

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>

Current positions

Aug 2019– University Lecturer (Assistant Professor) in Statistics, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge.

Feb 2020– Fellow, Corpus Christi College, Cambridge.

Oct 2021– Fellow, The Alan Turing Institute.

May 2021– Associate Faculty, Cambridge Centre for AI in Medicine.

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.

Previous positions

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

Sep 2011–Jun 2016 Teaching and research assistant, Department of Statistics, Stanford University.

Jun 2015–Sep 2015 Decision support analyst intern, Google (cloud analytics team, mentor: Sangho Yoon).

Jun 2014–Sep 2014 Decision support analyst intern, Google (advertiser analytics team, mentor: Daniel Percival).

Jun 2013–Sep 2013 Generative research intern, eBay (search ranking team, mentor: Shaji Sebastian).

Research grants

Sep 2020–Aug 2021 Isaac Newton Trust Early Career Grant (£54,962).

Supervision

PhD students

Oct 2021– Joakim Blach Anderson.

Oct 2020– Tobias Freidling.

Sep 2018– Matt Tudball (co-supervised with Kate Tilling, Rachel Hughes, Jack Bowden, and George Davey Smith).

Postdoctoral fellows

Sep 2020–Aug 2021 Katarzyna Reluga.

Summer undergraduate students

Jul 2021–Sep 2021 Thalia Seale (co-supervised with Alec Christie).

Jul 2021–Sep 2021 Naomi Wei (co-supervised with Jingshu Wang).

Jul 2020–Sep 2020 Etaash Katiyar.

Publications

Peer-reviewed articles: Statistical theory and methodology

1. 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](https://doi.org/10.1515/jci-2016-0010).
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](https://doi.org/10.1214/16-aos1511).
3. 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](https://doi.org/10.1080/01621459.2017.1407770).
4. 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](https://doi.org/10.1097/ede.0000000000000822).
5. 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](https://doi.org/10.1080/01621459.2018.1429277).
6. 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](https://doi.org/10.1214/18-aos1698).
7. 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](https://doi.org/10.1214/18-aos1702).

8. 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](https://doi.org/10.1080/01621459.2018.1497499).
9. 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](https://doi.org/10.1093/ije/dyy258).
10. 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](https://doi.org/10.1214/18-sts692).
11. 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](https://doi.org/10.1097/mlr.0000000000001040).
12. 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](https://doi.org/10.1111/rssb.12327).
13. 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](https://doi.org/10.1093/ije/dyz142).
14. 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](https://doi.org/10.1214/19-aos1866).
15. 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. 116, no. 533, pp. 295–308, 2020. DOI: [10.1080/01621459.2020.1736083](https://doi.org/10.1080/01621459.2020.1736083).
16. 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*, vol. 15, no. 1, pp. 363–390, 2021. DOI: [10.1214/20-aos1401](https://doi.org/10.1214/20-aos1401).
17. M. J. Tudball, J. Bowden, R. A. Hughes, A. Ly, M. R. Munafò, K. Tilling, Q. Zhao, and G. D. Smith, "Mendelian randomisation with coarsened exposures," *Genetic Epidemiology*, vol. 45, no. 3, pp. 338–350, 2021. DOI: [10.1002/gepi.22376](https://doi.org/10.1002/gepi.22376).
18. 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," *Observational Studies*, vol. 7, pp. 1–24, 2021.
19. 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," *PLOS Genetics*, in press, 2021. DOI: [10.1371/journal.pgen.1009575](https://doi.org/10.1371/journal.pgen.1009575).
20. Q. Zhao, L. J. Keele, D. S. Small, and M. M. Joffe, "A note on post-treatment selection in studying racial discrimination in policing," *American Political Science Review*, to appear, 2021. DOI: [10.1017/s0003055421000654](https://doi.org/10.1017/s0003055421000654).

