Zachary Ryan McCaw

Curriculum Vitae

GitHub, Google Scholar, LinkedIn, ORCID, Website

zmccaw@alumni.harvard.edu

September 2026

Education

Stanford University
01/2021 – 06/2022

Graduate Certificate in Artificial Intelligence

- Coursework: Computer Vision, Deep Learning, Reinforcement Learning.

Harvard University
08/2014 – 05/2019

Ph.D. in Biostatistics, A.M. in Biostatistics

- Dissertation: Transformation and multivariate methods for improving power in GWAS.
  - Studied operating characteristics of the rank-based inverse normal transformation for genome-wide association studies of quantitative traits. [19]
  - Developed multivariate regression methodology for leveraging a correlated surrogate outcome to improve inference on a partially missing target outcome. [15, 12]
- Advisor: Xihong Lin, Ph.D.
- Committee: Martin Aryee, Ph.D., and Jeffrey Miller, Ph.D.

University of North Carolina at Chapel Hill
08/2009 – 05/2013

B.S.P.H. in Biostatistics, B.S. in Quantitative Biology

- Graduated with highest distinction; Phi Beta Kappa Honors Society.

Professional Experience

Causal Labs
04/2026 – Present

Member of Technical Staff

- Researching and developing learned world models to support planning and policy learning in operations-research settings.

Voleon
01/2025 – 03/2026

Senior Member of Research Staff

- Projects
  - Conducted independent research on transformer architectures for financial time-series forecasting, comparing temporal and cross-asset attention as well as univariate and multivariate forecasting formulations.
  - Investigated self-supervised representation learning using variational, discrete, and masked autoencoders and contrastive objectives to support downstream forecasting.

Insitro

09/2021 – 01/2025

Staff Machine Learning Scientist: 04/23 – 01/25
Senior Machine Learning Scientist: 09/21 – 03/23

- Department: Statistical Genetics and Clinical Machine Learning
- Team Leads: Colm O'Dushlaine, Ph.D., and Thomas Soare, Ph.D.
- Projects
  - Evaluated self-supervised training strategies for vision transformers, including SimCLR and DINO, across diffusion MRI, OCT, and whole-slide histopathology; applied learned representations to GWAS, rare-variant association studies, and survival prediction.
  - Led the four-member Genetic Discovery team within Oncology, using imaging-derived phenotypes to identify rare germline associations and prioritize candidate therapeutic targets.
  - Co-led development of EmbedGEM, a framework for evaluating learned phenotypes based on both genetic association strength and clinical relevance, motivated by the validation challenges accompanying increased discovery yield.
  - Developed COAST end to end, including the statistical methodology, R software implementation, validation studies, and accompanying manuscript.
  - Co-developed machine-learning methods for predictive biomarker discovery and patient stratification from standard-of-care medical images, including candidate biomarkers for ADC response; named co-inventor on U.S. Patent 12,299,884 and co-author of the accompanying preprint. [5]
- Publications
  - Causal considerations can determine the utility of machine learning assisted GWAS. [4]
  - A scalable framework for identifying allelic series from summary statistics. [5]
  - EmbedGEM: a framework to evaluate the utility of embeddings for genetic discovery. [11]
  - Multi-artifact detection and filtering in digital pathology using intrinsic image properties. [3]
  - Machine learning enabled prediction of digital biomarkers from whole slide histopathology images. [5]
  - Pitfalls in performing genome-wide association studies on ratio traits. [7]
  - An allelic-series rare-variant association test for candidate-gene discovery. [13]

Google

09/2019 – 09/2021

Data Scientist

- Department: Genomic Medicine Team and DevIntel Data Science Team
- Team Leads: Babak Alipanahi, Ph.D. and Cory McLean, Ph.D. (Genomic Medicine); Heng Liu, Ph.D. (Core Developer)
- Projects
  - Conducted genetic analyses and contributed to manuscript development across projects using machine learning-derived phenotypes, including COPD, glaucoma, and physiological waveforms.
  - Co-developed DeepNull, a machine-learning method for modeling nonlinear covariate effects in GWAS and polygenic prediction, with central contributions to methodology and manuscript development.
  - Built reusable GWAS analysis tooling for locus formation, fine-mapping, replication analysis, and winner's-curse correction, supporting multiple projects within the Genomic Medicine team.

  - Developed internal R packages for causal analysis of matched pre/post designs and linear mixed-model analysis of repeated-measures data.

- Publications
  - Applying multimodal AI to physiological waveforms improves genetic prediction of cardiovascular traits. [6]
  - Unsupervised representation learning improves genomic discovery for lung function and respiratory disease prediction. [10]
  - High-throughput ML-guided design of diverse single-domain antibodies against SARS-CoV-2. [6]
  - Inference of chronic obstructive pulmonary disease with deep learning on raw spirograms identifies new genetic loci and improves risk models. [14]
  - DeepNull: Modeling non-linear covariate effects improves phenotype prediction and association power. [16]
  - Large-scale machine learning-based phenotyping significantly improves genomic discovery for optic nerve head morphology. [18]

Broad Institute 06/2019 – 09/2019

Visiting Scientist

- Department: Medical and Population Genetics
- Principal Investigator: Hilary Finucane, Ph.D.
- Project: Developed a cross-population extension of SuSiE that served as a precursor to the published MultiSuSiE method. [3]

