Qingyuan Zhao

Address:	Statistical Laboratory, University of Cambridge D1.01, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WB, United Kingdom.
Phone:	(+44) 1223337995
Email:	qyzhao@statslab.cam.ac.uk
Homepage:	http://www.statslab.cam.ac.uk/~qz280

Academic positions

Aug 2019— University Lecturer in Statistics, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge.

Feb 2020— Fellow, Corpus Christi College.

Aug 2016—Jul 2019 Postdoctoral Fellow, Department of Statistics, University of Pennsylvania.

Education

Sep 2011—Sep 2016 Ph.D. in Statistics, Stanford University.

Sep 2007—Jun 2011 B.S. in Mathematics, University of Science and Technology of China.

Publications

Peer-reviewed articles

1. Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec, "SEISMIC: A self-exciting point process model for predicting tweet popularity," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, ACM Press, 2015, ISBN: 9781450336642. DOI: 10.1145/2783258.2783401.

2. J. Wang, Q. Zhao, T. Hastie, and A. B. Owen, "Confounder adjustment in multiple hypothesis testing," *Annals of Statistics*, vol. 45, no. 5, pp. 1863–1894, 2017. DOI: 10.1214/16-aos1511.

3. Q. Zhao and D. Percival, "Entropy balancing is doubly robust," *Journal of Causal Inference*, vol. 5, no. 1, 2016. DOI: 10.1515/jci-2016-0010.

4. Q. Zhao, D. S. Small, and P. R. Rosenbaum, "Cross-screening in observational studies that test many hypotheses," *Journal of the American Statistical Association*, vol. 113, no. 523, pp. 1070–1084, 2018. DOI: 10.1080/01621459.2017.1407770.

5. Q. Zhao and D. S. Small, "Graphical diagnosis of confounding bias in instrumental variable analysis," *Epidemiology*, vol. 29, no. 4, e29–e31, 2018. DOI: 10.1097/ede.0000000000000822.

6. Q. Zhao, "On sensitivity value of pair-matched observational studies," *Journal of the American Statistical Association*, vol. 114, no. 526, pp. 713–722, 2018. DOI: 10.1080/01621459.2018.1429277.

7. Q. Zhao, "Covariate balancing propensity score by tailored loss functions," *Annals of Statistics*, vol. 47, no. 2, pp. 965–993, 2019. DOI: 10.1214/18-aos1698.

8. H. Y. He, K. Basu, Q. Zhao, and A. B. Owen, "Permutation p-value approximation via generalized Stolarsky invariance," *Annals of Statistics*, vol. 47, no. 1, pp. 583–611, 2019. DOI: 10.1214/18-aos1702.

9. Q. Zhao, D. S. Small, and W. Su, "Multiple testing when many p-values are uniformly conservative, with application to testing qualitative interaction in educational interventions," *Journal of the American Statistical Association*, vol. 114, no. 527, pp. 1291–1304, 2018. DOI: 10.1080/01621459.2018.1497499.

10. J. H. Silber, J. G. Reiter, P. R. Rosenbaum, Q. Zhao, D. S. Small, B. A. Niknam, A. S. Hill, L. L. Hochman, R. R. Kelz, and L. A. Fleisher, "Defining multimorbidity in older surgical patients," *Medical Care*, vol. 56, no. 8, pp. 701–710, 2018. DOI: 10.1097/mlr.0000000000000947.

11. J. Bowden, F. D. G. M, C. Minelli, Q. Zhao, D. A. Lawlor, N. A. Sheehan, J. Thompson, and G. D. Smith, "Improving the accuracy of two-sample summary-data Mendelian randomization: Moving beyond the nome assumption," *International Journal of Epidemiology*, vol. 48, no. 3, pp. 728–742, 2018. DOI: 10.1093/ije/dyy258.

12. Q. Zhao, J. Wang, W. Spiller, J. Bowden, and D. S. Small, "Two-sample instrumental variable analyses using heterogeneous samples," *Statistical Science*, vol. 34, no. 2, pp. 317–333, 2019. DOI: 10.1214/18-sts692.