Peer-reviewed articles: Statistical applications

21. 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](https://doi.org/10.1145/2783258.2783401).
22. 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](https://doi.org/10.1097/mlr.0000000000000947).
23. 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. McConaughy, 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](https://doi.org/10.1038/s41467-020-20142-y).
24. M. T. Patrick, P. E. Stuart, H. Zhang, Q. Zhao, X. Yin, K. He, X.-j. Zhou, N. N. Mehta, J. J. Voorhees, M. Boehnke, J. E. Gudjonsson, R. P. Nair, S. K. Handelman, J. T. Elder, D. J. Liu, and L. C. Tsoi, "Causal relationship and shared genetic loci between psoriasis and type 2 diabetes through trans-disease meta-analysis," *Journal of Investigative Dermatology*, vol. 141, no. 6, pp. 1493–1502, 2021. DOI: [10.1016/j.jid.2020.11.025](https://doi.org/10.1016/j.jid.2020.11.025).
25. Q. Zhao, J. Wang, Z. Miao, N. R. Zhang, S. Hennessy, D. S. Small, and D. J. Rader, "A Mendelian randomization study of the role of lipoprotein subfractions in coronary artery disease," *eLife*, vol. 10, e58361, 2021. DOI: [10.7554/eLife.58361](https://doi.org/10.7554/eLife.58361).
26. Q. Zhao, "Small data, big time—a retrospect of the first weeks of COVID-19 (with discussion and rejoinder)," *Journal of the Royal Statistical Society (Series A, Statistics in Society)*, to appear,

Invited contributions/Short letters

27. 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.
28. Zhao Q. & Panigrahi S. (2019). "Selective inference for effect modification: An empirical investigation," *Observational Studies* 5, 131–140.
29. 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](https://doi.org/10.1214/18-sts680).

30. Q. Zhao and T. Hastie, "Causal interpretations of black-box models," *Journal of Business & Economic Statistics*, vol. 39, no. 272-281, pp. 1–10, 2021. DOI: [10.1080/07350015.2019.1624293](https://doi.org/10.1080/07350015.2019.1624293).
31. 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](https://doi.org/10.2807/1560-7917.es.2020.25.29.2001381).
32. Q. Zhao, "Statistical modeling: Returning to its roots," *Observational Studies*, vol. 7, pp. 229–234, 1 2021. DOI: [10.1353/obs.2021.0014](https://doi.org/10.1353/obs.2021.0014).

Preprints

33. Q. Zhao, T. Hastie, and D. Pregibon, "Estimation and prediction in sparse and unbalanced tables," 2017. arXiv: [1703.02081](https://arxiv.org/abs/1703.02081) [stat.CO].
34. Q. Zhao, D. S. Small, and A. Ertefaie, "Selective inference for effect modification via the lasso," 2017. arXiv: [1705.08020](https://arxiv.org/abs/1705.08020) [stat.ME].
35. Y. Song and Q. Zhao, "Performance evaluation with latent factors," 2018. DOI: [10.2139/ssrn.3216272](https://doi.org/10.2139/ssrn.3216272). SSRN: [3216272](https://ssrn.com/abstract=3216272).
36. M. Tudball, Q. Zhao, R. Hughes, K. Tilling, and J. Bowden, "An interval estimation approach to sample selection bias," 2019. arXiv: [1906.10159](https://arxiv.org/abs/1906.10159) [stat.ME].
37. 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](https://doi.org/10.1101/2020.02.06.20020941).
38. 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](https://doi.org/10.1101/2020.02.11.943712).
39. D. Iong, Q. Zhao, and Y. Chen, "A latent mixture model for heterogeneous causal mechanisms in Mendelian randomization," 2020. arXiv: [2007.06476](https://arxiv.org/abs/2007.06476) [stat.AP].
40. T. Ye, J. Shao, Y. Yi, and Q. Zhao, "Toward better practice of covariate adjustment in analyzing randomized clinical trials," 2020. arXiv: [2009.11828](https://arxiv.org/abs/2009.11828) [stat.ME].
41. E. Katiyar and Q. Zhao, "On testing mean proportionality of multivariate normal variables," 2021. arXiv: [2103.05574](https://arxiv.org/abs/2103.05574) [math.ST].
42. Y. Zhang and Q. Zhao, "Multiple conditional randomization tests," 2021. arXiv: [2104.10618](https://arxiv.org/abs/2104.10618) [math.ST].