Technical Experience

- **AI & ML**: Computer vision, computational pathology, clinical ML, deep learning, transformers, representation learning, self-supervised learning, time-series forecasting, and world models.
- **Statistical Genetics**: Genome-wide association studies, fine-mapping, polygenic scoring, rare-variant association testing, and meta-analysis.
- **Prediction-Based Inference**: Valid statistical inference using machine-learned or surrogate outcomes.
- **Statistical Methods for Complex Data**: Development of novel regression-based statistical tests for complex settings, including the presence of related subjects, multivariate and repeated-measures data, and missing data.
- **Clinical Trials & Event-Time Analysis**: Survival and recurrent-event analysis, competing risks, and clinical-trial methods and estimands.
- **Statistical Interpretability**: Development of methods targeting interpretable estimands while avoiding unnecessary assumptions.
- **Causal Inference**: Matched pre/post designs and longitudinal observational data.
- **Programming & ML Frameworks**: R, Python, C++/Rcpp, PyTorch, and TensorFlow.
- **Research Computing & Data Tools**: Git, SQL, BigQuery, HPC clusters, AWS, GCP, Docker, and AI-assisted development workflows.

Scientific Software

Developer of the statistical methods and author of the R software implementations below.

- **COAST / AllelicSeries**: Gene-level rare-variant association testing designed to identify allelic series in which increasingly deleterious coding variants have increasingly large phenotypic effects. [13, 5]
- **RNOmni**: Robust association testing for continuous traits with non-normal residuals using direct, indirect, and omnibus rank-based inverse normal transformation tests. [19]
- **SurrogateRegression**: Prediction-based inference that jointly models a partially observed target outcome and a correlated surrogate outcome to improve power while maintaining valid inference. [15, 12]
- **MGMM**: Missingness-aware Gaussian mixture modeling for parameter estimation, clustering, and imputation with incomplete multivariate data. [7]
- **Temporal**: Parametric time-to-event analysis and between-group comparisons of interpretable survival summaries, including mean, median, and restricted mean survival.
- **MCC**: Estimation and inference for mean cumulative count curves and their areas for recurrent-event outcomes subject to censoring and terminal events. [2]

Patents

[1] Woicik A, Probert C, Akle Serrano S, **McCaw ZR**, Dulken B, and Narayanan S. "Systems and methods for determining tissue microarray sampling protocols." U.S. Patent 12,333,725 B2. Issued June 17, 2025.

[2] Kanwar V, Probert C, Dulken B, Woicik A, and **McCaw ZR**. "Systems and methods for artifact detection and removal from image data." U.S. Patent 12,299,856 B1. Issued May 13, 2025.

[3] Probert C, **McCaw ZR**, Koller D, and Shcherbina A. "Machine-learning-enabled predictive biomarker discovery and patient stratification using standard-of-care data." U.S. Patent 12,299,884 B2. Issued May 13, 2025.

Publications

Statistical Genetics

[1] A Frouin et al. "A versatile multi-components mixed model for bacterial-Genome Wide association studies". In: *Nature Communications* 17.1 (May 2026), p. 6171. DOI: 10.1038/s41467-026-72305-y.

[2] N Kui et al. "Large impact of genetic data processing steps on stability and reproducibility of set-based analyses in genome-wide association studies". In: *Genetics* 234.1 (Mar. 2026), iyag079. DOI: 10.1093/genetics/iyag079.

[3] J Rossen et al. "MultiSuSiE improves multi-ancestry fine-mapping in All of Us whole-genome sequencing data". In: *Nature Genetics* (Jan. 2026). DOI: 10.1038/s41588-025-02450-5.

[4] S Mukherjee et al. "Causal considerations can determine the utility of machine learning assisted GWAS". In: *Proceedings of the sixth Conference on Health, Inference, and Learning*. Vol. 287. Proceedings of Machine Learning Research. 2025, pp. 179–193. URL: https://proceedings.mlr.press/v287/mukherjee25a.html.

[5] ZR McCaw et al. "A Scalable Framework for Identifying Allelic Series from Summary Statistics". In: *American Journal of Human Genetics* 112.11 (Nov. 2025), pp. 2772–2788. DOI: 10.1016/j.ajhg.2025.09.012.

[6] Y Zhou et al. “Applying multimodal AI to physiological waveforms improves genetic prediction of cardiovascular traits”. In: *American Journal of Human Genetics* (July 2025). DOI: 10.1016/j.ajhg.2025.05.015.

[7] ZR McCaw et al. “Pitfalls in performing genome-wide association studies on ratio traits”. In: *Human Genetics and Genomics Advances* (Apr. 2025). DOI: 10.1016/j.xhgg.2025.100406.

[8] X Li, H Chen, MS Selvaraj, et al. “A statistical framework for powerful multi-trait rare variant analysis in large-scale whole-genome sequencing studies”. In: *Nature Computational Science* (Feb. 2025). DOI: 10.1038/s43588-024-00764-8.

[9] R Sun, ZR McCaw, and X Lin. “Testing a Large Number of Composite Null Hypotheses Using Conditionally Symmetric Multidimensional Gaussian Mixtures in Genome-Wide Studies”. In: *Journal of the American Statistical Association* (Oct. 2024). DOI: 10.1080/01621459.2024.2422124.

[10] T Yun et al. “Unsupervised representation learning improves genomic discovery for lung function and respiratory disease prediction”. In: *Nature Genetics* (July 2024). DOI: 10.1038/s41588-024-01831-6.

[11] S Mukherjee et al. “EmbedGEM: A framework to evaluate the utility of embeddings for genetic discovery”. In: *Bioinformatics Advances* 4.1 (July 2024), vbae135. DOI: 10.1093/bioadv/vbae135.

[12] ZR McCaw et al. “Synthetic surrogates improve power for genome-wide association studies of partially missing phenotypes in population biobanks”. In: *Nature Genetics* (June 2024). DOI: 10.1038/s41588-024-01793-9.

