamapsychiatry.2024.2744](https://doi.org/10.1001/jamapsychiatry.2024.2744)
19. Wu Q, Pajor NM, Lu Y, **Wolock CJ**, Tong J, Lorman V, Johnson KB, Moore JH, Forrest CB, Asch DA, Chen Y. (2024). A latent transfer learning method for estimating hospital-specific post-acute healthcare demands following SARS-CoV-2 infection. *Patterns* 5(11): 101079. doi:[10.1016/j.patter.2024.101079](https://doi.org/10.1016/j.patter.2024.101079)
20. **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. *Journal of Biomedical Informatics* 160: 104750. doi:[10.1016/j.jbi.2024.104750](https://doi.org/10.1016/j.jbi.2024.104750)
21. **Wolock CJ**, Gilbert PB, Simon N, Carone M. (2025). Assessing variable importance in survival analysis using machine learning. *Biometrika* 112(2). doi:[10.1093/biomet/asae061](https://doi.org/10.1093/biomet/asae061)
22. **Wolock CJ***, Jacob S*, Bennett JC, Elias-Warren A, O'Hanlon J, Kenny A, Jewell NP, Rotnitzky A, Cole SR, Weil AA, Chu HY, Carone M. (2025). Investigating symptom duration using current status data: a case study of post-acute COVID-19 syndrome. *Epidemiology*. doi:[10.1097/EDE.0000000000001882](https://doi.org/10.1097/EDE.0000000000001882)
23. Kenny A, **Wolock CJ**. (2025+). SimEngine: A Modular Framework for Statistical Simulations in R. [arXiv:2403.05698](https://arxiv.org/abs/2403.05698). To appear in *The R Journal*.

Under Revision/Under Review/Preprint

1. Zhang D, Zhang B, Wang H, Lu Y, **Wolock CJ**, Wang L, Schuemie M, Hripcsak G, Chen Y. (2025+) Negative control-calibrated difference-in-difference analyses: addressing unmeasured confounding in real-world data with application to quantifying the impact of pandemic on racial/ethnic differences. Invited revision, *npj Digital Medicine*.

2. Carone M, **Wollock CJ**, Olivas-Martinez A, Rotnitzky A, Gilbert PB. (2025+). Immune Correlates and Vaccine Immunobridging: Statistical Innovations, Challenges and Opportunities. Invited revision, ***The Journal of Infectious Diseases***.
3. Frivold C, Cox SN, Starita L, ..., **Wollock CJ**, ..., Englund JA, Naleway AL, Chu HY. (2025+). Correlates of Risk of Respiratory Syncytial Virus Disease: A Prospective Cohort Study. Invited revision, ***Nature Communications***.
4. Wollum A, Ralph L, Grossman D, **Wollock CJ**, Moseson H. (2025+) Evaluating a double list experiment to measure the population proportion of people who have attempted a self-managed abortion in the United States: analysis of a nationally representative cross-sectional survey in 2017. Invited revision, ***AJE Advances: Research in Epidemiology***.
5. Morenz ER*, **Wollock CJ***, Carone M. (2025+) Debiased machine learning for survival functionals based on left-truncated right-censored data. [arXiv.2411.09017](https://arxiv.org/abs/2411.09017).

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

(*invited, †upcoming)

1. Nonparametric approaches to assessing variable importance using health data.
 *†*ICSA Applied Statistics Symposium, Storrs, 2025.*
 **Department of Biostatistics, Columbia University, 2025.*
 **Department of Biostatistics and Computational Biology, University of Rochester, 2024.*
 **Division of Biostatistics and Health Data Science, University of Minnesota, 2024.*
2. Transfer learning for model-free variable importance.
ENAR, New Orleans, 2025.
3. Estimating causal effects from electronic health records data with underreported exposure.
JSM, Portland, 2024.
4. Nonparametric variable importance for time-to-event outcomes with application to prediction of HIV infection.
 **JSM, Toronto, 2023.*
ENAR, Baltimore, 2024.
5. Flexible estimation of the conditional survival function via observable regression models.
ENAR, Nashville, 2023.
JSM, Washington, D.C., 2022.
WNAR, virtual, 2022.

PROFESSIONAL AFFILIATIONS

American Statistical Association
International Biometric Society (ENAR)
Institute of Mathematical Statistics

SKILLS

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