Brian David Williamson

Curriculum Vitae

last updated: 21 November, 2022

Biographical Information

Brian D. Williamson, PhD
Assistant Investigator
Kaiser Permanente Washington Health Research Institute (KPWHRI)
Kaiser Foundation Health Plan of Washington
1730 Minor Ave, Ste 1600
Seattle, WA 98101-1448
(206) 310-4888
brian.d.williamson@kp.org
bdwilliamson
https://bdwilliamson.github.io

Education

2019 **Ph.D., Biostatistics**, *University of Washington*.
Committee: Marco Carone (co-chair), Noah Simon (co-chair), Scott Emerson, Peter Gilbert
Dissertation title: *A unified approach to model-agnostic variable importance*

2017 **M.S., Biostatistics**, *University of Washington*.

2014 **B.A., Mathematics**, *Pomona College*.

Licensure

Not applicable

Professional Positions

09/2021–Present **Assistant Investigator**, *Biostatistics Division, Kaiser Permanente Washington Health Research Institute*.

01/2022–Present **Affiliate Investigator**, *Biostatistics, Bioinformatics, and Epidemiology Program, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center*.

01/2020–09/2021 **Post-doctoral research fellow**, *Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center*.

2015–2019 **Research Assistant**, *Statistical Center for HIV/AIDS Research and Prevention, Fred Hutchinson Cancer Research Center*.

06/2013–08/2013 **Integrative Cancer Biology Program Research Fellow**, *Stanford University*.

Honors, Awards, and Scholarships

Research Communication and Travel Awards

04/2019 Nonparametrics Section Travel Award, American Statistical Association (ASA)

01/2018 Biometrics Section Travel Award, ASA

06/2017 Most Outstanding Oral Paper Award, Western North American Region (WNAR) of the International Biometric Society

06/2017 Graduate School Fund for Excellence and Innovation Travel Award, University of Washington (UW)

06/2017 Graduate and Professional Student Senate Travel Grant, UW

06/2013 Best Insight Award, UCLA DataFest

Teaching and Service Awards

07/2022 ICML Best Reviewer (top 10% of reviewers for ICML 2022)

06/2019 Exceptional Service in Biostatistics Award, UW Department of Biostatistics

09/2018 Excellence in Teaching Award, UW Department of Biostatistics

Academic Honors and Awards

09/2014 Top Scholar Incoming Student Award, UW Department of Biostatistics

05/2014 Distinction in the Senior Exercise, Pomona College

Organizational Service

At KPWHRI:

05/2022–09/2022 Member: Open-rank Collaborative Biostatistician Search Committee

11/2021–Present Member: Equity, Inclusion, and Diversity Standing Committee

At Fred Hutchinson Cancer Research Center:

03/2020–09/2021 Member: Hutch United Outreach Committee

At University of Washington:

02/2019–06/2019 Member: Chair's Task Force on Professionalism; resulted in the UW Biostatistics Code of Conduct

09/2016–12/2019 Member: Equity, Diversity, and Inclusion Committee

09/2016–12/2019 Peer Mentor

09/2015–09/2018 Member: Educational Policy and Teaching Evaluation Committee

External Professional Activities

Mentorship roles

2020, 2021 Mentor: Graduate Student Mentorship Initiative, Científico Latino

06/2015–Present Alumni Mentor: SagePost 47, Pomona College

Service in professional associations

2021 Member: ASA Biometrics Section Byar Award Committee

05/2021–Present Member: Justice, Equity, Diversity, and Inclusion Committee of WNAR

Manuscript reviewer

Annals of Statistics, Biometrics, Data Mining and Knowledge Discovery, Epidemiology, International Conference on Learning Representations (ICLR), International Conference on Machine Learning (ICML), International Journal of Biostatistics, Journal of the American Statistical Association (Theory & Methods), Journal of Machine Learning Research, Journal of the Royal Statistical Society: Series B (Statistical Methodology), Machine Learning in Computational Biology (2019–2021), Neural Information Processing Systems (NeurIPS), Observational Studies, Vaccine

