LINYING ZHANG

622 West 168th Street PH20, New York, NY 10032
linying.zhang@columbia.edu
HomePage: linyingzhang.com ⋄ Twitter: @Z_Linying ⋄ GitHub: https://github.com/zhangly811

EDUCATION

Columbia University in the City of New York *August 2018 - May 2023 (expected)*
Ph.D. in Biomedical Informatics GPA: 3.96/4.0
Advisors: George Hripcsak and David Blei

Harvard. T.H.Chan School of Public Health *August 2016 - May 2018*
M.S. in Computational Biology and Quantitative Genetics (CBQG) GPA: 3.79/4.0
Advisor: Giovanni Parmigiani
CBQG Program Student Committee Co-chair.
Thesis: *Interactions between multiple myeloma cells and bone marrow stromal cells impact epigenetic profiles of multiple myeloma.*

Boston University *August 2011 - May 2014*
B.A. with Honors (*Summa Cum Laude*) in Biochemistry and Molecular Biology GPA: 3.74/4.0
Advisor: Ulla Hansen
Awards: Dean's List (6/6 semesters); Senior Book Award.
Thesis: *Establishing hepatocellular carcinoma cell lines with inducible expression of degradable LSF to investigate LSF regulation in cell cycle.*

RESEARCH INTERESTS

My research focuses on causal inference from electronic health records: 1) developing statistical and machine learning models to support inference about treatment effects under unmeasured confounding; 2) assessing biases in the health care process and designing models to counteract biases. My long-term research goal is to develop models that are fair, interpretable, and robust across databases by incorporating causal inference into machine learning.

PUBLICATIONS & PREPRINTS

1. **Zhang L.**, Richter L.R., Wang Y., Ostropolets A., Elhadad N., Blei D.M., Hripcsak G. A Bayesian causal inference approach for assessing fairness in clinical decision-making. ***arXiv preprint***, 2022.
2. **Zhang L.**, Richter L.R., Kim T., Hripcsak G. Evaluating and improving performance and racial fairness of algorithms for GFR estimation. ***medRxiv preprint***, 2022.
3. **Zhang L.**, Wang Y., Schuemie M.J., Blei D.M., Hripcsak G. Adjusting for indirectly measured confounding using large-scale propensity score. ***Journal of Biomedical Informatics***, 2022.
4. **Zhang L.**, Wang Y., Ostropolets A., Mulgrave J.J., Blei D.M., and Hripcsak G. The medical deconfounder: Assessing treatment effects with electronic health records. ***Machine Learning for Healthcare Conference (MLHC)***, 2019.
5. Richter L.R., Albert B.I., **Zhang L.**, Ostropolets A., Zitsman J.L., 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.
6. Song W., **Zhang L.**, Liu L., Sainlaire M., Karvar M., Kang M., Pullman A., Lipsitz S., Massaro A., Patil N., Jasuja R., Dykes P.C. Predicting hospitalization of COVID-19 positive patients using clinician-guided machine learning methods. ***Journal of the American Medical Informatics Association***. 2022.

7. Song W., **Zhang L.**, Liu L., Sainlaire M., Karvar M., Kang M., Pullman A., Lipsitz S., Massaro A., Patil N., Jasuja R., Dykes P.C. Predicting pressure injury using nursing assessment phenotypes and machine learning methods. ***Journal of the American Medical Informatics Association***, 28 (4), 759-765, 2021.

8. 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. ***Journal of the American Medical Informatics Association***, 27 (12), 1968-1976, 2020.

9. 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***, 3 (2), 281-289, 2020.

10. Ostropolets A., Albogami Y., Conover M, Banda J.M., Baumgartner W.A. Jr., Blacketer C., Desai P., DuVall S.L., Fortin S., Gilbert J.P., Golozar A., Ide J., Kanter A.S., Kern D.M., Kim C., Lai L.Y.H., Li C., Liu F., Lynch K.E., Minty E., Ins Neves M., Ng D.Q., Obene T., Pera V., Pratt N., Rao G., Rappoport N., Reinecke I., Saroufim P., Shoaibi A., Simon K., Suchard M.A., Swerdel J.N., Voss E.A., Weaver J., **Zhang L.**, Hripcsak G., and Ryan P.B. Reproducible Variability: Assessing investigator discordance across nine research teams attempting to reproduce the same observational study. ***BMC Medical Research Methodology***. 2022 [Under review]

11. 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., Doshi-Velez F.. Evaluating reinforcement learning algorithms in observational health settings. ***arXiv preprint***, 2018.

CONFERENCE ABSTRACTS (PEER-REVIEWED)

1. **Zhang L.**, Richter L.R., Wang Y., Ostropolets A., Elhadad N., Blei D.M., 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]

2. **Zhang L.**, Richter L.R., Blei D.M., 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]

3. **Zhang L.**, Richter L.R., Hripcsak G. Assessing the impact of race on glomerular filtration rate (GFR) prediction. In *OHDSI Global Symposium*. 2021. [Presentation]

4. **Zhang L.**, Wang Y., Ostropolets A., Chen R., Blei D.M., and Hripcsak G. The Multi-Outcome Medical Deconfounder: Assessing Treatment Effects on Multiple Renal Measures. In *AMIA Annual Symposium*. 2020. [Poster]

5. **Zhang L.**, Wang Y., Ostropolets A., Chen R., Blei D.M., and Hripcsak G. The multi-outcome medical deconfounder: assessing treatment effects on multiple renal measures. In *OHDSI Global Symposium*. 2020. [Poster]

6. **Zhang L.**, Wang Y., Ostropolets A., Mulgrave J.J., Blei D.M., and Hripcsak G. The medical deconfounder: assessing treatment effects with electronic health records. In *Women in Machine Learning (WiML) Workshop*. Vancouver, Canada. 2019. [Poster]

