

Qi Zhang, MSPH

qi.zhang0305@gmail.com | (404) 384-8015 | Atlanta, GA | [LinkedIn](#) | [Google Scholar](#)

EDUCATION

Emory University, Rollins School of Public Health, Atlanta, GA

PhD in Epidemiology (GPA: 4.00/4.00)

August 2020 - Expected May 2025

- Thesis: Causal Machine Learning Methods with Healthcare Databases
- Relevant Coursework: Causal Inference, Data Mining, Machine Learning, Statistical Computing, Modern Regression

Master of Science in Public Health, Epidemiology (GPA: 3.94/4.00)

May 2019

- Thesis: Identifying High-Risk Population for Common Diseases Using Polygenic Risk Prediction

Tsinghua University, Beijing, China

June 2017

Master in Kinesiology (GPA: 3.92/4.00)

Beijing Sport University, Beijing, China

July 2014

Bachelor of Science in Sports Medicine (GPA: 3.91/4.00)

SKILLS

Programming: Python, R (Shiny and R markdown), SAS, SQL, JAVA, MATLAB; Linux Bash Shell, Git, LaTeX

Causal Inference: G-computation, Propensity Score Methods, Marginal Structural Models, Negative Control Methods, Instrumental Variable, Difference in Difference, Doubly Robust Estimation, Experimental Design, Target Trial Emulation

Statistical Methods: Generalized Linear Models, Longitudinal Analysis, MCMC, Survival Analysis, Multiple Imputation

Natural language processing (NLP): RoBERTa, SciBERT, BiLSTM, SVM, K-NN, Neural Network, Random Forest

PROFESSIONAL EXPERIENCE

Pre-Doctoral Fellow, Amgen, Inc., Thousand Oaks, CA

June 2023 - Present

Supervisor: Christopher Kim, PhD

- Developed plasmode simulation and generative models to generate synthetic data that mimic data structure of a real dataset used to evaluate conditional average treatment effects in prostate cancer with python
- Explored pros and cons for the utility of various healthcare databases in the prostate cancer study
- Estimated the heterogenous treatment effects of novel hormone therapies (abiraterone vs. enzalutamide) as first-line therapies in patients with metastatic castration-resistant prostate cancer using Optum claims data

Summer Data Scientist Intern, Merck & Co., Inc., North Wales, PA

May - August 2023

Supervisor: Linda Shoener Dunham, BS

- Characterized chronic heart failure patients treated with vericiguat (VERQUVO®) in terms of demographic and clinical features in support of utilization increase of vericiguat among target patient population
- Conducted a matched case-control study using healthcare claims data (MarketScan and Optum) to identify top differentiating diagnosis codes and procedure codes prior to treatment to enhance knowledge of patient profile

Senior Statistical Analyst, Epidemiologic Research & Methods, LLC, Atlanta, GA

July 2019 - Present

Supervisor: W. Dana Flanders, MD, DSc & Cathy Lally, MSPH

- Performed data management and analysis of large EHR databases (e.g., MarketScan Commercial and Medicaid Databases) and write study protocols for clients of pharmaceutical and biotechnology companies
- Used R and SAS SQL to implement web scraping and data wrangling to retrieve and recode 11 GB FDA NDC data in 2009-2018, merged with CMS Medicare data by fuzzy matching, reducing process time by 50% via automation
- Developed macro to project prevalence of platinum resistant ovarian cancer and determine the treatable population and sample size for the clinical trial

Data Analyst, Department of Epidemiology, Emory University, Atlanta, GA

July 2019 - August 2020

Supervisor: Michael Goodman, MD, MPH

- Performed pharmacoepidemiologic analysis on the effectiveness of cross-sex drugs (e.g., estrogen and testosterone) on transgender health in the areas of cardiovascular diseases, hematology, dermatology, mental health, etc., in a longitudinal matched cohort using electronic medical records (EMR) from Kaiser Permanente health systems
- Conducted conditional Poisson regression, Cox models with Heaviside function (to account for time-varying exposure), counting process models, linear mixed models and multiple imputation
- Wrote manuscripts investigating 1) association between hormone therapy and suicide attempts among young transgender people using time-to-event analysis for multiple events in the self-controlled design; and 2) changes of hematologic and metabolic parameters over time in relation to hormone therapy in transgender adults relative to their matched cis-gender referents using longitudinal analyses