Memberships in professional organizations

08/2014–Present WNAR of the International Biometric Society

03/2013–Present American Statistical Association

05/2014–Present Sigma Xi: the Scientific Research Honor Society

Bibliography

(a) Refereed research articles

(the symbol ★ denotes joint first-author contribution)

(the symbol ♣ denotes contribution as primary analyst)

(the symbol ♡ denotes contribution as a senior author)

(*Published or in press*)

(17) Huang Y, Zhang Y, Seaton KE, De Rosa S, Heptinstall J, Carpp LN, Randhawa AK, McKinnon LR, McLaren P, Viegas E, Gray GE, Churchyard G, Buchbinder SP, Edupuganti S, Bekker LG, Keefer MC, Hosseinipour MC, Goepfert PA, Cohen KW, ♡**Williamson BD**, McElrath MJ, Tomaras GD, Thakar J, Kobie J, HVTN 703/SAAVI 102 Protocol Leadership Team, HVTN 086/SAAVI 103 Protocol Leadership Team, HVTN 094 Protocol Leadership Team, HVTN 097 Protocol Leadership Team, HVTN 098 Protocol Leadership Team, HVTN 100 Protocol Leadership Team, HVTN 105 Protocol Leadership Team, HVTN 111 Protocol Leadership Team, and HVTN 205 Protocol Leadership Team. Baseline host determinants of robust human HIV-1 vaccine-induced immune responses: a meta-analysis of 26 vaccine regimens. *eBioMedicine*, 84, 2022. doi: 10.1016/j.ebiom.2022.104271

(16) Hughes JP, ♣**Williamson BD**, Krakauer C, Chau G, Ortiz B, Wakefield J, Hendrix C, Amico KR, Holtz TH, Bekker LG, and Grant R. Combining information to estimate adherence in studies of pre-exposure prophylaxis for HIV prevention: application to HPTN 067. *Statistics in Medicine*, 2022. doi: 10.1002/sim.9321

(15) Han S, **Williamson BD**, and Fong Y. Improving random forest predictions in small datasets from two-phase sampling designs. *BMC Medical Informatics and Decision Making*, 21:322, 2021. doi: 10.1186/s12911-021-01688-3

(14) **Williamson BD**, Gilbert PB, Simon NR, and Carone M. A general framework for inference on algorithm-agnostic variable importance. *Journal of the American Statistical Association (Theory & Methods)*, 2021. doi: 10.1080/01621459.2021.2003200

(13) Huang Y, ♣**Williamson BD**, Moodie Z, Carpp LN, Chambonneau L, DiazGranados CA, and Gilbert PB. Analysis of neutralizing antibodies as a correlate of instantaneous risk of hospitalized dengue in placebo recipients of dengue vaccine efficacy trials. *The Journal of Infectious Diseases*, 2021. doi: 10.1093/infdis/jiab342

(12) **Williamson BD**, Hughes JP, and Willis AD. A multi-view model for relative and absolute microbial abundances. *Biometrics*, 2021. doi: 10.1111/biom.13503

(11) **Williamson BD**, Magaret CA, Gilbert PB, Nizam S, Simmons C, and Benkeser D. Super LeArner Prediction of NAb Panels (SLAPNAP): a containerized tool for predicting combination monoclonal broadly neutralizing antibody sensitivity. *Bioinformatics*, 2021. doi: 10.1093/bioinformatics/btab398

(10) **Williamson BD**, Gilbert PB, Carone M, and Simon N. Nonparametric variable importance assessment using machine learning techniques (with discussion). *Biometrics*, 2020. doi: 10.1111/biom.13392

(one of the top-cited papers published in Biometrics during 2020–2021)

(9) Duke ER, **Williamson BD**, Borate B, Golob JL, Wychera C, Stevens-Ayers T, Huang M-L, Cossrow N, Wan H, Mast CT, Marks MA, Flowers ME, Jerome KR, Corey L, Gilbert PB, Schiffer JT, and Boeckh MJ. Cytomegalovirus viral load kinetics as surrogate endpoints after allogeneic transplantation. *The Journal of Clinical Investigation*, 2020. doi: 10.1172/JCI133960