13. L. Keele, Q. Zhao, R. R. Kelz, and D. Small, "Falsification tests for instrumental variable designs with an application to tendency to operate," *Medical Care*, p. 1, 2018. DOI: 10.1097/mlr.0000000000001040.

14. Q. Zhao, D. S. Small, and B. B. Bhattacharya, "Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 81, no. 4, pp. 735–761, 2019. DOI: 10.1111/rssb.12327.

15. Q. Zhao, Y. Chen, J. Wang, and D. S. Small, "Powerful three-sample genome-wide design and robust statistical inference in summary-data Mendelian randomization," *International Journal of Epidemiology*, vol. 48, no. 5, pp. 1478–1492, 2019. DOI: 10.1093/ije/dyz142.

16. Q. Zhao, J. Wang, G. Hemani, J. Bowden, and D. S. Small, "Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score," *Annals of Statistics*, vol. 48, no. 3, pp. 1742–1769, 2020. DOI: 10.1214/19-aos1866.

17. B. Zhang, J. Weiss, D. S. Small, and Q. Zhao, "Selecting and ranking individualized treatment rules with unmeasured confounding," *Journal of the American Statistical Association*, vol. to appear, 2020. DOI: 10.1080/01621459.2020.1736083.

18. Q. Zhao, N. Ju, S. Bacallado, and R. D. Shah, "BETS: The dangers of selection bias in early analyses of the coronavirus disease (COVID-19) pandemic," *Annals of Applied Statistics*, 2020, In press. arXiv: 2004.07743 [stat.AP].

19. A. P. Christie, D. Abecasis, M. Adjeroud, J. C. Alonso, T. Amano, A. Anton, B. P. Baldigo, R. Barrientos, J. E. Bicknell, D. A. Buhl, J. Cebrian, R. S. Ceia, L. Cibils-Martina, S. Clarke, J. Claudet, M. D. Craig, D. Davoult, A. D. Backer, M. K. Donovan, T. D. Eddy, F. M. França, J. P. A. Gardner, B. P. Harris, A. Huusko, I. L. Jones, B. P. Kelaher, J. S. Kotiaho, A. López-Baucells, H. L. Major, A. Mäki-Petäys, B. Martín, C. A. Martín, P. A. Martin, D. Mateos-Molina, R. A. McConnaughey, M. Meroni, C. F. J. Meyer, K. Mills, M. Montefalcone, N. Noreika, C. Palacín, A. Pande, C. R. Pitcher, C. Ponce, M. Rinella, R. Rocha, M. C. Ruiz-Delgado, J. J. Schmitter-Soto, J. A. Shaffer, S. Sharma, A. A. Sher, D. Stagnol, T. R. Stanley, K. D. E. Stokesbury, A. Torres, O. Tully, T. Vehanen, C. Watts, Q. Zhao, and W. J. Sutherland, "Quantifying and addressing the prevalence and bias of study designs in the environmental and social sciences," *Nature Communications*, vol. 11, no. 1, p. 6377, 2020. DOI: 10.1038/s41467-020-20142-y.

20. M. J. Tudball, J. Bowden, R. A. Hughes, A. Ly, M. R. Munafo, K. Tilling, Q. Zhao, and G. Davey Smith, "Mendelian randomisation with coarsened exposures," *Genetic Epidemiology*, to appear, Aug. 2020. DOI: 10.1101/2020.08.20.20178921.

Invited contributions/Short letters

21. Zhao, Q., Zheng, C., Hastie, T., & Tibshirani, R. (2016). Comment on "Causal inference using invariant prediction". *Journal of the Royal Statistical Society (Series B, Statistical Methodology)*, 78(5), 1005–1007.

22. Zhao Q. & Panigrahi S. (2019). Selective inference for effect modification: An empirical investigation. *Observational Studies* 5, 131–140.

