

LINYING ZHANG

linyinz@wustl.edu

<https://CausAiLab.github.io> ◊ Twitter: @Z_Linying

EDUCATION

- Columbia University** 2018 - 2023
PhD in Biomedical Informatics
Advisors: George Hripcsak and David Blei
Thesis: Causal machine learning for reliable real-world evidence generation in healthcare.
- Harvard University** 2016 - 2018
MS in Computational Biology and Quantitative Genetics
Advisor: Giovanni Parmigiani
Thesis: Interactions between multiple myeloma cells and bone marrow stromal cells impact epigenetic profiles of multiple myeloma.
- Boston University** 2011 - 2014
BA with Honors (*Summa Cum Laude*) in Biochemistry and Molecular Biology
Advisor: Ulla Hansen
Thesis: Establishing hepatocellular carcinoma cell lines with inducible expression of degradable LSF to investigate LSF regulation in cell cycle.

APPOINTMENT

- Washington University in St. Louis** October 2023 - Present
Assistant Professor of Biostatistics
Institute for Informatics, Data Science, and Biostatistics (I2DB)

HORNORS & AWARDS

- Edward H. Shortliffe Doctoral Dissertation Award Nominee 2024
- Symposium on Artificial Intelligence in Learning Health Systems (SAIL) Travel Award 2024
- OHDSI Best Community Contribution Award in Methodological Research 2022
- Women in Machine Learning Travel Award 2019
- Senior Book Award. Boston University. 2014
- Dean's List. Boston University. 2011-2014

PUBLICATIONS

1. Hripcsak G, **Zhang L**, Li K, Suchard MA, Ryan PB, Schuemie MJ. Assessing Covariate Balance with Small Sample Sizes. *medRxiv* 2024.04.23.24306230, 2024. [Under review]
2. Pang C, Jiang X, Pavinkurve NP, Kalluri KS, Minto EL, Patterson J, **Zhang L**, Hripcsak G, Elhadad N, Natarajan K. CEHR-GPT: Generating Electronic Health Records with Chronological Patient Timelines. *arXiv:2402.04400*, 2024. [Under review]
3. Jeong H, Jabbour S, Yang Y, Thapta R, Mozannar H, Han WJ, Mehandru N, Wornow M, Lialin V, Liu X, Lozano A, Zhu J, Kocielnik RD, Harrigian K, Zhang H, Lee E, Vukadinovic M, Balagopalan A, Jeanselme V, Matton K, Demirel I, Fries J, Rashidi P, Beaulieu-Jones B, Xu XO, McDermott M,