(8) ★**Williamson BD** and ★Feng J. Efficient nonparametric statistical inference on population feature importance using Shapley values. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 10282–10291, 2020. URL http://proceedings.mlr.press/v119/williamson20a.html

(7) Neidich SD, Fong Y, Li SS, Geraghty DE, **Williamson BD**, Young WC, Goodman D, Seaton KE, Shen X, Sawant S, Zhang L, deCamp AC, Blette BS, Shao M, Yates NL, Feely F, Pyo CW, Ferrari G, Frank I, Karuna ST, Swann E, Mascola J, Graham BS, Hammer SM, Sobieszczyk ME, Corey L, Janes H, McElrath MJ, Gottardo R, Gilbert PB, and Tomaras GD. Antibody Fc effector functions and IgG3 associate with decreased HIV-1 risk. *The Journal of Clinical Investigation*, 2019. doi: 10.1172/JCI126391

(6) ★Magaret CA, ★Benkeser DC, ★**Williamson BD**, Borate BR, Carpp LN, Georgiev IS, Setliff I, Dingens AS, Simon N, Carone M, Simpkins C, Montefiori D, Alter G, Yu WH, Juraska M, Edlefsen PT, Karuna S, Mgodi NM, Edugupanti S, and Gilbert PB. Prediction of VRC01 neutralization sensitivity by HIV-1 gp160 sequence features. *PLOS Computational Biology*, 2019. doi: 10.1371/journal.pcbi.1006952

(5) Hanscom B, Hughes JP, **Williamson BD**, and Donnell D. Adaptive non-inferiority margins under observable non-constancy. *Statistical Methods in Medical Research*, 2018. doi: 10.1177/0962280218801134

(4) ★Feng J, ★**Williamson BD**, Carone M, and Simon N. Nonparametric variable importance using an augmented neural network with multi-task learning. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1495–1504, 2018. URL http://proceedings.mlr.press/v80/feng18a.html

(3) Anchang B, Davis KL, Fienberg H, **Williamson B**, Bendall SC, Karacosta L, Tibshirani R, Nolan GP, and Plevritis SK. DRUG-NEM: optimizing drug combinations using single-cell perturbation response to account for intratumoral heterogeneity. *Proceedings of the National Academy of Sciences*, 115(18):E4294–E4303, 2018. doi: 10.1073/pnas.1711365115

(2) Safren SA, Hughes JP, Mimiaga MJ, Moore AT, Friedman RK, Srithanaviboonchai K, Limbada M, ♣**Williamson BD**, Elharrar V, Cummings V, Magidson JF, Gaydos CA, Celentano D, and Mayer KH for the HPTN063 Study Team. Frequency and predictors of estimated HIV transmissions and bacterial STI acquisition among HIV-positive patients in HIV care across three continents. *Journal of the International AIDS Society*, 19, 2016. doi: 10.7448/IAS.19.1.21096

(1) Ritchwood TD, Hughes JP, Jennings L, MacPhail C, ♣**Williamson B**, Selin A, Kahn K, Gómez-Olivé XF, and Pettifor A. Characteristics of age-discordant partnerships associated with HIV risk among young South African women (HPTN 068). *Journal of Acquired Immune Deficiency Syndromes*, 72:423–429, 2016. doi: 10.1097/QAI.0000000000000988

(*Submitted or under revision*)

(1) **Williamson BD** and Huang Y. Flexible variable selection in the presence of missing data. *arXiv*, 2022. URL https://arxiv.org/abs/2202.12989

(b) Other refereed scholarly publications

(1) **Williamson BD**, Gilbert PB, Carone M, and Simon N. Rejoinder to "Nonparametric variable importance assessment using machine learning techniques". *Biometrics*, 2020. doi: 10.1111/biom.13389