23. Q. Zhao, L. J. Keele, and D. S. Small, "Comment: Will competition-winning methods for causal inference also succeed in practice?" *Statistical Science*, vol. 34, no. 1, pp. 72–76, 2019. DOI: 10.1214/18-sts680.

24. Q. Zhao and T. Hastie, "Causal interpretations of black-box models," *Journal of Business & Economic Statistics*, pp. 1–10, 2019. DOI: 10.1080/07350015.2019.1624293.

25. S. Bacallado, Q. Zhao, and N. Ju, "Letter to the editor: Generation interval for COVID-19 based on symptom onset data," *Eurosurveillance*, vol. 25, no. 29, 2020. DOI: 10.2807/1560-7917.es.2020.25.29.2001381.

Preprints

26. Q. Zhao, T. Hastie, and D. Pregibon, *Estimation and prediction in sparse and unbalanced tables*, 2017. arXiv: 1703.02081 [stat.CO].

27. Q. Zhao, D. S. Small, and A. Ertefaie, *Selective inference for effect modification via the lasso*, 2017. arXiv: 1705.08020 [stat.ME].

28. Y. Song and Q. Zhao, *Performance evaluation with latent factors*, 2018. DOI: 10.2139/ssrn.3216272. SSRN: 3216272.

29. M. Tudball, Q. Zhao, R. Hughes, K. Tilling, and J. Bowden, "An interval estimation approach to sample selection bias," 2019. arXiv: 1906.10159 [stat.ME].

30. Q. Zhao, J. Wang, Z. Miao, N. Zhang, S. Hennessy, D. S. Small, and D. J. Rader, "The role of lipoprotein subfractions in coronary artery disease: A Mendelian randomization study," Jul. 2019. DOI: 10.1101/691089.

31. Q. Zhao, Y. Chen, and D. S. Small, "Analysis of the epidemic growth of the early 2019-ncov outbreak using internationally confirmed cases," Feb. 2020. DOI: 10.1101/2020.02.06.20020941.

32. C. Y. Shapland, Q. Zhao, and J. Bowden, "Profile-likelihood Bayesian model averaging for two-sample summary data Mendelian randomization in the presence of horizontal pleiotropy," Feb. 2020. DOI: 10.1101/2020.02.11.943712.

33. H. Kang, Y. Jiang, Q. Zhao, and D. S. Small, "ivmodel: An R package for inference and sensitivity analysis of instrumental variables models with one endogenous variable," 2020. arXiv: 2002.08457 [stat.AP].

34. J. Wang, Q. Zhao, J. Bowden, G. Hemani, G. D. Smith, D. S. Small, and N. R. Zhang, "Causal inference for heritable phenotypic risk factors using heterogeneous genetic instruments," May 2020. DOI: 10.1101/2020.05.06.077982. bioRxiv: 10.1101/2020.05.06.077982.

35. D. Iong, Q. Zhao, and Y. Chen, "A latent mixture model for heterogeneous causal mechanisms in Mendelian randomization," 2020. arXiv: 2007.06476 [stat.AP].

36. Q. Zhao, L. J. Keele, D. S. Small, and M. M. Joffe, "A note on post-treatment selection in studying racial discrimination in policing," 2020. arXiv: 2009.04832 [stat.AP].

37. T. Ye, J. Shao, and Q. Zhao, "Principles for covariate adjustment in analyzing randomized clinical trials," 2020. arXiv: 2009.11828 [stat.ME].

Talks and Presentations

Aug 2015 *"SEISMIC: A Self-Exciting Point Process Model for Predicting Tweet Popularity"*
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), Sydney, Australia.

Jan 2016 *"Confounder Adjustment in Multiple Hypothesis Testing"*
Rutgers University (Seminar in Department of Statistics and Biostatistics), New Brunswick, USA
New York University (Seminar in Department of Information, Operations and Management Science), New York, USA.

Aug 2016 *"Causal Interpretations of Black-Box Models"*
Joint Statistical Meetings (JSM) JBES invited session, Chicago, USA.