[13] ZR McCaw et al. “An allelic-series rare-variant association test for candidate-gene discovery”. In: *American Journal of Human Genetics* 110.8 (July 2023), pp. 1330–1342. DOI: 10.1016/j.ajhg.2023.07.001.

[14] J Cosentino et al. “Inference of chronic obstructive pulmonary disease with deep learning on raw spirograms identifies new genetic loci and improves risk models”. In: *Nature Genetics* (Apr. 2023). DOI: 10.1038/s41588-023-01372-4.

[15] ZR McCaw et al. “Leveraging a surrogate outcome to improve inference on a partially missing target outcome”. In: *Biometrics* (Feb. 2022). DOI: 10.1111/biom.13629.

[16] ZR McCaw et al. “DeepNull: Modeling non-linear covariate effects improves phenotype prediction and association power”. In: *Nature Communications* 13.1 (Jan. 2022), p. 241. DOI: 10.1038/s41467-021-27930-0.

[17] H Julienne et al. “Multitrait GWAS to connect disease variants and biological mechanisms”. In: *PLoS Genetics* 17.8 (Aug. 2021), e1009713. DOI: 10.1371/journal.pgen.1009713.

[18] B Alipanahi et al. “Large-scale machine learning-based phenotyping significantly improves genomic discovery for optic nerve head morphology”. In: *American Journal of Human Genetics* (May 2021). DOI: 10.1016/j.ajhg.2021.05.004.

[19] ZR McCaw et al. “Operating Characteristics of the Rank-Based Inverse Normal Transformation for Quantitative Trait Analysis in Genome-Wide Association Studies”. In: *Biometrics* (Dec. 2019). DOI: 10.1111/biom.13214.

Applied Machine Learning

[1] J Gao et al. “What Is Fair? Defining Fairness in Machine Learning for Health”. In: *Stat Med* 44.20-22 (Sept. 2025), e70234. DOI: 10.1002/sim.70234.

[2] J Gronsbell et al. “Another look at inference after prediction”. In: *arXiv* (Nov. 2024). DOI: 10.48550/arXiv.2411.19908.

[3] V Kanwar et al. “Multi-Artifact Detection and Filtering in Digital Pathology Using Intrinsic Image Properties”. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*. Aug. 2024, pp. 1–5. DOI: 10.1109/ISBI56570.2024.10635902.

[4] A Woicik et al. “In Silico Optimization of Tissue Microarray Design for Machine Learning Analysis”. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*. Aug. 2024, pp. 1–5. DOI: 10.1109/ISBI56570.2024.10635208.

[5] ZR McCaw et al. “Machine learning enabled prediction of digital biomarkers from whole slide histopathology images”. In: *medRxiv* (Jan. 2024). DOI: 10.1101/2024.01.06.24300926.

[6] C Angermueller, Z Mariet, B Jester, et al. “High-throughput ML-guided design of diverse single-domain antibodies against SARS-CoV-2”. In: *bioRxiv* (Dec. 2023). DOI: 10.1101/2023.12.01.569227.

[7] ZR McCaw, H Julienne, and H Aschard. “Fitting Gaussian mixture models on incomplete data”. In: *BMC Bioinformatics* 23.1 (June 2022), p. 208. DOI: 10.1186/s12859-022-04740-9.

Biostatistics & Clinical Trials

[1] ZR McCaw et al. “Using Days Alive and Out of Hospital as the Study Endpoint in Cardiovascular Heart Failure Clinical Trials”. In: *JACC: Heart Failure* 14.7 (July 2026), p. 103013. DOI: 10.1016/j.jchf.2026.103013.

[2] ZR McCaw et al. “Pseudo-value Based Mean Cumulative Count Regression”. In: *arXiv* (June 2026). DOI: 10.48550/arXiv.2606.24024.

[3] J Gronsbell et al. “Nonparametric estimation of the total treatment effect with multiple outcomes in the presence of terminal events”. In: *Biometrics* 82.2 (Apr. 2026), ujag053. DOI: 10.1093/biomtc/ujag053.

[4] AD Sherry et al. “Interpreting Treatment Effects Using Posterior Probabilities: A Bayesian Reanalysis of 230 Phase III Oncology Trials”. In: *JCO Clinical Cancer Informatics* 10.2 (Apr. 2026), e2500303. DOI: 10.1200/CCI-25-00303.

[5] A Ocampo et al. “Revealing the Truth: Calculating True Values in Causal Inference Simulation Studies via Gaussian Quadrature”. In: *arXiv* (Jan. 2026). DOI: 10.48550/arXiv.2601.05128.

[6] EJ Hsu et al. “Differential censoring in overall survival in phase 3 oncology clinical trials”. In: *Journal of the National Cancer Institute* djaf378 (Dec. 2025). DOI: 10.1093/jnci/djaf378.

[7] S Armbruster, ZR McCaw, K Jering, et al. “Capturing Totality of Treatment Effects: A Model-Free Approach to Analyzing Multiple Event-Time Data From Heart Failure Studies”. In: *JACC Heart Failure* (Aug. 2025). DOI: 10.1016/j.jchf.2025.102605.

[8] AD Sherry, P Msaouel, AM Miller, et al. “Reproducibility of statistically significant phase III oncology trials: An In Silico meta-epidemiological analysis”. In: *European Journal of Cancer* (Aug. 2025). DOI: 10.1016/j.ejca.2025.115596.

[9] AD Sherry et al. “Bayesian Interim Analysis and Efficiency of Phase III Randomized Trials”. In: *Br J Cancer* (Aug. 2025). DOI: 10.1038/s41416-025-03156-5.