Talks and Presentations

2015 ACM KDD Conference.

2016 Rutgers University (Department of Statistics); New York University (Department of Information, Operations and Management Science); Joint Statistical Meeting (JSM).

- 2017** Atlantic Causal Inference Conference; University of Bristol (MRC Integrative Epidemiology Unit).
- 2018** University of Minnesota (Department of Statistics); Johns Hopkins University (Department of Biostatistics); UC Berkeley (School of Public Health, Division in Biostatistics); Stanford University (Department of Statistics); Atlantic Causal Inference Conference; International Conference on Econometrics and Statistics (EcoSta); ICSA Applied Statistics Symposium; University of British Columbia (Department of Statistics); Rutgers University (Department of Biostatistics and Epidemiology); Cornell University (Department of Statistics).
- 2019** University of Washington (Department of Biostatistics); University of Southern California (Marshall School); Yale University (Department of Biostatistics); Rutgers University (Department of Statistics); University of Minnesota (Department of Statistics); University of Michigan (Department of Statistics); University of Cambridge (DPMMS); Harvard University (Department of Statistics); University of Wisconsin (Department of Statistics); Carnegie Mellon University (Department of Statistics); Bayesian Causal Inference Workshop, Ohio State University; Banff International Research Station (BIRS) Workshop; WNAR Annual Meeting; CCI Causal Inference Summer Institute, Rutgers University; Mendelian Randomization Conference, University of Bristol; MRC Biostatistics Unit, University of Cambridge; SAMSI workshop in Causal Inference.
- 2020** University of Cambridge (CCIMI Seminar); Yale University (Department of Biostatistics); Joint Statistical Meeting (JSM); Annual Conference of International Society for Clinical Biostatistics (ISCB); Yale University (Department of Biostatistics); Ohio State University (Department of Statistics); Pacific Causal Inference Conference (PCIC); University of Science and Technology of China (Department of Statistics and Finance); Online Causal Inference Seminar.
- 2021** International Selective Inference Seminar; University of Oslo (Center for Lifespan Changes in Brain and Cognition); Infectious disease outbreaks webinar; EPFL (Department of Statistics); University of Pennsylvania (Center for Statistics in Big Data, Open Insights Seminar); RSS Annual Conference (Research Section Discussion Meeting).

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).

Professional activities

Peer review for scientific journals

1. American Journal of Epidemiology.
2. American Journal of Human Genetics.
3. Annals of Applied Statistics.
4. Annals of Statistics.
5. Arteriosclerosis, Thrombosis, and Vascular Biology.
6. Bioinformatics.
7. Biometrics.
8. Biometrika.
9. Canadian Journal of Statistics.
10. Electronic Journal of Statistics.
11. eLife.
12. IEEE Transactions on Knowledge and Data Engineering.
13. IEEE Transactions on Pattern Analysis and Machine Intelligence.
14. International Journal of Data Science and Analytics.
15. Journal of the Americal Statistical Association.
16. Journal of Business & Economic Statistics.
17. Journal of Causal Inference.
18. Journal of Econometrics.
19. Journal of Machine Learning Research.
20. Journal of the Royal Statistical Society (Series A, Statistics in Society).
21. Journal of the Royal Statistical Society (Series B, Statistical Methodology).
22. Nature Communications.
23. Observational Studies.
24. Proceedings of the National Academy of Sciences.
25. PLOS ONE.

26. PLOS Genetics.
27. Scandinavian Journal of Statistics.
28. Statistics in Medicine.
29. Statistica Sinica.

Peer review for grant proposals

1. Singapore Ministry of Education Academic Research Fund.
2. UK Medical Research Council, Better Methods, Better Research Panel.

Other services

Jul 2021– Moderator, Online Causal Inference Seminar.

Teaching

Courses

1. Statistical Modelling (Part II), University of Cambridge, Michaelmas 2020.
2. Causal Inference (Part III), University of Cambridge, Michaelmas 2019, 2020.

Guest lectures

1. Graphical models, Systems Biology Part III, University of Cambridge, January 2021.
2. 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.