May 2017 *"Selective Inference for Effect Modification"*
Atlantic Causal Inference Conference (ACIC) 2017 (invited session), Chapel Hill, USA.

Nov 2017 *"Two Sample Instrumental Variable Analysis: Challenges and Some Progress"*
University of Bristol (seminar in MRC Integrative Epidemiology Unit), Bristol, UK.

Mar 2018 *"Mendelian Randomization: From Genetic Association to Epidemiological Causation"*
University of Minnesota (Statistics Seminar), Minneapolis, USA
Johns Hopkins University (Statistical Genetics and Causal Inference working groups meeting), Baltimore, USA

Apr 2018 *"Mendelian Randomization: From Genetic Association to Epidemiological Causation"*
UC Berkeley (Biostatistics Seminar), Berkeley, USA
Stanford University (Statistics Seminar), Stanford, USA

May 2018 *"Bootstrapping Sensitivity Analysis"*
Atlantic Causal Inference Conference (invited session), Pittsburgh, USA

Jun 2018 *"Confounder Adjustment in Multiple Hypothesis Testing"*
EcoSta 2018 (invited session), Hong Kong, China.

Jun 2018 *"Selective Inference for Effect Modification"*
ICSA Applied Statistics Symposium (invited session), New Brunswick, USA.

Oct 2018 *"Genome-wide Mendelian randomization under pervasive pleiotropy"*
American Society of Human Genetics Annual Meeting 2018 ("Reviewer's choice" poster presentation, top 10% of poster abstracts), San Diego, USA

Dec 2018 *"Mendelian randomization: A comprehensive statistical approach and applications to preventing heart disease"*
University of British Columbia (Statistics Seminar), Vancouver, Canada
Rutgers University (Biostatistics and Epidemiology Seminar), Piscataway, New Jersey, USA
Cornell University (Statistics Seminar), Ithaca, New York, USA

Jan 2019 *"Mendelian randomization: A comprehensive statistical approach and applications to preventing heart disease"*
University of Washington (Biostatistics Seminar), Seattle, USA
University of Southern California (Marshall School Statistics Seminar), Los Angeles, USA
Yale University (Biostatistics Seminar), New Heaven, USA
Rutgers University (Statistics Seminar), Piscataway, USA
University of Minnesota (Statistics Seminar), Minneapolis, USA
University of Michigan (Statistics Seminar), Ann Arbor, USA

Feb 2019 *"Mendelian randomization: A comprehensive statistical approach and applications to preventing heart disease"*
University of Cambridge (Statistics Seminar), Cambridge, UK
Harvard University (Statistics Seminar), Cambridge, USA
University of Wisconsin (Statistics Seminar), Madison, USA
Carnegie Mellon University (Statistics Seminar), Pittsburgh, USA.

Jun 2019 *"Bootstrapping Sensitivity Analysis"*
Bayesian Causal Inference Workshop (invited session), Ohio State University, Columbus, USA

Apr 2019 *"Is (empirical) Bayes the future for instrumental variable estimation?"*
Banff International Research Station (BIRS) Workshop on "New and Evolving Roles of Shrinkage in Large-Scale Prediction and Inference", Banff, Canada.

Jun 2019 *"Statistical considerations in Summary-Data Mendelian randomization"*
WNAR Annual Meeting (invited session), Portland, USA

Jul 2019 *"Mendelian randomization tutorial"*
CCI Causal Inference Summer Institute, Rutgers University, USA

Jul 2019 *"The statistics of summary-data Mendelian randomization"*
Mendelian Randomization Conference (invited), Bristol, UK

Oct 2019 *"Using sparsity to overcome unmeasured confounding: Two examples"*
MRC Biostatistics Unit Seminar, University of Cambridge, UK

Dec 2019 *"Bootstrapping sensitivity analysis for IPW estimators"*
Opening workshop in Causal Inference, Statistics and Applied Mathematical Sciences Institute (SAMSI), Durham, North Carolina, USA

Feb 2020 *"The cycle of statistical research"*
CCIMI seminar, Cambridge, UK

May 2020 *"BETS: The dangers of selection bias in early analyses of the coronavirus disease (COVID-19) pandemic"*
Yale University (Biostatistics Seminar), online.