(c) Books and book chapters

(d) Other non-refereed scholarly publications

(1) Gilbert PB, Fong Y, Benkeser D, Andriesen J, Borate B, Carone M, Carp LN, Díaz I, Fay MP, Fiore-Gartland A, Hejazi NS, Huang Y, Huang Y, Hyrien O, Janes HE, Juraska M, Li K, Luedtke A, Nason M, Randhawa AK, van der Laan L, **Williamson BD**, Zhang W, and Follman D. USG COVID-19 Response Team / CoVPN Vaccine Efficacy Trial Immune Correlates Statistical Analysis Plan. 2021. doi: 10.6084/m9.figshare.13198595

Software

`vimp` perform inference on algorithm-agnostic variable importance (available on CRAN)

`vimpy` perform inference on algorithm-agnostic variable importance in Python (available on PyPI)

`paramedic` Predicting Absolute and Relative Abundance by Modeling Efficiency to Derive Intervals and Concentrations

SLAPNAP Super Learner Predictions using NAb Panels (available on DockerHub)

`rigr` regression, inference, and general data analysis tools for R

`flevr` flexible, ensemble-based variable selection with potentially missing data

Patents and Intellectual Property

Not applicable

Funding History

(a) Current funded projects

(1) **Evaluating Influenza, SARS-CoV-2, and Other Respiratory Virus Vaccine Effectiveness in Prevention of Acute Illness in Washington State 2022–2027**, CDC (U01 IP001191), PI: Wernli, Co-I, 0.1–0.2 FTE, 01/2023–09/2027

(2) **Sentinel Scalable NLP**, FDA, PIs: Carrell and Smith, Co-I, 0.05–0.15 FTE, 08/2022–01/2023.

(3) **Reducing CNS-active Medications to Prevent Falls and Injuries in Older Adults**, CDC (U01 CE002967), PI: Phelan, Co-I, 0.20–0.40 FTE, 09/30/2018–09/29/2023.

(4) **Engaging Staff to Improve COVID-19 Vaccination Rates at Long Term Care Facilities (ENSPIRE)**, PCORI (COVID-2021C2-1368), PI: Hsu, Co-I, 0.20–0.40 FTE, 8/1/2021–1/31/2025.

(5) **Adaptive Treatment Strategy Methods for Electronic Health Records**, NIMH, PIs: Shortreed and Moodie, Co-I, 0.25 FTE, 09/14/2021–09/29/2022.

(6) **Sentinel Advanced Phenotyping**, FDA, PIs: Carrell and Nelson, Co-I, 0.05–0.15 FTE, 09/14/2021–09/29/2022.

(b) Completed or past funded projects

(1) **SDMC: HIV Vaccine Trials Network**, NIAID (UM1 AI068635), PIs: Gilbert, Huang, and Janes, Postdoctoral Research Fellow, 07/2021–09/2021.

(2) **Nutrition and Physical Activity Assessment Study**, NCI (R01 CA119171), PI: Neuhouser, Postdoctoral Research Fellow, 03/2021–09/2021.

(3) **Statistical Methods for HIV-1 Immune Correlates Studies**, NIAID (R01 AI122991), PI: Fong, NIAID, Postdoctoral Research Fellow, 04/2021–07/2021.

(4) **Detection and Prognosis of Early-Stage Pancreatic Cancer by Interdependent Plasma Markers**, NCI (U01 CA152653), PI: Haab, Postdoctoral Research Fellow, 03/2020–07/2021.

(5) **A Clinical Validation Center for Early Detection of Pancreatic Cancer**, NCI (U01 CA152653), PI: Maitra, Postdoctoral Research Fellow, 02/2020–03/2021.

(6) **Statistical Methods for Selection and Evaluation of Biomarkers**, NIGMS (R01 GM106177), PI: Huang, Postdoctoral Research Fellow, 01/2020–01/2021.

(7) **Statistical Methods in HIV Vaccine Efficacy Trials**, NIAID (R37 AI054165), PI: Gilbert, Postdoctoral Research Fellow, 03/2020–06/2020.