- Naumann T, Agrawal M, Zitnik M, Ustun B, Choi E, Yeom K, Gursoy G, Ghassemi M, Pierson E, Chen G, Kanjilal S, Oberst M, **Zhang L**, Singh H, Hartvigsen T, Zhou H, Okolo CT. Recent Advances, Applications, and Open Challenges in Machine Learning for Health: Reflections from Research Roundtables at ML4H 2023 Symposium. *arXiv:2403.01628*, 2024.
4. Cai CX, Nishimura A, Bowring MG, Westlund E, Tran D, Ng JH, Nagy P, Cook M, McLeggon JA, DuVall SL, Matheny ME, Golozar A, Ostropolets A, Minty E, Desai P, Bu F, Toy B, Hribar M, Falconer T, **Zhang L**, Lawrence-Archer L, Boland MV, Goetz K, Hall N, Shoaibi A, Reys J, Sena AG, Blacketer C, Swerdel J, Jhaveri KD, Lee E, Gilbert Z, Zeger SL, Crews DC, Suchard MA, Hripcsak G, Ryan PB. Similar risk of kidney failure among patients with blinding diseases who receive ranibizumab, aflibercept, and bevacizumab: an OHDSI Network Study. *Ophthalmology Retina*, 2024.
 5. **Zhang L**, Richter LR, Kim T, Hripcsak G. Evaluating and improving performance and racial fairness of algorithms for GFR estimation. *IEEE International Conference on Artificial Intelligence × Medicine, Health, and Care (AIMHC)*, 2024.
 6. **Zhang L**, Richter LR, Wang Y, Ostropolets A, Elhadad N, Blei DM, Hripcsak G. Causal fairness assessment of treatment allocation with electronic health records. *Journal of Biomedical Informatics*, 2024.
 7. Song W, Liu L, Rice H, Sainlaire M, Min L, **Zhang L**, Thai T, Kang MJ, Li S, Tejada C, Lipsitz S, Samal L, Carroll D, Adkison L, Herlihy L, Ryan V, Bates D, Latham N, Dykes P. From traditional fall injury risk screening to a temporal machine learning-based approach: improving algorithm generalizability and clinical action. *Journal of the American Geriatrics Society*, 2024.
 8. Ostropolets A, Albogami Y, Conover M, Banda JM, Baumgartner WA Jr, Blacketer C, Desai P, DuVall SL, Fortin S, Gilbert JP, Golozar A, Ide J, Kanter AS, Kern DM, Kim C, Lai L.Y.H, Li C, Liu F, Lynch K.E, Minty E, Ins Neves M, Ng DQ, Obene T, Pera V, Pratt N, Rao G, Rappoport N, Reinecke I, Saroufim P, Shoaibi A, Simon K, Suchard MA, Swerdel JN, Voss EA, Weaver J, **Zhang L**, Hripcsak G, and Ryan PB. Reproducible Variability: Assessing investigator discordance across nine research teams attempting to reproduce the same observational study. *JAMIA*, 2023.
 9. **Zhang L**, Wang Y, Schuemie MJ, Blei DM, and Hripcsak G. Adjusting for indirectly measured confounding using large-scale propensity score. *Journal of Biomedical Informatics*, 2022.
 10. Richter LR, Albert BI, **Zhang L**, Ostropolets A, Zitsman JL, Fennoy I, Albers D, Hripcsak G. Data assimilation on mechanistic models of glucose metabolism predicts glycemic states in adolescents following bariatric surgery. *Frontiers in Physiology*, 2022.
 11. Song W, **Zhang L**, Liu L, Sainlaire M, Karvar M, Kang M, Pullman A, Lipsitz S, Massaro A, Patil N, Jasuja R, Dykes PC. Predicting hospitalization of COVID-19 positive patients using clinician-guided machine learning methods. *JAMIA*, 2022.
 12. Song W, Kang MJ, **Zhang L**, Jung W, Song J, Bates DW, Dykes PC. Predicting pressure injury using nursing assessment phenotypes and machine learning methods. *JAMIA*, 2021.
 13. Ostropolets A, **Zhang L**, and Hripcsak G. A scoping review of clinical decision support tools that generate new knowledge to support decision making in real time. *JAMIA*, 2020.
 14. Ostropolets A, Chen R, **Zhang L**, and Hripcsak G. Characterizing physicians information needs related to a gap in knowledge unmet by current evidence. *JAMIA Open*, 2020.
 15. **Zhang L**, Wang Y, Ostropolets A, Mulgrave JJ, Blei DM, and Hripcsak G. The medical deconfounder: Assessing treatment effects with electronic health records. *Machine Learning for Healthcare Conference (MLHC)*, 2019.

16. Gottesman O, Johansson F, Meier J, Dent J, Lee D, Srinivasan S, **Zhang L**, Ding Y, Wihl D, Peng X, Yao J, Lage I, Mosch C, Lehman L.H, Komorowski M, Faisal A, Celi L, Sontag D, and Doshi-Velez F. Evaluating reinforcement learning algorithms in observational health settings. *arXiv preprint*, 2018.

CONFERENCE ABSTRACTS (PEER-REVIEWED)