7. **Zhang L.**, Wang Y., Ostropolets A., Mulgrave J.J., Blei D.M., and Hripcsak G. The medical deconfounder: assessing treatment effects with electronic health records. In *Machine Learning for Health Workshop*. Vancouver, Canada. 2019. [Poster]

8. **Zhang L.**, Samur M.K., Szalat R., Epstein C.B., Prabhala R., Fulciniti M., Munshi N.C.*, 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]

9. **Zhang L.**, Samur M.K., Szalat R., Epstein C.B., Prabhala R., Fulciniti M., Munshi N.C.*, 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]

10. Song W., **Zhang L.**, Sainlaire M., Karvar M., Kang M., Pullman A., Massaro A., Patil N., Jasuja R., Dykes P.C. Predicting hospitalization of COVID-19 positive patients using machine learning methods. In *AMIA Annual Symposium*. 2021. [Presentation]

11. 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]

12. Ostropolets A., **Zhang L.**, Mulgrave J.J., Hripcsak G. Investigating female-male differences in risk factors for myocardial infarction using OHDSI tools. In *AMIA Annual Symposium*. 2019. [Poster]

13. 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]

14. Szalat R., Samur M.K., Ott C.J., Lawlor M., Epstein C., Abraham B.J., Lin C.Y., **Zhang L.**, Prabhala R., Farrell N., Wes K., Tai Y.T., Fulciniti M., Parmigiani G., Young R.A., Anderson K.C., and Munshi N.C. 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]

INVITED TALKS

When does statistical equality meet health equity: developing analytical pipelines to compare associational and causal fairness in their application to EHR data.
OHDSI Global Symposium. 2022.

Evaluating and improving performance and racial fairness of algorithms for GFR estimation.
OHDSI Global Symposium Health Equity Work Group. 2022.

Adjusting for indirectly measured confounding using large-scale propensity score.
OHDSI Global Symposium Population-Level Estimation Work Group. 2022.

Algorithmic fairness in medicine: A case study in glomerular filtration rate (GFR) prediction.
DBMI Research Seminar at Columbia University. 2021.

Assessing the impact of race on glomerular filtration rate (GFR) prediction.
OHDSI Health Equity Work Group. 2021.

Adjusting for unobserved confounding using large-scale propensity score.
DBMI Research Seminar at Columbia University. 2020.

The medical deconfounder: assessing treatment effects with electronic health records.
Causal Inference Panel at the AMIA Annual Symposium. 2020.

The medical deconfounder: assessing treatment effects with electronic health records.
SC-TRM Working Group Meeting at Perelman School of Medicine, University of Pennsylvania. 2020.

PROFESSIONAL ACTIVITIES

Volunteer for NeurIPS. 2022.

Mentor for the AMIA Annual Symposium Career Development for Women Event. 2022.

Panelist for the First Look Program at the AMIA Annual Symposium. 2022.

Volunteer for the OHDSI Global Symposium. 2022.

Mentor for the Columbia DBMI Summer Research Program. 2022.

Reviewer for Health Informatics Journal. 2022-Present.

Reviewer for Applied Clinical Informatics Journal. 2022-Present.

Reviewer for Machine Learning for Healthcare (MLHC) 2020 and 2022.

Reviewer for AMIA Annual Symposium 2022.

Reviewer for International Conference on Machine Learning (ICML) 2020.

Reviewer for WiML Workshop 2019.

Reviewer for Machine Learning for Health (ML4H) Workshop 2019 and 2020.

AWARDS

OHDSI Best Community Contribution Award in Methodological Research. 2022.

Women in Machine Learning Travel Award. 2019.

Senior Book Award. Boston University Department of Biology. 2014.

Dean's List. Boston University College of Arts and Sciences. Every semester between 2011 and 2014.

RESEARCH & WORK EXPERIENCE

Dana-Farber Cancer Institute April 2017 - May 2018
Graduate Researcher (Advisor: Giovanni Parmigiani)

· Built a pipeline to integrate RNA-Seq and ChIP-Seq data to study the impact of epigenetic changes on gene regulation in multiple myeloma (MM).
· Identified epigenetic changes in both MM and its neighboring cells via differential binding analysis and motif analysis.

Alnylam Pharmaceuticals Inc. June 2014 - June 2016
Research Scientist (Supervisor: Stuart Milstein)

· Supported drug candidate selection by performing high-throughput screening and data analysis.
· Supported Clinical Trial Application filling of ALN-AAT by assay development.
· Investigated the possibility of targeting APP with siRNA for treatment of Alzheimers disease.

Boston University March 2013 - May 2014
Undergraduate Researcher (Advisor: Ulla Hansen)

· Identified the role of transcription factor LSF in cell cycle progression in liver cancer cells.
· Initiated the establishment of an inducible cell line by optimizing transfection condition.
· Designed DNA constructs to achieve instant protein degradation in cell cycle by PCR.

TEACHING EXPERIENCE

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.

REFERENCES

George Hripcsak, MD, MS
Chair and Vivian Beaumont Allen Professor of Biomedical Informatics
622 West 168th Street, PH20
New York, NY 10032
hripcsak@columbia.edu
+1 212 305 5712

Noémie Elhadad, PhD
Associate Professor of Biomedical Informatics
622 West 168th Street, PH20
New York, NY 10032
noemie.elhadad@columbia.edu
+1 212 305 0509

David M. Blei, PhD
Professor of Statistics and of Computer Science
1255 Amsterdam Avenue
Room 1005 SSW
New York, NY 10016
david.blei@columbia.edu
+1 212 851 2132