(8) **How Did a Vaccine Enhance HIV Acquisition?**, NIAID (R01 AI118590), PI: Miller, Postdoctoral Research Fellow, 01/2020–05/2020.

(9) **Evaluating predictors of HIV vaccine efficacy: Statistical methods for estimation, testing, and inference**, NIAID (F31 AI140846), PI: Williamson, PI, 09/2018–12/2019.

(10) **Statistical Methods for AIDS Research**, NIAID (R01 AI029168), PI: Hughes, Research Assistant, 0.5 FTE, 09/2015–09/2018.

Conferences and Symposiums

Conference organization:

1. Organizer and session chair, “Methods for inference on variable importance using machine learning”, International Chinese Statistical Association (ICSA) Applied Statistics Symposium, 2022 (invited session)
2. Co-chair, Keynote and short talks, Hutch United Symposium, 2022 (invited session)
3. Session chair, “Quantification, Association Testing, and Integration of the Microbiome”, Joint Statistical Meetings, 2017 (invited session)

Oral presentations: (the symbol [I] denotes an invited presentation)

23. “Inference for Model-Agnostic Longitudinal Variable Importance”. 15th International Conference of the European Consortium for Informatics and Mathematics Working Group on Computational and Methodological Statistics (CMStatistics). London, UK, 2022. [I]
22. “Scalable Phenotyping for Safety Outcomes Using Electronic Health Record Data”. Forum on the Integration of Observational and Randomized Data. Washington, DC, 2022. [I] *(upcoming)*
21. “Inference for Model-Agnostic Variable Importance”. ICSA Applied Statistics Symposium. Gainesville, FL, 2022. [I]
20. “Flexible Variable Selection in the Presence of Missing Data”. WNAR of the International Biometric Society. Virtual, 2022.
19. “Model-Agnostic Variable Importance and Selection”. University of Washington Department of Biostatistics Seminar. Virtual, 2022. [I]
18. “Inference for Model-Agnostic Variable Importance”. ASA Statistical Learning and Data Science Section Webinar Series. Virtual, 2022. [I]
17. “Flexible Variable Selection in the Presence of Missing Data”. Joint Statistical Meetings. Virtual, 2021.
16. “Inference for Algorithm-Agnostic Variable Importance”. WNAR of the International Biometric Society. Virtual, 2021. [I] *(presented in place of Marco Carone)*
15. “Inference for Model-Agnostic Variable Importance”. Third Annual Hutch United Symposium. Virtual, 2021. *(keynote)* [I]

14. "Inference for Model-Agnostic Variable Importance". Kaiser Permanente Washington Health Research Institute Seminar. Virtual, 2021. [I]

13. "Statistical Inference and Containerization in Computational Pipelines". Fred Hutchinson Cancer Research Center Biostatistics Program Seminar. Virtual, 2021. [I]

12. "A Unified Approach to Inference on Algorithm-Agnostic Variable Importance". Vanderbilt University Department of Biostatistics Seminar. Virtual, 2020. [I]

11. "Efficient Nonparametric Statistical Inference on Population Feature Importance using Shapley Values". Thirty-seventh International Conference on Machine Learning. Virtual, 2020.

10. "Guiding HIV-1 Antibody Regimen Down-Selection and Prevention Efficacy Trial Design Using Machine Learning". 27th International Dynamics and Evolution of HIV and Other Human Viruses Meeting. Virtual, 2020.

9. "A Unified Approach to Nonparametric Variable Importance Assessment". Joint Statistical Meetings. Denver, CO, 2019 (selected for an ASA Nonparametrics Section Travel Award).

8. "Assessing Variable Importance Nonparametrically using Machine Learning Techniques". University of Washington Department of Biostatistics Colloquium. Seattle, WA, 2018. [I]

7. "Assessing Variable Importance Nonparametrically using Machine Learning Techniques". Joint Statistical Meetings. Vancouver, BC, Canada, 2018 (selected for an ASA Biometrics Section Travel Award).