1. Fan R, Banks D, Song W, Wilcox A, **Zhang L**. Real-world Comparative Effectiveness of Pharmacological Treatments for Opioid Use Disorder. In *OHDSI Global Symposium*, 2024. [Submitted]
2. Lee H, Gupta S, Blacketer C, Cook M, Naka S, Fan R, Martin B, Aziz K, **Zhang L**, Nagy P. Comparative Analysis of OMOP CDM Profiles Across Institutions and Future Research Implications. In *OHDSI Global Symposium*, 2024. [Submitted]
3. **Zhang L**, Jiang X, Natarajan K, Hripcsak G. Causal Fairness for Decomposing Racial and Sex Disparities in Treatment Allocation Using Real-World Data. In *AMIA Annual Symposium*, 2024. [Presentation]
4. Fan R, Banks D, Song W, Wilcox A, **Zhang L**. Real-world Comparative Effectiveness of Pharmacological Treatments for Opioid Use Disorder. In *AMIA Annual Symposium*, 2024. [Poster]
5. Chen HY, **Zhang L**, Hripcsak G. Learning Latent Confounding Representation in High-dimensional Observational Studies with EHRs via Variational Autoencoder. In *AMIA Annual Symposium*, 2024. [Poster]
6. Song W, Kang MJ, Liu L, Sainlaire M, Lowenthal G, Baris VK, Cho S, Carroll D, Furlong D, Gilles-Fowler W, Goncalves L, Lipsitz S, Melanson B, Morrow L, Massaro J, Martel T, Wolski P, **Zhang L**, Dykes PC. Multi-state Modeling of Pressure Injury Staging Transition Trajectories. In *AMIA Annual Symposium*, 2024. [Presentation]
7. **Zhang L**, Jiang X, Natarajan K, Hripcsak G. Building Causally Explainable Fair Learning Health System. In *Symposium on Artificial Intelligence in Learning Health Systems (SAIL)*, 2024. [Poster]
8. **Zhang L**, Jiang X, Natarajan K, Hripcsak G. Explaining Treatment Disparities from a Causal Perspective with EHRs. In *AMIA Annual Symposium*, 2023. [Presentation]
9. Schuemie M, Suchard MA, Nishimura A, **Zhang L**, Hripcsak G. Evaluating Confounding Adjustment When Sample Size is Small. In *OHDSI Global Symposium*, 2023. [Poster]
10. Sena AG, Reys J, Kim C, Brewster J, Black A, **Zhang L**, Cook M, Phuc PH, Suchard MA. Save Our Sisyphus Challenge: Lessons learned from Strategus execution on the OHDSI Network. In *OHDSI Global Symposium*, 2023. [Presentation]
11. Song W, Liu L, Sainlaire M, Cho S, Furlong D, Gilles-Fowler W, Herlihy L, Kang M.J, Lipsitz S, Melanson B, Massaro J, Martel T, Wolski P, **Zhang L**, Dykes P. An EHR-based Comparative Analysis of the Distribution of Pressure Injury Anatomical Locations and Stages and Associated Disparities Across a Large Healthcare System. In *AMIA Annual Symposium*, 2023. [Presentation]
12. **Zhang L**, Richter LR, Wang Y, Ostropolets A, Elhadad N, Blei DM, and Hripcsak G. A Bayesian causal inference approach for assessing fairness in clinical decision-making. In *Algorithmic Fairness through the Lens of Causality and Privacy Workshop, NeurIPS*, 2022. [Poster]
13. **Zhang L**, Richter LR, Blei DM, Wang Y, Ostropolets A, Elhadad N, and Hripcsak G. Assessing racial fairness of dialysis allocation in end-stage renal disease. In *OHDSI Global Symposium*, 2022. [Presentation]