Data Analyst Intern, HealthMPowers, Inc., Atlanta, GA

April 2018 - May 2019

Supervisor: Julie Gazmararian, PhD, MPH

- Survey data manipulation and multivariate analysis with R and SQL to evaluate effectiveness of interventions in the child care program on improving nutrition and physical activity of children
- Designed and maintained visualizations with Tableau on survey data of 86 child care centers across Georgia, provided training to health educators, prepared and presented annual report

Healthcare Analyst Intern, Emory Healthcare, Atlanta, GA

September 2018 - May 2019

- Extracted data on service lines performing precertification functions from 246 entities and facilities across Emory Healthcare, collaborated with technology team to create a GIS map of service areas to streamline the process
- Conducted phone interviews with around 50 staff in multiple departments regarding precertification process
- Presented project and strategy to improve efficiency at the Emory Healthcare Revenue Cycle Opportunities meeting

RESEARCH EXPERIENCE**Research Assistant, Department of Epidemiology, Emory University, Atlanta, GA***A Comparison of Causal Forests and the DR-Learner for Estimating Conditional Average Treatment Effects*

Supervisor: Ashley Naimi, PhD

March 2022 - Present

- Applied Causal Forest and doubly robust learner to estimate the conditional average treatment effect of dietary intake on adverse pregnancy outcome using R (*grf* package)
- Designed simulation studies to automatically detect effect modifiers under different scenarios of effect heterogeneity

Text Classification Models for Natal Sex Identification from Electronic Health Records of Transgender Population

Supervisor: Abeed Sarker, PhD & Michael Goodman, MD, MPH

October 2021 - Present

- Identified transgender cohort using natural language processing models based on free-text clinical notes, and compared the classification performances of the transformer-based model with traditional machine learning models (including support vector machine, K-nearest Neighbor, neural network and random forest)
- Investigated the optimal prediction threshold selection by evaluating multiple metrics (recall, precision and F1-score)
- Designed a practical validation study by stratified sampling and manually review of clinical notes to assess the trasportibility of the model

Bayesian Priors for Bias Parameters of Binary Exposure Misclassification in Probabilistic Bias Analysis

Supervisor: Timothy L. Lash, PhD

September 2020 - September 2024

- Implemented probabilistic bias analysis with vectorized R code on HPC cluster based on Monte Carlo simulations
- Compared the performances of a Bayesian prior and continuity correction in parameterizing beta distributions of bias parameters yielded from the validation study with balanced sampling design in misclassification of exposure

Research Assistant, Department of Computer Science, Emory University, Atlanta, GA September 2022 - Present*Data-driven Heterogeneous Treatment Effect Estimation*

Supervisor: Carl Yang, PhD

- Implemented causal structure (DAG) learning for selecting optimal adjustment set to quantify heterogeneity of causal effects in synthetic data
- Wrote a survey for causal machine learning methods in dynamic treatment regime and heterogeneous treatment effects (HTE), and compared performances of different HTE estimators with synthetic data

Thesis, Department of Epidemiology, Emory University, Atlanta, GA

April 2018 - May 2019

Impact of Predictive Ability on Identifying Higher-Risk Population for Common Diseases in Polygenic Risk Prediction

Supervisor: A. Cecile J.W. Janssens, PhD

- Investigated the influence of improvement of the area under the receiver operating characteristic curve (AUC) on clinical performance of polygenic risk scores by assessing accuracy and reclassification measures
- Developed multiple R functions to simulate hypothetical risk data with genetic variants and disease status for 100,000 individuals, constructed prediction models based on the risk genetic variants using Bayes' theorem

Graduate Research Analyst

September 2014 - July 2017

Research Center of Sports and Health Science, Tsinghua University, Beijing, China*Body Composition Analysis Using Bioelectrical Impedance in Chinese Population*

Supervisor: Jingmin Liu, PhD

- Designed and performed data collection and analysis for the development and application of a body composition analyzer (BCA) product of Tongfang Health Technology Co., Ltd.
- Collaborated with researchers from 5 institutions to implement a longitudinal cohort study on association between sarcopenia and nutrition, physical activity and drugs among people aged 50 years and older in Jiangsu Province

Principal Investigator

December 2011 - September 2013

Laboratory of Neurophysiology, Beijing Sport University, Beijing, China*Effects of "Acupuncture Serum" on Proliferation and Differentiation of C₂C₁₂ Myoblasts*

Supervisor : Li Zhao, PhD

- Received a \$4,300 grant from National College Students' Innovative Entrepreneurial Training Plan Program to support a study which investigated the mechanism of acupuncture inducing regeneration of skeletal muscle from the perspectives of humoral regulation and cellular effects

