

Amy D Willis, Ph.D.

University of Washington
Department of Biostatistics
F-657 Health Sciences Building
Seattle, WA 98195

Phone: (206) 543-8027
Email: adwillis@uw.edu
Homepage: statisticaldiversitylab.com

Education

- Australian National University, B Actuarial Studies (Honours), 12/2011
- Cornell University, M.S., Statistics, 01/2015
- Cornell University, Ph.D., Statistics, 05/2017

Professional Positions

- Senior Analyst, Macroeconomics, Canberra, Australia, 08/2009-06/2012
- Research, Data Management and Audit Officer, ACT NOWaste, Canberra, Australia, 11/2011-06/2012
- Statistical Consultant, Zadeh Laboratory, College of Human Ecology, Cornell University, 12/2014-05/2015
- Quantitative Analyst Intern, Google, Mountain View, CA, 05/2015-07/2015
- Assistant Professor, Department of Biostatistics, University of Washington, Seattle, WA, 06/2017-present
 - Genentech Endowed Professorship in Biostatistics, 07/2018-present
 - Adjunct Assistant Professor, Department of Statistics, 10/2017-present
 - Member, Center for Microbiome Sciences and Therapeutics (CMiST), UW Medicine, 2017-present
 - Program Faculty, Quantitative Ecology and Resource Management (QERM), UW College of the Environment, 2019-present
 - Affiliate Faculty, UW eScience Institute, 2019-present

Honors, Awards, Scholarships

- National Undergraduate Scholarship, Australian National University, 2008-2011
- Australian National University College of Business and Economics Summer Research Scholarship, 2009-2010.
- John XXIII College Academic Excellence Scholarship, 2008, 2009
- First Class Honours, Australian National University, 2011
- 3-Minute Thesis Finalist, Cornell University, 2015
- College Banner Bearer, Cornell University, 2017

Professional Activities

- Referee Service: *Annals of Applied Statistics*; *Annals of Statistics*; *Bayesian Analysis*; *Bioinformatics*; *Biometrics*; *Biostatistics*; *Environmental and Ecological Statistics*; *Environmental Microbiology*; *Genome Biology*; *Journal of Agricultural, Biological, and Environmental Statistics*; *Journal of Non-parametric Statistics*; *Microbiome*; *mSystems*; *Nature Ecology & Evolution*; *Nature Medicine*; *PLoS Biology*; *PLoS Computational Biology*; *Scandinavian Journal of Statistics*
- Guest Editor Service: *PLoS Computational Biology* (2017)
- Associate Editor: *Biostatistics* (2018+)
- Associate Editor: *Microbiome* (2018+)
- Associate Editor: *Journal of the American Statistical Association, Applications and Case Studies* (2019+)
- Professional Memberships: American Statistical Association; International Biometric Society (Western North American Region); National Geographic Society

Bibliography

(a) Refereed research articles

- 1) John Bunge, Amy Willis, and Fiona Walsh. Estimating the number of species in microbial diversity studies. *Annual Review of Statistics and Its Application*, 1:427–445, 2014
- 2) Jacob P Christ, Amy D Willis, Eric D Brooks, Heidi Vanden Brink, Brittany Y Jarrett, Roger A Pierson, Donna R Chizen, and Marla E Lujan. Follicle number, not assessments of the ovarian stroma, represents the best ultrasonographic marker of polycystic ovary syndrome. *Fertility and sterility*, 101(1):280–287, 2014
- 3) Arindam RoyChoudhury, Amy Willis, and John Bunge. Consistency of a phylogenetic tree maximum likelihood estimator. *Journal of Statistical Planning and Inference*, 161:73–80, 2015
[Corresponding author]
- 4) Amy Willis and John Bunge. Estimating diversity via frequency ratios. *Biometrics*, 71(4):1042–1049, 2015
- 5) Gennady Samorodnitsky, Sidney Resnick, Don Towsley, Richard Davis, Amy Willis, and Phyllis Wan. Nonstandard regular variation of in-degree and out-degree in the preferential attachment model. *Journal of Applied Probability*, 53(1):146–161, 2016
- 6) Heidi Vanden Brink, Amy D Willis, Brittany Y Jarrett, Annie W Lin, Steven Soler, Siân Best, Erica L Bender, Andrew K Peppin, Kathleen M Hoeger, and Marla E Lujan. Sonographic markers of ovarian morphology, but not hirsutism indices, predict serum total testosterone in women with regular menstrual cycles. *Fertility and sterility*, 105(5):1322–1329, 2016
- 7) Amy Willis. Extrapolating abundance curves has no predictive power for estimating microbial biodiversity. *Proceedings of the National Academy of Sciences*, 113(35):E5096–E5096, 2016