14. **Zhang L**, Richter LR, Hripcsak G. Assessing the impact of race on glomerular filtration rate (GFR) prediction. In *OHDSI Global Symposium*. 2021. [Presentation]
15. Song W, **Zhang L**, Sainlaire M, Karvar M, Kang M, Pullman A, Massaro A, Patil N, Jasuja R, Dykes PC. Predicting hospitalization of COVID-19 positive patients using machine learning methods. In *AMIA Annual Symposium*. 2021. [Presentation]
16. **Zhang L**, Wang Y, Ostropolets A, Chen R, Blei DM, and Hripcsak G. The Multi-Outcome Medical Deconfounder: Assessing Treatment Effects on Multiple Renal Measures. In *AMIA Annual Symposium*. 2020. [Poster]
17. **Zhang L**, Wang Y, Ostropolets A, Chen R, Blei DM, and Hripcsak G. The multi-outcome medical deconfounder: assessing treatment effects on multiple renal measures. In *OHDSI Global Symposium*. 2020. [Poster]
18. Chen R, Schuemie M, Suchard M, Ostropolets A, **Zhang L**, Ryan P, Hripcsak G. Evaluation of large-scale propensity score modeling and covariate balance on potential unmeasured confounding in observational research. In *AMIA Annual Symposium*. 2020. [Poster]
19. **Zhang L**, Wang Y, Ostropolets A, Mulgrave JJ, Blei DM, and Hripcsak G. The medical deconfounder: assessing treatment effects with electronic health records. In *Women in Machine Learning (WiML) Workshop*. Vancouver, Canada. 2019. [Poster]
20. **Zhang L**, Wang Y, Ostropolets A, Mulgrave JJ, Blei DM, and Hripcsak G. The medical deconfounder: assessing treatment effects with electronic health records. In *Machine Learning for Health Workshop*. Vancouver, Canada. 2019. [Poster]
21. Ostropolets A, **Zhang L**, Mulgrave JJ, Hripcsak G. Investigating female-male differences in risk factors for myocardial infarction using OHDSI tools. In *AMIA Annual Symposium*. 2019. [Poster]
22. Song W, **Zhang L**, E. Gill, J.Z. Liu, A. Wright. Personalized treatment for type 2 diabetes using weighted k-nearest neighbors. In *AMIA Annual Symposium*. 2019. [Poster]
23. Szalat R, Samur MK, Ott CJ, Lawlor M, Epstein C, Abraham BJ, Lin CY, **Zhang L**, Prabhala R, Farrell N, Wes K, Tai YT, Fulciniti M, Parmigiani G, Young RA, Anderson KC, and Munshi NC. Integrative oncogenomic analysis combining whole genome, transcriptome and epigenome identifies altered chromatin accessibility landscape in multiple myeloma. In *American Society of Hematology Annual Meeting*. 2018. [Presentation]
24. **Zhang L**, Samur MK, Szalat R, Epstein CB, Prabhala R, Fulciniti M, Munshi NC.*, Parmigiani G.*. Interactions between multiple myeloma cells and bone marrow stromal cells impact epigenetic profiles of multiple myeloma. In *Program in Quantitative Genomics Conference*. Boston, MA. 2017. [Poster]
25. **Zhang L**, Samur MK, Szalat R, Epstein CB, Prabhala R, Fulciniti M, Munshi NC.*, Parmigiani G.*. Interactions between multiple myeloma cells and bone marrow stromal cells impact epigenetic profiles of multiple myeloma. In *Dana-Farber/Harvard Cancer Center Celebration of Junior Investigators*. Boston, USA. 2017. [Poster]

INVITED TALKS

1. **Washington University in St. Louis**
CS531a AI for Health Guest Lecture
 Reliable real-world evidence for equitable health care

2024

2. **Washington University in St. Louis** 2024
Institute for Informatics, Data Science, and Biostatistics Research Seminar
Reliable real-world evidence with EHR
3. **Washington University in St. Louis** 2024
CSE521S Wireless Sensor Networks Guest Lecture
Causal machine learning for real-world evidence generation
4. **International Chinese Statistical Association (ICSA) International Conference** 2023
Probabilistic machine learning on unstructured data
Adjusting for indirectly measured confounding using large-scale propensity score
5. **University of Colorado School of Medicine** 2023
Department of Biomedical Informatics Special Seminar
Improving the Reliability of Real-World Evidence Generation from Large-Scale Observational Data
6. **Ohio State University College of Medicine** 2023
Department of Biomedical Informatics Special Seminar
Improving the Reliability of Real-World Evidence Generation from Large-Scale Observational Data
7. **Vanderbilt University** 2023
Department of Biomedical Informatics Special Seminar
Improving the Reliability of Real-World Evidence Generation from Large-Scale Observational Data
8. **Washington University in St. Louis** 2023
Institute for Informatics, Data Science, and Biostatistics Special Seminar
Improving the Reliability of Real-World Evidence Generation from Large-Scale Observational Data
9. **Northeastern University** 2023
Department of Health Sciences Special Seminar
Improving the Reliability of Real-World Evidence Generation from Large-Scale Observational Data
10. **OHDSI Global Symposium** 2022
Lightening talk
When does statistical equality meet health equity: developing analytical pipelines to compare associational and causal fairness in their application to EHR data.
11. **OHDSI Global Symposium** 2022
Health Equity Workgroup
Evaluating and improving performance and racial fairness of algorithms for GFR estimation.
12. **OHDSI Global Symposium** 2022
Population-Level Estimation Workgroup
Adjusting for indirectly measured confounding using large-scale propensity score.
13. **Columbia University** 2021
Department of Biomedical Informatics Research Seminar
Algorithmic fairness in medicine: A case study in glomerular filtration rate (GFR) prediction.
14. **OHDSI Health Equity Workgroup Monthly Meeting** 2021
Assessing the impact of race on glomerular filtration rate (GFR) prediction.
15. **Columbia University** 2020
Department of Biomedical Informatics Research Seminar
Adjusting for unobserved confounding using large-scale propensity score.