PUBLICATIONS

- Zhang Q**, MacLehose RF, Collin LJ, Ahern TP, Lash TL. Parameterization of beta distributions for bias parameters of binary exposure misclassification in probabilistic bias analysis. *Epidemiology*. 2025 Mar 1;36(2):237-44.
- Anike O, **Zhang Q**, Bhasin S, Flanders WD, Getahun D, Haw JS, Huybrechts KF, Lash TL, McCracken CE, Roblin D, Silverberg MJ. Association between gender-affirming hormone therapy and measures of glucose metabolism: A longitudinal study. *The Journal of Clinical Endocrinology & Metabolism*. 2024 May 30:dgae353.
- Flanders WD, Waller LA, **Zhang Q**, Getahun D, Silverberg M, Goodman M. Negative control exposures: causal effect identifiability and use in probabilistic-bias and Bayesian analyses with unmeasured confounders. *Epidemiology*. 2022 Jul 27;33(6):832-9.
- Pampati S, Flanders WD, **Zhang Q**, Ahmedani BK, Bhasin S, Getahun D, Lash TL, McCracken C, Sandberg DE, Silverberg MJ, Tangpricha V. Prevalence of Self-Inflicted Injuries Among Transgender and Gender Diverse Adolescents and Young Adults Compared to Their Peers: An Examination of Interaction with Mental Health Morbidity. *Annals of Epidemiology*. 2023 Mar 11.
- Lee YT, **Zhang Q**, Elmontser M, Stevenson E, Gazmararian JA. Association between early care and education centres policies and practices and children's healthy behaviours. *The International Journal of Health Planning and Management*. 2023 Nov;38(6):1743-56.
- Antun A, **Zhang Q**, Bradlyn A, Flanders WD, Getahun D, Lash TL, Nash R, Roblin D, Silverberg MJ, Tangpricha V, Vupputuri S, Goodman M. Longitudinal changes in hematologic parameters among transgender people receiving hormone therapy. *Journal of the Endocrine Society* 4, no. 11 (2020): bvaa119.
- Hashemi L, **Zhang Q**, Getahun D, Jasuja GK, McCracken C, Pisegna J, Roblin D, Silverberg MJ, Tangpricha V, Vupputuri S, Goodman M. Longitudinal Changes in Liver Enzyme Levels Among Transgender People Receiving Gender Affirming Hormone Therapy. *The Journal of Sexual Medicine*. 2021 Aug 5.
- Zhang Q**, Rechler W, Bradlyn A, Flanders WD, Getahun D, Lash TL, McCracken C, Nash R, Panagiotakopoulos L, Roblin D, Sandberg DE. Changes in size and demographic composition of transgender and gender non-binary population receiving care at integrated health systems. *Endocrine Practice*. 2020 Dec 28.
- Zhang Q**, Goodman M, Adams N, Banik S, Cornell T, Kreukels B, Motmans J, Snyder R, Coleman E. Epidemiological considerations in transgender health: A systematic review with focus on higher quality data. *International Journal of Transgender Health* 21, no. 2 (2020): 125-137.
- Mak J, Shires DA, **Zhang Q**, Prieto L, Ahmedani BK, Becerra-Culqui TA, Bradlyn A, Flanders WD, Getahun D, Giammattei SV, Hunkeler E, Lash TL, Millman A, Nash R, Quinn VP, Robinson B, Roblin D, Silverberg MJ, Slovis J, Tangpricha V, Vupputuri S, Goodman M. Suicide attempts among a cohort of transgender and gender diverse people. *American journal of preventive medicine*. 2020 Oct 1;59(4):570-7.
- Braun H, **Zhang Q**, Getahun D, Silverberg MJ, Tangpricha V, Goodman M, Yeung H. Moderate-to-Severe Acne and Mental Health Symptoms in Transmasculine Persons Who Have Received Testosterone. *JAMA dermatology*. 2021 Jan 20.