[10] AD Sherry, Y Liu, P Msaouel, et al. “Survival-inferred fragility of statistical significance in phase III oncology trials”. In: *NPJ Precision Oncology* (July 2025). DOI: 10.1038/s41698-025-01024-2.

[11] TJ Kleber, AD Sherry, AJ Arifin, et al. “Justification, margin values, and analysis populations for oncologic noninferiority and equivalence trials: a meta-epidemiological study”. In: *Journal of the National Cancer Institute* (May 2025). DOI: 10.1093/jnci/djae318.

[12] J Gronsbell et al. “Exact Inference for Random Effects Meta-Analyses for Small, Sparse Data”. In: *Stats* 8.1 (Mar. 2025). DOI: 10.3390/stats8010005.

[13] AD Sherry et al. “Evidence-Based Prior for Estimating the Treatment Effect of Phase III Randomized Trials in Oncology”. In: *JCO Precision Oncology* 8 (Oct. 2024), e2400363. DOI: 10.1200/PO.24.00363.

[14] AD Sherry et al. “Increasing Power in Phase III Oncology Trials With Multivariable Regression: An Empirical Assessment of 535 Primary End Point Analyses”. In: *JCO Clinical Cancer Informatics* 8 (Sept. 2024), e2400102. DOI: 10.1200/CCI.24.00102.

[15] AD Sherry et al. “Off-Protocol Radiation Therapy in Phase 3 Metastatic Solid Tumor Trials”. In: *Int J Radiat Oncol Biol Phys* (Aug. 2024). DOI: 10.1016/j.ijrobp.2024.05.033.

[16] AD Sherry et al. “Improving the clinical meaning of surrogate endpoints: An empirical assessment of clinical progression in phase III oncology trials”. In: *International Journal of Cancer* (Aug. 2024). DOI: 10.1002/ijc.35129.

[17] TA Lin et al. “Proportional Hazards Violations in Phase 3 Cancer Clinical Trials: A Potential Source of Trial Misinterpretation”. In: *Clinical Cancer Research* (Aug. 2024). DOI: 10.1158/1078-0432.CCR-24-0566.

[18] EJ Hsu et al. “Association of differential censoring with survival and suboptimal control arms among oncology clinical trials”. In: *Journal of the National Cancer Institute* 116.6 (June 2024), pp. 990–994. DOI: 10.1093/jnci/djae028.

[19] AD Sherry, P Msaouel, TA Lin, et al. “Postprogression therapy and confounding for the estimated treatment effect on overall survival in phase III oncology trials”. In: *BMJ Oncology* (Apr. 2024). DOI: 10.1136/bmjonc-2024-000322.

[20] AD Sherry, AW Hahn, ZR McCaw, et al. “Differential Treatment Effects of Subgroup Analyses in Phase 3 Oncology Trials From 2004 to 2020”. In: *JAMA Network Open* 7.3 (Mar. 2024), e243379. DOI: 10.1001/jamanetworkopen.2024.3379.

[21] AD Sherry, P Msaouel, ZR McCaw, et al. “Prevalence and implications of significance testing for baseline covariate imbalance in randomised cancer clinical trials: The Table 1 Fallacy”. In: *European Journal of Cancer* 194 (Nov. 2023), p. 113357. DOI: 10.1016/j.ejca.2023.113357.

[22] X Wang et al. “Using a Clinically Interpretable End Point Composed of Multiple Outcomes to Evaluate Totality of Treatment Effect in Comparative Oncology Studies”. In: *JAMA Network Open* 6.6 (June 2023), e2319055. DOI: 10.1001/jamanetworkopen.2023.19055.

[23] PS Jhund et al. “Effect of Dapagliflozin on Total Heart Failure Events in Patients With Heart Failure With Mildly Reduced or Preserved Ejection Fraction: A Prespecified Analysis of the DELIVER Trial”. In: *JAMA Cardiology* (Apr. 2023). DOI: 10.1001/jamacardio.2023.0711.

[24] A Das et al. “Assessment of Median and Mean Survival Time in Cancer Clinical Trials”. In: *JAMA Network Open* 6.4 (Apr. 2023), e236498. DOI: 10.1001/jamanetworkopen.2023.6498.

[25] HM Dehbi, A Embleton-Thirsk, and ZR McCaw. “Sample size calculation for randomized selection trials with a time-to-event endpoint and a margin of practical equivalence”. In: *Statistics in Medicine* (June 2022). DOI: 10.1002/sim.9490.

[26] BL Claggett et al. “Quantifying Treatment Effects in Trials with Multiple Event-Time Outcomes”. In: *NEJM Evidence* 1.10 (June 2022). DOI: 10.1056/EVIDoa2200047.

[27] ZR McCaw, DH Kim, and LJ Wei. “Pitfall in the Design and Analysis of Comparative Oncology Trials With a Time-to-Event Endpoint and Recommendations”. In: *JNCI Cancer Spectrum* 6.1 (Feb. 2022), pkac007. DOI: 10.1093/jncics/pkac007.

[28] ZR McCaw et al. “Practical Recommendations on Quantifying and Interpreting Treatment Effects in the Presence of Terminal Competing Risks: A Review”. In: *JAMA Cardiology* (Dec. 2021). DOI: 10.1001/jamacardio.2021.4932.

[29] ZR McCaw et al. “Choosing clinically interpretable summary measures and robust analytic procedures for quantifying the treatment difference in comparative clinical studies”. In: *Statistics in Medicine* 40.28 (Dec. 2021), pp. 6235–6242. DOI: 10.1002/sim.8971.