- 8) Nancy Chan, Amy Willis, Naomi Kornhauser, Maureen M Ward, Sharrell B Lee, Eleni Nackos, Bo Ri Seo, Ellen Chuang, Tessa Cigler, Anne Moore, and others. Influencing the Tumor Microenvironment: A Phase II Study of Copper Depletion Using Tetrathiomolybdate in Patients with Breast Cancer at High Risk for Recurrence and in Preclinical Models of Lung Metastases. *Clinical Cancer Research*, 23(3):666–676, 2017
 - 9) Clément P Bataille, Amy Willis, Xiao Yang, and Xiao-Ming Liu. Continental igneous rock composition: A major control of past global chemical weathering. *Science Advances*, 3(3):e1602183, 2017
 - 10) Amy Willis, John Bunge, and Thea Whitman. Improved detection of changes in species richness in high-diversity microbial communities. *Journal of the Royal Statistical Society: Series C*, 66(5):963–977, 2016
 - 11) Sheena Sahota, Naomi Kornhauser, Amy Willis, Maureen M Ward, Tessa Cigler, Anne Moore, Eleni Andreopoulou, Veronica Fitzpatrick, Sarah E Schneider, Alysia Wiener, Anna Rubinchik, Sharrell Lee, Maureen E Lane, Vivek Mittal, and Linda T Vahdat. A phase II study of copper-depletion using tetrathiomolybdate (TM) in patients (pts) with breast cancer (BC) at high risk for recurrence: Updated results. *Journal of Clinical Oncology*, 35(15):2557–2557, 2017
 - 12) Amy D Willis and Rayna Bell. Uncertainty in phylogenetic tree estimates. *Journal of Computational and Graphical Statistics*, 27(3):542–552, 2018
 - 13) Amy D Willis. Confidence Sets for Phylogenetic Trees. *Journal of the American Statistical Association*, 114(525):235–244, 2019
 - 14) Nicolas Tromas, Zofia E Taranu, Bryan D Martin, Amy Willis, Nathalie Fortin, Charles W Greer, and B Jesse Shapiro. Niche Separation Increases With Genetic Distance Among Bloom-Forming Cyanobacteria. *Frontiers in Microbiology*, 9:109, 2018
 - 15) Amy D. Willis. Rigorous statistical methods for rigorous microbiome science. *mSystems*, In Press, 2019
 - 16) Anna E. Sarfaty, Caroline J. Zeiss, Amy D. Willis, Jorgen Harris, and Peter C. Smith. Concentration dependent toxicity following subcutaneous meloxicam administration in C57Bl/6N Mice. *Journal of the American Association of Laboratory Animal Science*, In Press, 2019
 - 17) Samuel S Minot and Amy D Willis. Clustering co-abundant genes identifies components of the gut microbiome that are reproducibly associated with colorectal cancer and inflammatory bowel disease across independent cohorts. *Under revision*.
 - bioRxiv: doi.org/10.1101/567818
 - 18) Michael McLaren, Amy Willis and Benjamin Callahan. Consistent and comprehensible bias in metagenomic sequencing experiments. *Under revision*.
 - bioRxiv: doi.org/10.1101/559831
- (b) Other refereed scholarly publications

- 1) Amy Donaldson Willis. Confidence procedures for phylogenetic trees. May 2017 (PhD Thesis, Cornell University)
- (c) Other non-refereed published scholarly publications
 - 1) Amy Willis. Species richness estimation with high diversity but spurious singletons.
 - arXiv: [arXiv:1604.02598](https://arxiv.org/abs/1604.02598)
 - 2) Amy Willis. Rarefaction, alpha diversity, and statistics.
 - bioRxiv: doi.org/10.1101/231878
 - 3) Amy D Willis, and Bryan D Martin. DivNet: Estimating diversity in networked communities.
 - bioRxiv: doi.org/10.1101/305045
 - 4) Evan Bolyen, Jai Ram Rideout, Matthew R Dillon, Nicholas A Bokulich, ..., Bryan Martin, ..., Pauline Trinh, ..., Amy D Willis, ..., Rob Knight, J Gregory Caporaso. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science.
 - PeerJ: doi.org/10.7287/peerj.preprints.2729502
 - 5) Bryan Martin, Daniela Witten* and Amy Willis*. Modeling microbial abundances and dysbiosis with beta-binomial regression. (* = joint last author)
 - arXiv: [arXiv:1902.02776](https://arxiv.org/abs/1902.02776)
 - 6) Jean Feng, Will DeWitt, Aaron McKenna, Noah Simon, Amy Willis and Frederick A Matsen IV. Inferring cell lineage trees from CRISPR arrays using statistical phylogenetics.
 - bioRxiv: doi.org/10.1101/595215