- Ragmanauskaite L, **Zhang Q**, Kim J, Getahun D, Silverberg MJ, Goodman M, Yeung H. Survey of Dermatologic Procedures in Transgender Adults. *Dermatologic Surgery*. 2021 Oct 1;47(10):1379-83.
- To M, **Zhang Q**, Bradlyn A, Getahun D, Giammattei S, Nash R, Owen-Smith AA, Roblin D, Silverberg MJ, Tangpricha V, Vupputuri S. Visual Conformity with Affirmed Gender or “Passing”: Its Distribution and Association with Depression and Anxiety in a Cohort of Transgender People. *The Journal of Sexual Medicine*. 2020 Oct 1;17(10):2084-92.
- Braun H, Thompson EC, **Zhang Q**, Tangpricha V, Goodman M, Yeung H. Prevalence of Psoriasis and Perceived Association with Hormone Therapy in Transgender Adults. *Transgender Health*. 2022 Mar 28.
- Islam N, Nash R, **Zhang Q**, Panagiotakopoulos L, Daley T, Bhasin S, Getahun D, Sonya Haw J, McCracken C, Silverberg MJ, Tangpricha V. Is There a Link Between Hormone Use and Diabetes Incidence in Transgender People? Data From the STRONG Cohort. *The Journal of Clinical Endocrinology & Metabolism*. 2021 Nov 30.
- Ragmanauskaite L, Kim J, **Zhang Q**, Luk K, Getahun D, Silverberg MJ, Goodman M, Yeung H. Self-reported tattoo prevalence and motivations in transgender adults: a cross-sectional survey. *Dermatology Online Journal*. 2020;26(12).
- Yeung H, Ragmanauskaite L, **Zhang Q**, Kim J, Tangpricha V, Getahun D, Silverberg MJ, Goodman M. Prevalence of moderate to severe acne in transgender adults: A cross-sectional survey. *Journal of the American Academy of Dermatology*. 2020 Nov 1;83(5):1450-2.
- Becerra-Culqui T, Swiatek D, Dizon B, Getahun D, Silverberg M, **Zhang Q**, Im T, Goodman M. Challenging Norms: The Impact of Transgender and Gender-Diverse Realities on Work and School Participation. *The American Journal of Occupational Therapy*. 2024 May 1;78(3):7803205150.
- Hensley M, Goodman M, Madani R, Jaggi P, Keesari R, **Zhang Q**, Oster ME. Cardiac complications in children with acute COVID-19 vs multisystem inflammatory syndrome in children (MIS-C). *American Heart Journal*. 2023 Sep 1;263:177-82.
- Goodman M, Li J, Flanders WD, Mahood D, Anthony LG, **Zhang Q**, LaKind JS. Epidemiology of PCBs and neurodevelopment: Systematic assessment of multiplicity and completeness of reporting. *Global Epidemiology*. 2020 Nov 4:100040.
- Goodman M, Yacoub R, Getahun D, McCracken CE, Vupputuri S, Lash TL, ..., **Zhang Q**, ..., Sandberg DE. Cohort profile: pathways to care among people with disorders of sex development (DSD). *BMJ Open*. 2022 Sep 21;12(9):e063409-.

CONFERENCES & PRESENTATIONS

- Zhang Q**, MacLehose RF, Collin LJ, Ahern TP, Lash TL. Bayesian Priors for Bias Parameters of Binary Exposure Misclassification in Probabilistic Bias Analysis. Oral presentation at International Society of Pharmacoepidemiology (ISPE) Annual meeting, Berlin, Germany, August 2024
- Zhang Q**, Giulio Biscardi, Simone Colombara, Giovanni Toto, Carlo Zaccardi, Corwin Zigler. Heterogeneity of Comparative Treatment Effects on Patients with Lower Grade Glioma. Oral presentation at Data Research Camp (supported by the excellence program “Rethinking Data Science: advances in causal inference and statistical learning across domains”), Florence, Italy, July 2024
- Zhang Q**, Yu YH, Naimi AI. A Comparison of Causal Forests and the DR-Learner for Estimating Conditional Average Treatment Effects. Oral presentation at Society for Epidemiologic Research (SER), Austin, Texas, June 2024
- Zhang Q**, Guo Y, Lash TL, Sarker A, Goodman M. Text Classification Models for Natal Sex Identification from Electronic Health Records of Transgender Population, Poster presentation at Society for Epidemiologic Research (SER), Austin, Texas, June 2024
- Zhang Q**, Yu YH, Naimi AI. A Comparison of Causal Forests and the DR-Learner for Estimating Conditional Average Treatment Effects. Poster presentation at American Causal Inference Conference (ACIC), Seattle, Washington, May

2024

Zhang Q, Guo Y, Al-Garadi MA, Goodman M, Sarker A. Text Classification Models for Creating a Transgender Cohort from Electronic Health Records, Poster presentation at American Medical Informatics Association Annual Symposium (AMIA), New Orleans, LA, November 2023

Zhang Q. Causal Machine Learning Methods in Epidemiology: An Application to Transgender Health Research. Oral presentation at Society for Epidemiologic Research (SER) Dissertation Workshop, Portland, Oregon, June 2023

Zhang Q, MacLehose RF, Collin LJ, Ahern TP, Lash TL. Parameterization of Beta Distributions for Bias Parameters of Binary Exposure Misclassification in Probabilistic Bias Analysis. Poster presentation at Society for Epidemiologic Research (SER), Portland, Oregon, June 2023

Zhang Q, Goodman M. Hormone therapy and risk of venous thromboembolism in transgender women. Oral presentation at Mid-Year Meeting of the International Society of Pharmacoepidemiology (ISPE), Reykjavik, Iceland, April 2023.

