Daiwei (David) Zhang

Department of Biostatistics, Epidemiology and Informatics
Perelman School of Medicine, University of Pennsylvania
423 Guardian Drive, Philadelphia, PA 19104
daiwei.zhang@pennmedicine.upenn.edu

Research Interests

Methodology: machine learning, vision-language models, generative AI

Applications: spatial multi-omics, computational pathology, medical imaging

Education and Training

Postdoctoral Training in Biostatistics, University of Pennsylvania, Philadelphia, PA — 2021 – Present

Advisor: Mingyao Li

Ph.D. in Biostatistics and Scientific Computing, University of Michigan, Ann Arbor — 2016 – 2021

Thesis: Novel statistical learning methods for high-dimensional complex biomedical data analysis

Advisors: Jian Kang & Seunggeun Lee

M.S. in Biostatistics, University of Michigan, Ann Arbor — 2016 – 2019

Budapest Semester in Mathematics, Hungarian Academy of Sciences — 2014

B.S. in Mathematics (with Honors), Calvin College, Grand Rapids, MI — 2012 – 2016

Research Experience

Research Associate in Biostatistics — 2022 – Present
Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Postdoctoral Researcher in Biostatistics — 2021 – 2022
Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA

Research Assistant in Biostatistics — 2016 – 2021
University of Michigan, Ann Arbor, MI

Research Assistant in Mathematics — 2013 – 2016
Calvin College, Grand Rapids, MI

Teaching Experience

Teaching Assistant (Calculus, Differential Equations) 2013 – 2014
Department of Mathematics, Calvin College, Grand Rapids, MI

Teaching Assistant (Introduction to Programming, Information Technology) 2012 – 2013
Department of Computer Science, Calvin College, Grand Rapids, MI

Awards

Penn Conference on Big Data in Biomedical and Population Health Sciences Best Poster Award 2022

International Biometric Society ENAR Distinguished Student Paper Award 2021

Michigan Student Symposium for Interdisciplinary Statistical Sciences Best Oral Presentation 2021

Rackham Conference Travel Grant 2018, 2020

William Rinck Memorial Prize (for the best graduating mathematics major) 2016

W. L. Putnam Mathematical Competition (ranked top 9% internationally) 2013

William Rinck Memorial Scholarship (for outstanding mathematics majors) 2013 – 2016

Trustee Scholarship (full tuition coverage) 2012 – 2016

Professional Service

Journal Referee

Journal of Machine Learning Research

Journal of the American Statistical Association

Biometrics

Statistics in Medicine

Annals of Applied Statistics

Biostatistics

Publications

*Equal contribution or alphabetical authorship

Journal Articles

1. **D. Zhang**, A. Schroeder, H. Yan, H. Yang, J. Hu, M. Y. Y. Lee, K. S. Cho, K. Susztak, G. X. Xu, M. D. Feldman, E. B. Lee, E. E. Furth, L. Wang, and M. Li. Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. *Nature Biotechnology*, in press, 2023.

2. **D. Zhang**, L. Li, C. Sripada, and J. Kang. Image response regression via deep neural networks. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, page qkad073, 2023.

3. **D. Zhang**, T. Liu, and J. Kang. Density regression and uncertainty quantification with bayesian deep noise neural networks. *Stat*, 12(1):e604, 2023.

4. J. Hu, K. Coleman, **D. Zhang**, E. B. Lee, H. Kadara, L. Wang, and M. Li. Deciphering tumor ecosystems at super resolution from spatial transcriptomics with TESLA. *Cell Systems*, 14(5):404–417, 2023.

5. **D. Zhang** and J. Kang. Discussion of "LESA: Longitudinal elastic shape analysis of brain subcortical structures". *Journal of the American Statistical Association*, 118(541):22–24, 2023.

6. X. Yin, L. S. Chan, D. Bose, A. U. Jackson, P. VandeHaar, A. E. Locke, C. Fuchsberger, H. M. Stringham, R. Welch, K. Yu, L. Fernandes Silva, S. K. Service, **D. Zhang**, E. C. Hector, E. Young, L. Ganel, I. Das, H. Abel, M. R. Erdos, L. L. Bonnycastle, J. Kuusisto, N. O. Stitziel, I. M. Hall, G. R. Wagner, J. Kang, J. Morrison, C. F. Burant, F. S. Collins, S. Ripatti, A. Palotie, N. B. Freimer, K. L. Mohlke, L. J. Scott, X. Wen, E. B. Fauman, M. Laakso, and M. Boehnke. Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. *Nature Communications*, 13(1):1644, March 2022.

7. L. G. Fritsche, Y. Ma, **D. Zhang**, M. Salvatore, S. Lee, X. Zhou, and B. Mukherjee. On cross-ancestry cancer polygenic risk scores. *PLOS Genetics*, 17(9):e1009670, September 2021.

8. **D. Zhang**, R. Dey, and S. Lee. Fast and robust ancestry prediction using principal component analysis. *Bioinformatics*, 36(11):3439–3446, 2020.

9. S. Auyeung*, J. Ruiter*, and **D. Zhang***. An algebraic characterization of highly connected 2n-manifolds. *Rose-Hulman Undergraduate Mathematics Journal*, 17(2):5, 2016.

10. C. Blom*, T. DeVries*, A. Hayes*, and **D. Zhang***. Analytic extension and conformal mapping in the dual and the double planes. *Rose-Hulman Undergraduate Mathematics Journal*, 14(2):9, 2013.