Aug 2020 *"Bootstrapping sensitivity analysis"*
Joint Statistical Meeting. Invited session, online.

Aug 2020 *"Using sparsity to overcome unmeasured confounding"*
Annual Conference of the International Society of Clinical Biostatistics (ISCB). Invited session, online.

Sep 2020 *"Sensitivity analysis for observational studies"*
Yale University (Biostatistics Seminar), online.

Sep 2020 *"BETS: The dangers of selection bias in early analyses of the coronavirus disease (COVID-19) pandemic"*
Ohio State University (Statistics Seminar), online.

Sep 2020 *"Discovering mechanistic heterogeneity using Mendelian randomization"*
Pacific Causal Inference Conference, online.

Dec 2020 *"Selection bias in 2020"*
Causal Inference Online Seminar, online; University of Science and Technology of China, online.

Software

1. bets.covid19 on CRAN for likelihood inference for early epidemic data (author and maintainer). p
2. bootsens on GitHub for sensitivity analysis using bootstrap (author and maintainer).

3. mr.raps on CRAN for Mendelian randomization via robust adjusted profile score (author and maintainer).
4. ivmodel on CRAN for instrumental variable modeling (author of two diagnostics functions, `iv.diagnosis` and `iv.diagnosis.plot`).
5. CrossScreening on CRAN for cross-screening in observational studies that test many hypotheses (author and maintainer).
6. cate on CRAN for high dimensional factor analysis and confounder adjusted testing and estimation (author and maintainer).
7. seismic on CRAN for predicting information cascade by self-exciting point process (author and maintainer).

Reviewer

I have reviewed research papers for

1. American Journal of Epidemiology.
2. American Journal of Human Genetics.
3. Annals of Statistics.
4. Arteriosclerosis, Thrombosis, and Vascular Biology.
5. Bioinformatics.
6. Biometrics.
7. Biometrika.
8. Canadian Journal of Statistics.
9. eLife.
10. IEEE Transactions on Knowledge and Data Engineering.
11. IEEE Transactions on Pattern Analysis and Machine Intelligence.
12. International Journal of Data Science and Analytics.
13. Journal of the Americal Statistical Association.
14. Journal of Business & Economic Statistics.
15. Journal of Causal Inference.
16. Journal of Machine Learning Research.
17. Journal of the Royal Statistical Society (Series B, Statistical Methodology).
18. Nature Communications.

19. Observational Studies.
20. Proceedings of the National Academy of Sciences.
21. PLOS ONE.
22. PLOS Genetics.
23. Scandinavian Journal of Statistics.
24. Statistics in Medicine.
25. Statistica Sinica.

Teaching

At University of Cambridge

1. Statistical Modelling (Part II, undergraduate) 2020-2021.
2. Causal Inference (Part III, postgraduate) 2019–2020, 2020-2021.
3. Causal Inference in the Social Sciences (Social Science Research Methods Programme, bite-sized modules), March 2020.

Awards

1. Best Potential Prize in Stanford-Columbia datafest, 2013.
2. Bronze Medal (No.3) in Applied Mathematics and Statistics in S. T. Yau College Student Mathematics Competition, 2010.
3. National Student Scholarship, 2010.
4. Microsoft Research Asia Young Fellowship, 2010.

Previous Employment

Stanford University: Teaching and research assistant. 2011.9 – 2016.6.

Google Inc.: Decision support analyst intern (cloud analytics team, mentor: Sangho Yoon). 2015.6 – 2015.9.

Google Inc.: Decision support analyst intern (advertiser analytics team, mentor: Daniel Percival). 2014.6 – 2014.9.

eBay Inc.: Generative research intern (search ranking team, mentor: Shaji Sebastian). 2013.6 – 2013.9.