Zhang Q, Ramos R, Dong X. Effects of Valsalva Maneuver on Myoelectrical Activities of Trunk Muscles during the Transferring of Patients between Hospital Bed and Wheelchair. Poster presentation at American College of Sports Medicine Annual Meeting Texas Chapter, Waco, TX, February 2017

Zhang Q, Liu J, Zhao L. Effects of “Acupuncture Serum” on Proliferation and Differentiation of C₂C₁₂ Myoblasts. Poster presentation at International Convention on Science, Education and Medicine in Sport, São Paulo, Brazil, September 2016

Zhang Q, Sheng Y, He Y. Blind navigation: a belt compass system. Oral presentation at CERN Summer Webfest (Annual Hackathon), Meyrin, Switzerland, July 2016

PEER REVIEW

Journal

- *IEEE Transactions on Neural Networks and Learning Systems*
- *Epidemiology*
- *Epidemiologic Methods*

Conference

- American Causal Inference Conference
- Society for Epidemiologic Research Annual Meeting
- American Medical Informatics Association Annual Symposium

LEADERSHIP & SERVICES

Organizer, Causal Inference Reading Group

March 2022 - Present

- Organized a student-led causal inference reading group, comprising students in biostatistics, epidemiology, and biomedical informatics from multiple universities, that meet and discuss articles and books biweekly
- Invited guest speakers from the area of causal inference to give talks and attend discussions

Rollins Ambassador, RSPH Office of Admissions & Student Services, Emory University January 2018 - May 2019

- Supported recruitment efforts by leading campus tours for prospective students and served as a representative of Rollins to answer questions related to admissions and student life

Event Chair, Emory Global Health Organization, Emory University January 2018 - December 2018

- Organized and coordinated events to promote global health awareness and advocacy

Volunteer, Zijing Volunteer Team, Tsinghua University November 2015 - January 2016

- Narrated audiobooks for the visually disabled children
- Organized promotion of blood donation

Vice President, Graduate Union in School of Social Sciences, Tsinghua University July 2015 - July 2016

- Developed a sports club, recruited 80 members, set up membership system, and organized weekly training
- Received \$1,000 competitive funding from Graduate Students' Affairs Committee by presenting a pitch

HONORS

- Recipient of Best Application Award, the Bumblekite Machine Learning Summer School in Health, Care and Biosciences (ETH Zurich), 2024
- Recipient of Cambridge Centre for AI in Medicine Summer School Scholarship, 2024
- Recipient of Oxford Machine Learning Summer School Scholarship, 2024
- Recipient of ACIC Travel Scholarship, American Causal Inference Conference, 2024
- Recipient of Penn Causal Inference Summer Institute Scholarship, 2024
- Recipient of Global Student Travel Support Scholarship, Emory university, 2024
- Recipient of ISPE Mid-Year Meeting Scholarship, International Society of Pharmacoepidemiology, 2023
- Recipient of SER Dissertation Workshop Scholarship, The Society for Epidemiologic Research, 2023
- Recipient of Outstanding Thesis Award, Tsinghua University, 2017
- Excellent Graduate, Tsinghua University, 2017
- Recipient of Qinghe Scholarship, Tsinghua University, 2016
- Recipient of the Best Innovative Project Award in CERN Summer Webfest, the European Organization for Nuclear Research, 2016
- Excellent Graduate of Beijing, Beijing City Board of Education, 2014
- Recipient of Outstanding Thesis Award, Beijing Sport University, 2014
- Excellent Graduate, Beijing Sport University, 2014
- Dean's list, Beijing Sport University, 2011, 2012 & 2013
- Recipient of National Scholarship, The Ministry of Education of China, 2011, 2012 & 2013
- Recipient of National College Students' Innovative Entrepreneurial Training Plan Program Grant, The Ministry of Education of China, 2011 & 2013

TEACHING

Teaching Assistant (lab lecturer), Emory University

- Epidemiologic Methods I (EPI 530), 2019
- Epidemiologic Methods II (EPI 540), 2020
- Epidemiologic Methods III (EPI 550), 2021
- Epidemiologic Methods IV (EPI 560), 2023