[30] R Sun et al. “Moving beyond conventional stratified analysis to assess the treatment effect in a comparative oncology study”. In: *Journal for ImmunoTherapy of Cancer* 9.11 (Nov. 2021), e003323. DOI: 10.1136/jitc-2021-003323.

[31] ZR McCaw et al. “Neoadjuvant chemotherapy in bladder cancer: Clinical benefit observed in prospective trials computed with restricted mean survival times”. In: *Urologic Oncology* S1078-1439.20 (Jan. 2021), pp. 30640–30642. DOI: 10.1016/j.urolonc.2020.12.012.

[32] ZR McCaw et al. “Survival analysis of treatment efficacy in comparative COVID-19 studies”. In: *Clinical Infectious Diseases* (Oct. 2020). DOI: 10.1093/cid/ciaa1563.

[33] C Perego et al. “Utility of Restricted Mean Survival Time Analysis for Heart Failure Clinical Trial Evaluation and Interpretation”. In: *JACC Heart Failure* (Oct. 2020). DOI: 10.1016/j.jchf.2020.07.005.

[34] ZR McCaw et al. “Selecting Appropriate Endpoints for Assessing Treatment Effects in Comparative Clinical Studies for COVID-19”. In: *Contemporary Clinical Trials* (Sept. 2020). DOI: 10.1016/j.cct.2020.106145.

[35] ZR McCaw et al. “How to Quantify and Interpret Treatment Effects in Comparative Clinical Studies of COVID-19”. In: *Annals of Internal Medicine* (July 2020). DOI: 10.7326/M20-4044.

[36] B Huang et al. “Analysis of Response Data for Assessing Treatment Effects in Comparative Clinical Studies”. In: *Annals of Internal Medicine* (July 2020). DOI: 10.7326/M20-0104.

[37] ZR McCaw, G Yin, and LJ Wei. “Using the Restricted Mean Survival Time Difference as an Alternative to the Hazard Ratio for Analyzing Clinical Cardiovascular Studies”. In: *Circulation* 140.17 (Oct. 2019), pp. 1366–1368. DOI: 10.1161/CIRCULATIONAHA.119.040680.

[38] ZR McCaw et al. “Applying Evidence-Based Medicine to Shared Decision Making: Value of Restricted Mean Survival Time”. In: *The American Journal of Medicine* 132.1 (Jan. 2019), pp. 13–15. DOI: 10.1016/j.amjmed.2018.07.026.

Biology

[1] L Siraj et al. “Functional dissection of complex trait variants at single-nucleotide resolution”. In: *Nature* (Feb. 2026). DOI: 10.1038/s41586-026-10121-6.

[2] J Marzec et al. “Toll-like receptor 4-mediated respiratory syncytial virus disease and lung transcriptomics in differentially susceptible inbred mouse strains”. In: *Physiological Genomics* (Nov. 2019). DOI: 10.1152/physiolgenomics.00101.2019.

[3] M High et al. “Determinants of host susceptibility to murine respiratory syncytial virus (RSV) disease identify a role for the innate immunity scavenger receptor MARCO gene in human infants”. In: *EBioMedicine* S2352-3964.16 (2016), pp. 30360–7. DOI: 10.1016/j.ebiom.2016.08.011.

[4] JM Ciencewicki et al. “Effects of mannose-binding lectin on pulmonary gene expression and innate immune inflammatory response to ozone”. In: *American Journal of Physiology-Lung Cellular and Molecular Physiology* 311.2 (2016), pp. 280–91. DOI: 10.1152/ajplung.00205.2015.

[5] BP Kleinstiver et al. “Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells”. In: *Nature Biotechnology* 34.8 (2016), pp. 869–74. DOI: 10.1038/nbt.3620.

[6] KC Verhein et al. “Novel Roles for Notch3 and Notch4 Receptors in Gene Expression and Susceptibility to Ozone-Induced Lung Inflammation in Mice”. In: *Environmental Health Perspectives* 123.8 (2015), pp. 799–805. DOI: 10.1289/ehp.1408852.

[7] J Krishnaswamy et al. “Coincidental loss of DOCK8 function in NLRP10-deficient and C3H/HeJ mice results in defective dendritic cell migration”. In: *PNAS* 112.10 (2015), pp. 3056–61. DOI: 10.1073/pnas.1501554112.

[8] H Cho et al. “Association of Nrf2 polymorphism haplotypes with acute lung injury phenotypes in inbred strains of mice”. In: *Antioxidants and Redox Signaling* 22.4 (2015), pp. 325–38. DOI: 10.1089/ars.2014.5942.

[9] KC Verhein et al. “Genetic Factors Involved in Susceptibility to Lung Disease”. In: *The Lung Second Edition: Development, Aging and the Environment*. Ed. by Plopper CG Harding R Pinkerton KE. London: Academic Press, 2014. DOI: 10.1016/B978-0-12-799941-8.00020-1.

Correspondence

[1] ZR McCaw, TW Soare, and C O’Dushlaine. “Ratio Variables in Regression: Implications for Cardiovascular Research”. In: *Journal of Clinical Cardiology* 7.1 (Apr. 2026), pp. 58–61. DOI: 10.33696/cardiology.7.090.

[2] AD Sherry et al. “Interpretation of long-term survival data from MONARCH 3”. In: *Annals of Oncology* S0923-7534.24 (June 2024), pp. 00718–X. DOI: 10.1016/j.annonc.2024.05.543.

[3] ZR McCaw, L Tian, and LJ Wei. “Evaluating the Duration of Response With Mirvetuximab Soravtansine for Treating Platinum-Resistant Ovarian Cancer”. In: *Journal of Clinical Oncology* (Aug. 2023), JCO2300288. DOI: 10.1200/JCO.23.00288.