Patents and Other Intellectual Property

- Publicly available software
 1. breakaway, an R package for estimating and modelling microbial diversity
 - Stable release: cran.r-project.org/web/packages/breakaway/index.html
 - Development version: github.com/adw96/breakaway
 - QIIME2 plug-in: github.com/statdivlab/q2-breakaway
 2. CatchMore, an R package for species richness estimation
 - Development version: github.com/statdivlab/CatchMore
 3. DivNet, an R package for estimating diversity in networked communities
 - Stable release: github.com/adw96/DivNet
 4. corncob, an R package for modeling microbial abundances
 - Stable release: github.com/bryandmartin/corncob (mentored student author: Bryan Martin)
 - QIIME2 plug-in: github.com/statdivlab/q2-corncob (mentored student author: Pauline Trinh)

Funding History

1. Funded projects

- (a) University of Wisconsin-Madison Data Science Initiative. Role: co-PI. Title: Addressing misclassification in the microbiome: A data-scientific approach to propagating uncertainty in microbial community composition. Collaborators: Thea Whitman (co-PI, University of Wisconsin-Madison) and Karl Broman (co-I, University of Wisconsin-Madison).
- (b) National Institutes of Health, NIEHS, R01. Role: co-I. Title: Developmental PBDE exposure, gut microbiome, and diabetes. Collaborators: Julia Yue Cai (PI, University of Washington).
- (c) Amazon AWS Cloud Credits for Research. Role: PI. Title: Scaling statistical methods for the analysis of shotgun metagenomics data using distributed computing.

Conferences and Symposia

* = invited

- 1. March 2014: Ordered Data Analysis, Models, and Health Research Methods: HNN60; A ratio-based method for estimating an unknown number of classes
- 2. August 2014: Joint Statistical Meetings; id.
- 3. October 2014: Université de Montréal, Microbial Evolutionary Genomics Group*; Statistical inference for estimating and modeling diversity
- 4. June 2015: Google HQ, Decision Support Research Group*; Species Richness Estimation
- 5. August 2015: Joint Statistical Meetings, Microbiome in Epidemiology Section*; Righting some wrongs in the statistical analysis of microbial community data
- 6. December 2015: Queensland University of Technology, ARC Centre of Excellence for Mathematical & Statistical Frontiers*; Taxonomic richness: estimation and comparison
- 7. September 2016: Pontificia Universidad Javeriana, Departamento de Matemáticas*; Teoremas del límite central y conjuntos de confianza para árboles
- 8. September 2016: Universidad de los Andes, Departamento de Matemáticas*; id.
- 9. October 2016: University of Glasgow, School of Mathematics and Statistics*; Biodiversity Statistics
- 10. October 2016: Newcastle University, School of Mathematics and Statistics*; Confidence Sets for Phylogenetic Trees
- 11. December 2016: Université de Montreal, Département de mathématiques et de statistique*; id.
- 12. January 2017: Pennsylvania State University, Department of Statistics*; id.
- 13. January 2017: University of Texas at Austin, Department of Statistics and Data Sciences*; id.
- 14. January 2017: University of Michigan, Department of Statistics*; id.
- 15. January 2017: Carnegie Mellon University, Department of Statistics*; id.
- 16. February 2017: University of Wisconsin-Madison, Department of Biostatistics and Medical Informatics*; id.