Under Review

11. K. Coleman, **D. Zhang**, J. Hu, and M. Li. Integrating gene expression and histology images for spatial domain detection in spatial transcriptomics. *Nature Communications*, under review.

12. G. Pei, V. Branchi, J. Min, K. I. Rajapakshe, D. Sadeghian, **D. Zhang**, Y. Liu, K. S. Cho, Y. Chu, E. Dai, G. Han, M. Li, C. Yee, P. M. Grandgenett, M. A. Hollingsworth, P. A. Guerrero, L. Wang, and A. Maitra. Spatial mapping of transcriptomic and lineage diversity in metastatic pancreatic cancer. *Nature*, under review.

13. Y. Liu, Y. Zhang, **D. Zhang**, G. Han, J. Min, G. Pei, K. S. Cho, E. Dai, Y. Chu, J. Jiang, R. Wang, P. A. Guerrero, A. Sinjab, J. Fujimoto, L. M. S. Soto, A. Serrano, J. A. Ajani, C. Yee, P. Mazur, A. J. Lazar, K. Lu, H. Kadara, M. Li, C. Fridman, W. H. Fridman, A. A. Jazaeri, A. Maitra, and L. Wang. Conserved cellular neighborhoods of cancer-associated fibroblasts revealed by high-plex pan-cancer single-cell spatial imaging analysis. *Cancer Cell*, under review.

14. **D. Zhang***, J. Jiang*, K. S. Cho*, Y. Liu, J. Chen, R. L. Segura, X. Yan, G. Pei, L. M. S. Soto, Y. Chu, A. F. Sinjab, C. Yee, S. Kopetz, A. Maitra, A. A. Jazaeri, A. Futreal, H. Kadara, A. Lazar, J. Gao, M. Li, and L. Wang. iStarTLS: Advanced detection and phenotyping of tertiary lymphoid structures. *Cancer Cell*, under review.

15. **D. Zhang**, B. Gao, Q. Feng, A. Manichaikul, G. M. Peloso, R. P. Tracy, P. Durda, K. D. Taylor, Y. Liu, W. C. Johnson, S. Gabriel, N. Gupta, J. D. Smith, F. Aguet, K. G. Ardlie, T. W. Blackwell, R. E. Gerszten, S. S. Rich, J. I. Rotter, L. J. Scott, X. Zhou, and S. Lee. Proteome-wide association studies for blood lipids and comparison with transcriptome-wide association studies. *bioRxiv*, 10.1101/2023.08.17.553749, 2023.

16. E. C. Hector*, **D. Zhang***, X. Yin, M. Laakso, T. Xu, Y. Bai, M. Boehnke, J. Kang, and T. Yu. A dmGWAS approach to understanding genetic regulators of metabolic networks. Under review.

In Preparation

17. A. Schroeder, **D. Zhang**, M. Loth, C. Luo, S. Piya, H.-H. Thai, B. Zhang, and M. Li. Predicting super-resolution gene expression in large tissues with iStarXL. In preparation.

18. J. Hu, K. Coleman, **D. Zhang**, E. Peyster, A. Janowczyk, and M. Li. Heart quality prediction by integrating histology with proteomics. In preparation.

19. M. Li*, **D. Zhang***, Y. Guo, and J. Kang. Neural network-guided independent component analysis with application to neuroimaging. In preparation.

Patents

1. **D. Zhang** and M. Li. Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. US Patent, 2023.

Presentations

Invited

1. **D. Zhang**, A. Schroeder, H. Yan, H. Yang, J. Hu, M. Y. Y. Lee, K. S. Cho, K. Susztak, G. X. Xu, M. D. Feldman, E. B. Lee, E. E. Furth, L. Wang, and M. Li. *Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology*. Joint Statistical Meetings, Portland, OR, 2024.

2. **D. Zhang**, L. Li, C. Sripada, and J. Kang. *Image-on-Scalar Regression via Deep Neural Networks*. Distinguished Student Paper Award Presentation, Spring Meeting, ENAR, International Biometric Society, Virtual Event, 2021.

3. **D. Zhang**, R. Dey, and S. Lee. *Fast and robust ancestry prediction with FRAPOSA*. Tools and Technology Seminar, Department of Internal Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, 2020.

CONTRIBUTED

4. **D. Zhang**, A. Schroeder, H. Yan, H. Yang, J. Hu, M. Y. Y. Lee, K. S. Cho, K. Susztak, G. X. Xu, M. D. Feldman, E. B. Lee, E. E. Furth, L. Wang, and M. Li. *Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology*. Annual Meeting, American Society of Human Genetics, Washington, DC, 2023.

5. **D. Zhang**, L. Li, C. Sripada, and J. Kang. *Image-on-Scalar Regression via Deep Neural Networks*. Michigan Student Symposium for Interdisciplinary Statistical Sciences, Virtual Event, 2021.

6. **D. Zhang**, Y. Guo, and J. Kang. *Neural network-guided independent component analysis with application to neuroimaging*. Spring Meeting, ENAR, International Biometric Society, Nashville, TN, 2020.

7. **D. Zhang**, R. Dey, and S. Lee. *Fast and robust ancestry inference using principal component analysis*. Annual Meeting, American Society of Human Genetics, San Diego, CA, 2018.