[4] ZR McCaw, PG Richardson, and LJ Wei. “Assessing the Ability of Long Noncoding RNA Expression to Predict Patient Outcomes in Pediatric AML”. In: *Journal of Clinical Oncology* (June 2023), JCO2300465. DOI: 10.1200/JCO.23.00465.

[5] TA Lin et al. “Determining the Efficacy of Pembrolizumab in Patients With Previously Treated Advanced Hepatocellular Carcinoma”. In: *Journal of Clinical Oncology* (Apr. 2023). DOI: 10.1200/JCO.23.00072.

[6] ZR McCaw, EB Ludmir, and LJ Wei. “Assessing the Clinical Utility of Oral Paclitaxel Plus Encequidar Versus Intravenous Paclitaxel in Patients With Metastatic Breast Cancer”. In: *Journal of Clinical Oncology* 41.6 (Feb. 2023), p. 1323. DOI: 10.1200/JCO.22.01759.

[7] ZR McCaw and LJ Wei. “Clinical Utility Assessment of Gonadotropin-Releasing Hormone Analogs Among Women Younger Than 35 Years”. In: *JAMA Oncology* 8.6 (Apr. 2022), pp. 943–944. DOI: 10.1001/jamaoncol.2022.0488.

[8] ZR McCaw and LJ Wei. “Questions About a Risk Prediction Model of Mortality After Esophagectomy for Cancer”. In: *JAMA Surgery* 157.3 (Nov. 2021), pp. 279–280. DOI: 10.1001/jamasurg.2021.5701.

[9] ZR McCaw, L Tian, and LJ Wei. “Quantifying the Effect of Lower vs Higher Positive End-Expiratory Pressure on Ventilator-Free Survival in ICU Patients”. In: *JAMA* 325.15 (Apr. 2021), pp. 1566–1567. DOI: 10.1001/jama.2021.1700.

[10] ZR McCaw, MA Liu, and LJ Wei. “Olaparib in Metastatic Castration-Resistant Prostate Cancer”. In: *New England Journal of Medicine* 384.12 (Mar. 2021), p. 1174. DOI: 10.1056/NEJMc2100225.

[11] ZR McCaw, EB Ludmir, and LJ Wei. “Quantifying the Long-term Survival Benefit of Pembrolizumab for Patients With Advanced Gastric Cancer”. In: *JAMA Oncology* 7.4 (Feb. 2021). DOI: 10.1001/jamaoncol.2020.8002.

[12] ZR McCaw, G Fitzmaurice, and LJ Wei. “The COMPASS Trial: Net Clinical Benefit of Low-Dose Rivaroxaban Plus Aspirin as Compared With Aspirin in Patients With Chronic Vascular Disease”. In: *Circulation* 143.1 (Jan. 2021), e1–e2. DOI: 10.1161/CIRCULATIONAHA.120.050723.

[13] RR Patel et al. “Transparency in Reporting of Phase 3 Cancer Clinical Trial Results”. In: *Acta Oncologica* 60.2 (Dec. 2020). DOI: 10.1080/0284186X.2020.1856410.

[14] EB Ludmir, ZR McCaw, and LJ Wei. “Interpreting the Effect of Ipilimumab Following Radiotherapy for Patients with Postdocetaxel Metastatic Castration-resistant Prostate Cancer”. In: *European Urology* 79.1 (Oct. 2020), e10–e11. DOI: 10.1016/j.eururo.2020.09.049.

[15] ZR McCaw, L Tian, and LJ Wei. "What We Learned from Recent COVID-19 Clinical Studies Regarding Statistical Methodology". In: *Biopharmaceutical Report* 27.3 (Oct. 2020).

[16] ZR McCaw, L Tian, and LJ Wei. "Appropriate Analysis of Duration of Response Data in Cancer Trials". In: *JAMA Oncology* 6.12 (Oct. 2020), p. 1978. DOI: 10.1001/jamaoncol.2020.4657.

[17] EB Ludmir et al. "Progression-free survival in the ICON8 trial". In: *Lancet* 396.10253 (Sept. 2020), p. 756. DOI: 10.1016/S0140-6736(20)31175-2.

[18] ZR McCaw et al. "Further clinical interpretation and implications of KEYNOTE-048 findings". In: *Lancet* 396.10248 (Aug. 2020), pp. 378–379. DOI: 10.1016/S0140-6736(20)30904-1.

[19] ZR McCaw, DH Kim, and LJ Wei. "Remdesivir for the Treatment of Covid-19 - Preliminary Report". In: *New England Journal of Medicine* 383 (July 2020), pp. 992–994. DOI: 10.1056/NEJMc2022236.

[20] ZR McCaw, DH Kim, and LJ Wei. "Risk-Benefit Comparisons Between Shorter and Longer Durations of Adjuvant Chemotherapy in High-Risk Stage II Colorectal Cancer". In: *JAMA Oncology* 6.8 (June 2020), pp. 1301–1302. DOI: 10.1001/jamaoncol.2020.2256.

[21] EB Ludmir et al. "Fulvestrant plus capivasertib for metastatic breast cancer". In: *Lancet Oncology* 21.5 (May 2020), e233. DOI: 10.1016/S1470-2045(20)30228-X.

[22] ZR McCaw, LJ Wei, and EB Ludmir. "Interpreting the impact of apalutamide on overall survival among patients with non-metastatic castration-resistant prostate cancer". In: *Annals of Oncology* 31.3 (Mar. 2020), pp. 438–440. DOI: 10.1016/j.annonc.2019.11.020.