17. February 2017: University of Pennsylvania, The Wharton School, Statistics Department*; id.
18. February 2017: Columbia University, Department of Statistics*; id.
19. February 2017: University of Washington, Department of Biostatistics*; id.
20. February 2017: Broad Institute of MIT, Second Workshop on Statistical and Algorithmic Challenges in Microbiome Data Analysis*; Visualizing uncertainty in phylogenetic trees.
21. March 2017: Cornell University, Department of Statistics; Confidence Procedures for Phylogenetic Trees.
22. October 2017: University of Washington, Department of Statistics*; id.
23. October 2017: University of Washington, Centre for Microbiome Sciences and Therapeutics; Statistical Analysis of Alpha Diversity.
24. November 2017: Rice University, Department of Statistics*; Confidence Sets for Phylogenetic Trees.
25. November 2017: University of Connecticut, Department of Statistics*; id.
26. November 2017: University of Washington, Centre for Microbiome Sciences and Therapeutics*; Adjusting for Multiple Testing in Microbiome Data Analysis.
27. December 2017: University of Washington, Microbiome Symposium*; Statistical Considerations for Microbiome Research.
28. January 2018: University of Washington, Mini-Symposium: Emerging Opportunities in Microbiomes and Environmental Health*; Statistical summaries and uncertainties in microbiome data analysis.
29. March 2018: Fred Hutchinson Cancer Research Centre, Public Health Sciences Division: Microbiome Data to Knowledge*; Microbial community structure data analysis.
30. June 2018: 4th Conference of the International Society for Nonparametric Statistics; Invited Session *Graphs and Trees in Object Oriented Data Analysis**. Estimation and testing of functions of phylogeny and abundance.
31. January 2019: 2019 Workshop on Genomics, Český Krumlov; **Invited to teach session** on *Microbiome Data& Analysis*.
32. February 2019: University of Washington, Department of Genomes Sciences and the Computational Molecular Biology Program, COMBI seminar*. Estimating diversity and relative abundance in microbial communities.
33. February 2019: Indiana University, Department of Statistics*; id.
34. February 2019: Fred Hutchinson Cancer Research Center, Biostatistics division*; id.
35. March 2019: Institute for Systems Biology, Seattle, Washington*; id.
36. March 2019: University of Georgia, Department of Marine Sciences*; id.
37. March 2019: University of Georgia, Department of Statistics; id.
38. April 2019: Third Workshop on Statistical and Algorithmic Challenges in Microbiome Data Analysis*; Measurement Error and Inference in Human 16S Studies (collaborative work presented by trainee David Clausen).

39. (May 2019: Symposium on Data Science and Statistics; Invited session *Computational Statistics for Large-Scale Biological Data*; organiser: Jacob Bien*)
40. (May 2019: Symposium on Data Science and Statistics; Invited session *Deciphering Biological Systems via Innovative Statistical Learning Methods*; organiser: Tian Zhang*)
41. (June 2019: 2019 WNAR/IMS/JR Annual Meeting; Invited session *Microbiome Data Analysis*; organiser: Debashis Mondal*)
42. (December 2019: 4th Thünen Symposium on Soil Metagenomics; Braunschweig, Germany; **Keynote speaker**; organiser: Christoph Tebbe*)

University Service

- Faculty Fellows Program 2017, Participant
- Department of Biostatistics Diversity Committee, Member (07/2017-08/2018)
- Department of Biostatistics Seminar Organizer (01/2018-03/2018)
- Panelist at University of Washington Career and Internship Center's 2018 workshop *Academic Salary Negotiations*
- Search committee member, Center for Microbiome Sciences and Therapeutics, UW Medicine (07/2018-present)
- Womxn in Biostatistics and Statistics Faculty Liaison (08/2018-present)

Other Information

- Popular media
 - Invited interviews
 - * Podcast on bioinformatics.chat. *Rarefaction, alpha diversity, and statistics with Amy Willis*. Interviewer: Roman Cheplyaka, January 2018.
 - * Interview on feminasci.com. *Interview: Early Career Quantitative Scientists*. Moderator: Erin Eggleston, April 2018
 - Twitter: @AmyDWillis

Teaching History

1. Formal courses
 - Summer Institute in Statistical Genetics, and Summer Institute in Statistics and Modelling in Infectious Diseases (crosslisted courses)
 - (a) Module 1: Probability and Statistical Inference, July 2017
 - Overall Summative Rating by students: 4.3/5
 - (b) Module 1: Probability and Statistical Inference, July 2018
 - Overall Summative Rating by students: 4.5/5

- Department of Biostatistics, University of Washington
 - (a) BIOST 561: Computational Skills for Biostatistics I, Autumn 2017
 - Advanced R programming for graduate students (1 credit)
 - Overall Summative Rating by students: 4.6/5
 - (b) BIOST 582: Student Seminar, Winter 2018
 - (c) (BIOST 561: Computational Skills for Biostatistics I, Spring 2019)
 - (d) (BIOST 533: Theory of Linear Models, Spring 2019)