16. **AMIA Annual Symposium** 2020
Causal Inference Panel
 The medical deconfounder: assessing treatment effects with electronic health records.
17. **University of Pennsylvania School of Medicine** 2020
SC-TRM Working Group Meeting
 The medical deconfounder: assessing treatment effects with electronic health records.

PROFESSIONAL ACTIVITIES

Conference & Workshop Organizing

- Scientific program committee member, AMIA Annual Symposium 2024
- Workshop organizer, OHDSI Global Symposium Tutorial: An Introduction to the Journey 2024
- Workshop chair, International Conference on AI in Medicine (AIME)
Reliable and Equitable Real-World Evidence Generation Workshop 2024
- Co-chair, I2BD Annual Symposium 2024
- Committee member, IEEE Conference on AI x Medicine, Health, and Care (AIMHC) 2024
- Session chair, Machine Learning for Health (ML4H) Symposium
Causality in Health AI Research Roundtable 2023
- Session chair, AMIA Annual Symposium
Precision Medicine and Disease Subtyping 2023

Grant Reviewing

- Reviewer, Patient-Centered Outcomes Research Institute (PCORI) 2024-Present

Journal & Conference Reviewing

- JAMIA Open 2024-Present
- Journal of Biomedical Informatics (JBI) 2023-Present
- Journal of Medical Internet Research (JMIR) 2023-Present
- Health Informatics Journal 2022-Present
- AMIA Annual Symposium 2019-Present
- Applied Clinical Informatics Journal (ACI) 2022-Present
- American Causal Inference Conference (ACIC) 2024
- Machine Learning for Healthcare (MLHC) 2020, 2022, 2023
- International Conference on Machine Learning (ICML) 2020
- Women in Machine Learning (WiML) Workshop 2019
- Machine Learning for Health (ML4H) Workshop 2019, 2020

TRAINEE MENTORSHIP

- Siqi Sun, postdoctoral fellow Aug 2024-present
- Mark-Daniels Tamkloe, PhD student Jun 2024-present

- Hsin Yi (Cindy) Chen, MD/PhD student Sep 2023-present
- Zhen Luo, Master's student Jun 2024-present
- Yiou Zhang, Master's student Jun-Aug 2024
- Ayushman Choudhury, undergraduate student May-Aug 2024
- Anand Ramabadrán, undergraduate student May-Aug 2024
- Yichen Sun, PhD student Feb-Apr 2024
- Tevin Kim, high school student Jun-Aug 2022

UNIVERSITY SERVICE

- Member, Biomedical Informatics and Data Science (BIDS) PhD Admission Committee
Division of Biology & Biomedical Sciences (DBBS), WashU Medicine 2023-present
- Member, BIDS@I2 Summer Research Internship Admission Committee
I2DB, WashU Medicine 2023-present
- Reviewer, Geospatial Research Initiative
Washington University 2024

TEACHING

- Computational Methods. *Columbia University*. Spring 2020
- Computer Applications in Health Care and Biomedicine. *Columbia University*. Fall 2019
- Principles of Biostatistics I&II. *Harvard T.H.Chan School of Public Health*. Summer 2017
- General Physics I&II. *Boston University*. 2012-2013

COMMUNITY OUTREACH & VOLUNTEER ACTIVITIES

- Mentor, AMIA Annual Symposium Career Development for Women Event 2022
- Mentor, Columbia DBMI Summer Research Program 2022
- Volunteer, OHDSI Global Symposium 2022
- Volunteer, NeurIPS 2022