[23] D Li, ZR McCaw, and LJ Wei. "Interpreting the Benefit of Simvastatin-Ezetimibe in Patients 75 Years or Older". In: *JAMA Cardiology* 5.2 (Jan. 2020), p. 235. DOI: 10.1001/jamacardio.2019.5200.

[24] EB Ludmir et al. "Quantifying the Benefit of Non-small-cell lung Cancer Immunotherapy". In: *Lancet* 394.10212 (Nov. 2019), p. 1904. DOI: 10.1016/S0140-6736(19)32503-6.

[25] ZR McCaw and LJ Wei. "P2Y12 Inhibitor Monotherapy vs Dual Antiplatelet Therapy After Percutaneous Coronary Intervention". In: *JAMA* 322.16 (Oct. 2019), p. 1607. DOI: 10.1001/jama.2019.13159.

[26] ZR McCaw, Z Meng, and LJ Wei. "A Shorter Regimen for Rifampin-Resistant Tuberculosis". In: *New England Journal of Medicine* 381.11 (Sept. 2019), e22. DOI: 10.1056/NEJMc1905782.

[27] G Yin and ZR McCaw. "Design of Noninferiority Trials for Hypofractionated vs Conventional Radiotherapy Among Patients With Cancer". In: *JAMA Oncology* 5.10 (Aug. 2019). DOI: 10.1001/jamaoncol.2019.2391.

[28] ZR McCaw, DH Kim, and LJ Wei. "Analysis of Long-term Benefits of Intensive Blood Pressure Control". In: *JAMA* 322.2 (July 2019), pp. 169–170. DOI: 10.1001/jama.2019.5840.

[29] Z Yang, ZR McCaw, and G Yin. "Caplacizumab for Acquired Thrombotic Thrombocytopenic Purpura". In: *New England Journal of Medicine* 380.18 (May 2019), e32. DOI: 10.1056/NEJMc1902336.

[30] ZR McCaw, LJ Wei, and DH Kim. "Effects of Aspirin in the Healthy Elderly". In: *New England Journal of Medicine* 380.18 (May 2019), pp. 1775–1776. DOI: 10.1056/NEJMc1901774.

[31] Z Yang, ZR McCaw, and G Yin. "Radical Surgery or Watchful Waiting in Prostate Cancer". In: *New England Journal of Medicine* 380.11 (Mar. 2019), pp. 1083–1084. DOI: 10.1056/NEJMc1900410.

[32] ZR McCaw and LJ Wei. "Interpreting the Survival Benefit From Neoadjuvant Chemoradiotherapy Before Surgery for Locally Advanced Squamous Cell Carcinoma of the Esophagus". In: *Journal of Clinical Oncology* 37.12 (Mar. 2019), pp. 1032–1033. DOI: 10.1200/JCO.18.01164.

[33] ZR McCaw, LJ Wei, and DH Kim. "Interpreting the Prognostic Value of Unrecognized Myocardial Infarction Among Older Adults". In: *JAMA Cardiology* 4.4 (Mar. 2019), p. 391. DOI: 10.1001/jamacardio.2019.0184.

[34] ZR McCaw, JL Vassy, and LJ Wei. "Palbociclib and Fulvestrant in Breast Cancer". In: *New England Journal of Medicine* 380.8 (Feb. 2019), p. 796. DOI: 10.1056/NEJMc1816595.

[35] ZR McCaw, DH Kim, and LJ Wei. "Evaluating Treatment Effect of Transcatheter Interatrial Shunt Device Using Heart Failure Event Rates". In: *JAMA Cardiology* 4.3 (Feb. 2019), p. 299. DOI: 10.1001/jamacardio.2019.0001.

[36] ZR McCaw et al. "Trifluridine/tipiracil in metastatic gastric cancer". In: *Lancet Oncology* 20.1 (Jan. 2019), e8. DOI: 10.1016/S1470-2045(18)30908-2.

[37] ZR McCaw, F Jiang, and LJ Wei. "Trastuzumab Therapy for 9 Weeks vs 1 Year for Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer". In: *JAMA Oncology* 3.1 (Dec. 2018), pp. 117–118. DOI: 10.1001/jamaoncol.2018.5730.

[38] ZR McCaw et al. "Interpreting Clinical Benefits of Neoadjuvant Chemoradiation With Gemcitabine Versus Upfront Surgery in Patients With Borderline Resectable Pancreatic Cancer (BRPC)". In: *Annals of Surgery* 270.2 (Nov. 2018), e48–e50. DOI: 10.1097/SLA.0000000000003115.

[39] ZR McCaw, S Piantadosi, and LJ Wei. "Quantifying the Added Value of Low-Molecular-Weight Heparin to Intermittent Pneumatic Compression for Preventing Venous Thromboembolic Events Under the Risk-Benefit Perspective". In: *JAMA Surgery* 154.3 (Nov. 2018), pp. 270–271. DOI: 10.1001/jamasurg.2018.4294.

[40] ZR McCaw, D Liu, and LJ Wei. "Body Composition and Overall Survival in Patients With Nonmetastatic Breast Cancer". In: *JAMA Oncology* 5.1 (Nov. 2018), pp. 114–115. DOI: 10.1001/jamaoncol.2018.5290.

[41] ZR McCaw, LJ Wei, and DH Kim. "Gene Expression-Guided Adjuvant Chemotherapy in Breast Cancer". In: *New England Journal of Medicine* 379.17 (Oct. 2018), p. 1681. DOI: 10.1056/NEJMc1810515.

[42] BL Claggett et al. "Sex as a predictor of response to cancer immunotherapy". In: *Lancet Oncology* 19.8 (Aug. 2018), e377. DOI: 10.1016/S1470-2045(18)30517-5.