2. Other teaching

- Teaching Assistantships at Cornell University
 - (a) STSCI2100: Introductory Statistics; Fall 2012
 - (b) STSCI4100: Multivariate Analysis; Spring 2013
 - (c) BTRY6010: Statistical Methods 1; Fall 2013
 - (d) STSCI4140: Applied Design; Spring 2015, Spring 2017
 - (e) STSCI3100: Statistical Sampling; Fall 2014, Fall 2015
- Teaching Assistantships at Australian National University
 - (a) STAT1008: Quantitative Research Methods; Semester 2 2009, Semester 1 2010, Semester 2 2010, Semester 2 2011
 - (b) STAT1003: Statistical Techniques; Semester 1 2011
- Lecturing at Marine Biological Laboratory
 - (a) Strategies and Techniques for Analyzing Microbial Population Structures
 - i. August 2013: *breakaway: Diversity estimation software*
 - ii. August 2014: *The peanut butter sandwich of beta diversity*
 - iii. August 2015: *Biodiversity: A practical guide*
 - iv. August 2016: *Honest statistics for microbial ecologists; Species richness for microbial ecologists; Rarefaction and why you shouldn't*
 - v. August 2017: *Introduction to R; R lab; alpha and beta diversity analysis; Diversity lab*
 - vi. August 2018: *Statistics & statistical thinking; Statistical estimation: diversity, abundance, shotgun data; Diversity lab*
 - vii. (July 2019)

3. Independent Study Supervision

- University of University of Washington, Department of Biostatistics
 - (a) Edward Zhao, University of Washington, Department of Biostatistics, PhD Candidate, Winter Quarter 2018.
 - (b) Kaleigh Ervin, University of Washington, Department of Biostatistics, PhD Candidate, Winter Quarter 2018.
 - (c) Alex Paynter, University of Washington, Department of Biostatistics, MS Candidate, Spring Quarter 2018, Summer Quarter 2018, Fall Quarter 2018, Winter Quarter 2019 & Spring Quarter 2019.

- (d) Pauline Trinh, University of Washington, Department of Environmental and Occupational Health Sciences, PhD Candidate, Spring Quarter 2018, Summer Quarter 2018, Fall Quarter 2018, Winter Quarter 2019 & Spring Quarter 2019.
- (e) Jean Feng, University of Washington, Department of Biostatistics, PhD Candidate, Spring Quarter 2018.
- (f) Kendrick Qijun Li, University of Washington, Department of Biostatistics, PhD Candidate, Spring Quarter 2018, Summer Quarter 2018, Fall Quarter 2018 & Winter Quarter 2019.
- (g) Bryan Martin, University of Washington, Department of Statistics, PhD Candidate, Summer Quarter 2018, Fall Quarter 2018, Winter Quarter 2019 & Spring Quarter 2019.
- (h) David Clausen, University of Washington, Department of Biostatistics, PhD Candidate, Summer Quarter 2018, Fall Quarter 2018, Winter Quarter 2019 & Spring Quarter 2019.
- (i) Charles Wolock, University of Washington, Department of Biostatistics, PhD Candidate, Fall Quarter 2018, Winter Quarter 2019 & Spring Quarter 2019.

Advising and Formal Mentoring

1. PhD Dissertations, chair

- (a) Bryan Martin, University of Washington, Department of Statistics, PhD Candidate (co-chair with Daniela Witten)
- (b) David Clausen, University of Washington, Department of Biostatistics, PhD Candidate

2. Masters Theses, chair

- (a) Alex Paynter, University of Washington, Department of Biostatistics, MS Candidate
 - Alex was the Winner of the Outstanding MS Student Award in Biostatistics, 2019

3. Mentored Scientists and Postdoctoral Fellows

- Postdoctoral Supervision
 - (a) Jacob Price, University of Wisconsin, Department of Soil Science (co-mentored with Thea Whitman), December 2018+

4. MS and PhD committees in non-chair role

- Doctoral Committee Member
 - (a) Kelsey Grinde, University of Washington, Department of Biostatistics, PhD Candidate
- Graduate Student Representative
 - (a) Andria Ellis, University of Washington, Department of Genome Sciences, PhD Candidate
 - (b) Leandra Brettner, University of Washington, Department of Bioengineering, PhD Candidate
 - (c) Andrew Gustin, University of Washington, Department of Pathobiology, PhD Candidate

5. Other Mentoring

- Project Supervision (outside of independent study supervision)
 - (a) Jean Feng, University of Washington, Department of Biostatistics, PhD Candidate (project co-supervised with Erick Matsen and Noah Simon)
 - (b) Brian Williamson, University of Washington, Department of Biostatistics, PhD Candidate (project co-supervised with Jim Hughes)

6. Academic Advising

- Department-assigned Academic Advisor
 - (a) Alexander Paynter, University of Washington, Department of Biostatistics, MS Candidate
 - (b) Taylor Okonek, University of Washington, Department of Biostatistics, PhD Candidate

Last updated: May 24, 2019