6. "Nonparametric Variable Importance Using an Augmented Neural Network with Multi-Task Learning". Thirty-fifth International Conference on Machine Learning. Stockholm, Sweden, 2018.

5. "Assessing Variable Importance Nonparametrically using Machine Learning Techniques". WNAR of the International Biometric Society. Santa Fe, NM, 2017 (selected as the Most Outstanding Oral Paper).

4. "Assessing Variable Importance Nonparametrically using Machine Learning Techniques". University of Washington Department of Biostatistics Student Seminar. Seattle, WA, 2017.

3. "An Introduction to Targeted Learning". University of Washington Department of Biostatistics Student Seminar. Seattle, WA, 2017.

2. "Shrinkage Estimators for High-Dimensional Covariance Matrices". Pomona College Mathematics Seminar. Claremont, CA, 2014.

1. "Automating Cell Gating and Creating a Nested Effects Model to Compare Drug Effects". Stanford University Center for Cancer Systems Biology Meeting. Stanford, CA, 2013.

Poster presentations:

5. "Nonparametric Variable Importance using an Augmented Neural Network with Multi-Task Learning". Thirty-fifth International Conference on Machine Learning. Stockholm, Sweden, 2018.

4. "HIV-1 Sequence Predictors of VRC01 Neutralization Sensitivity". 25th International Dynamics and Evolution of HIV and Other Human Viruses Meeting. Leavenworth, WA, 2018.

3. "Assessing Variable Importance Nonparametrically using Machine Learning Techniques". University of Washington Biostatistics Department Retreat. Seattle, WA, 2017.

2. "Assessing Variable Importance Nonparametrically using Machine Learning Techniques". Joint Statistical Meetings. Baltimore, MD, 2017.

1. "Assessing Variable Importance Nonparametrically". University of Washington Biostatistics Department Retreat. Blaine, WA, 2015.

Non-technical, Outreach, and Mentoring presentations:

5. Moderator for "Interdisciplinary Science Panel". Seattle Central College MESA-LSAMP Program. Virtual, 2021.
4. "Statistics in Infectious Disease Research". Roanoke Valley Governor's School for Science and Technology Computational Biology Course. Virtual, 2020.
3. Panelist for "LSAMP Virtual Research Panel". Seattle Central College MESA-LSAMP Program. Virtual, 2020.
2. "Fellowships, scholarships, and grants". Biostatistics Student Seminar, University of Washington. Seattle, WA, 2018.
1. "Travel grants and conference funding". University of Washington Department of Biostatistics. Seattle, WA, 2018.

Teaching History

(a) Formal Courses, including Distance Learning

In the University of Washington School of Public Health Math and R Skills Preparatory Workshop

Summer 2016 (co-taught with G Tarr): 175 students

Summer 2017 (co-taught with J Williams-Nguyen): 175 students

Summer 2018 (co-taught with J Williams-Nguyen): 175 students

At the University of Washington

(1) Pre-Doctoral Instructor, Biostatistics 311 – Regression Methods in the Health Sciences Spring 2018 (co-taught with K Grinde), enrollment: 12; median evaluation: 4.9/5.0

(b) Other Teaching

(c) Independent Study

(1) Ziyi Chen, MS Capstone candidate (with RY Coley and S Shortreed), 01/2022–06/2022
(2) Emily Minus, MS Thesis candidate (with RY Coley and S Shortreed), 01/2022–09/2022
(3) Runjia Zou, MS Capstone candidate (with RY Coley and S Shortreed), 01/2022–06/2022

Advising and Formal Mentoring

(a) PhD Dissertations, chair

(b) Master's Theses, chair

(1) Emily Minus (co-chaired with RY Coley)

(c) Mentored Scientists and Postdocs

(d) MS and PhD committees in non-chair role

(e) Other mentoring

(1) Drew King, Fred Hutch / Seattle Central College MESA Intern (with Y Huang), 03/2021–08/2021
(2) Courtney Simmons, Emory University MS Capstone Project (with D Benkeser), 02/2021–05/2021