Professional Activities

- **ML4H Section Chair** 2025

- **Peer Review** 2025
  Journals: Bioinformatics, BMC Bioinformatics,
  BMC Medical Research Methodology, Cancers, Clinical Trials, EJMR,
  iScience, ISMB, Nature Aging, Nutrition,
  Nature Communications, npj Health Systems
- **Peer Review** 2024
  Journals: Bioinformatics, Briefings in Bioinformatics, BMC Medical Research Methodology, ISCB, ISMB, Statistics in Medicine, Stats
- **Peer Review** 2023
  Journals: Current Cancer Drug Targets, ISCB, RECOMB
- **Peer Review** 2022
  Journals: Axioms, ISCB, Life, Statistics in Biopharmaceutical Research,
  Statistics in Medicine, TEST, Viruses
- **Peer Review** 2021
  Journals: Circulation: Cardiovascular Quality and Outcomes,
  Frontiers in Genetics, ISCB, Statistics in Medicine
- **Peer Review** 2020
  Journals: ISCB, Statistics in Medicine

- **JSM Section Chair** 2019
  Regression Methods for Longitudinal Data
- **JSM Section Chair** 2018
  Gene-Gene and Gene-Environment Interactions

Conference Presentations

- **American Society of Human Genetics** 11/2024
  Unveiling the power of allelic series: enhancements and applications of COAST.
- **American Association for Cancer Research** 04/2024
  Machine learning enabled prediction of digital biomarkers from whole slide histopathology images.
- **American Society of Human Genetics** 10/2023
  Synthetic slope analysis for progression GWAS.
- **American Association for Cancer Research** 04/2023
  Learned phenotypic embeddings enable scalable imputation of high-content molecular data elucidating prognostic chromatin signatures.
- **American Society of Human Genetics** 10/2022
  An allelic series rare variant association test for candidate gene discovery.
- **American Society of Human Genetics** 10/2019
  Cross-population fine-mapping to identify shared and population-specific causal effects.
- **Joint Statistical Meetings** 07/2019
  Cross-tissue eQTL calling via surrogate expression analysis.
- **Harvard School of Public Health, Program in Quantitative Genomics** 11/2018
  Leveraging the UKB to empower association testing on scarce phenotypes.
- **Joint Statistical Meetings** 07/2018
  Leveraging surrogate phenotypes to improve inference on a partially missing target phenotype.
- **Joint Statistical Meetings** 07/2017
  Inverse normal transformation for genome-wide association testing of quantitative traits.
- **American Thoracic Society** 05/2014
  Gene expression profiling predicts response to respiratory syncytial virus (RSV) in mice.
- **National Institute of Environmental Health Sciences** 07/2011
  Identifying candidate susceptibility genes for respiratory syncytial virus (RSV) disease severity.
- **National Institute of Environmental Health Sciences** 07/2010
  Characterization of transcriptional networks underlying Tlr4-mediated respiratory syncytial virus (RSV) disease in mice.

Awards and Distinctions

- **Distinguished Student Paper Award** 07/2019
  Joint Statistical Meetings, Section on Statistics in Genomics and Genetics.
- **Stellar Abstract Award** 11/2018
  Harvard School of Public Health, Program in Quantitative Genomics

- **Ruth L. Kirschstein National Research Service Award (F31)** 03/2018
  Innovations in Genome-Wide Association Testing Inspired by Obstructive Sleep Apnea Phenotypes
- **Teaching Fellow** 11/2017
  Global Initiative for Neuropsychiatric Genetic Education in Research
- **NIH Pre-Doctoral Training Grant** 08/2016
  Statistical and Quantitative Training in Big Data Health Science
- **NIH Pre-Doctoral Training Grant** 08/2014
  Interdisciplinary Training Grant in Biostatistics and Computational Biology
- **NIH Post-Baccalaureate Research Fellow** 09/2013
  National Institute of Environmental Health Sciences
- **Undergraduate Academic Achievement Award** 04/2013
  UNC Department of Biostatistics
- **Phi Beta Kappa National Honors Society** 11/2011
- **NIH Summer Internship** 05/2011
  National Institute of Environmental Health Sciences 05/2010

Teaching Experience

UNC-Chapel Hill 06/2023 – Present

Adjunct Instructor

- Department: MPH@UNC
- Classes: Section instructor for SPHG 711, Data Analysis for Public Health.

UNC-Chapel Hill

- Class: Data Analysis for Public Health (SPHG 711) 08/2024 – 12/2024
  08/2023 – 12/2023

Harvard University

- Class: Inference II (BST 241) 02/2019 – 05/2019
  Instructor: Rui Wang, Ph.D.

- Class: Introduction to Biostatistics 02/2019
  Instructor: Lori Chibnik, Ph.D.
  Location: University of KwaZulu-Natal, Durban, South Africa

- Class: Multivariate and Longitudinal Analysis (BST 245) 02/2018 – 05/2018
  Instructor: Sebastien Haneuse, Ph.D.

- Class: Inference I (BST 231) 02/2017 – 05/2017
  Instructor: Judith Lok, Ph.D.

- Class: Statistical Genetics (BST 227) 10/2016 – 12/2016
  Instructor: Martin Aryee, Ph.D.

- Class: Computational Biology (STAT 215) 02/2016 – 05/2016
  Instructor: X. Shirley Liu, Ph.D.

UNC-Chapel Hill

- Class: General Chemistry I (CHEM 101) 08/2012 – 12/2012
  Instructor: Jennifer Krumper, Ph.D.

- Class: Organic Chemistry II (CHEM 262) 08/2011 – 12/2011
  Instructor: Jennifer Krumper, Ph.D.