Fan Yang

Professor
Yau Mathematical Sciences Center
Tsinghua University
Email: yangfan1987@tsinghua.edu.cn

Education

2005 - 2009   **B.S.**, Mathematics and Applied Mathematics, Zhejiang University

2009 - 2014   **Ph.D.**, Applied Mathematics and Computational Science, University of Pennsylvania

- Advisors: *Dylan Small* and *Edward George*

- Dissertation title: *Causal Inference Methods for Addressing Censoring by Death and Unmeasured Confounding Using Instrumental Variables.*

Academic Appointments and Positions

2014 - 2017   **Assistant Professor**, Department of Public Health Sciences, The Division of Biological Sciences, The University of Chicago

2017 - 2021   **Assistant Professor**, Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus

2021 - 2022   **Associate Professor**, Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus

2022 -   **Professor**, Yau Mathematical Sciences Center, Tsinghua University

Awards

- J. Parker Bursk Memorial Prize, The Wharton School, University of Pennsylvania, 2012

- American Statistical Association (ASA) Section on Statistics in Epidemiology Young Investigators Award, 2013

- ASA Section on Health Policy Statistics Student Paper Competition Award, 2014

- International Genetic Epidemiology Society Robert C. Elston Best Paper Award, 2022

Professional Service

2019   International Chinese Statistical Association (ICSA) Local Committee at 2019 Joint Statistical Meetings

2019   Western North American Region of The International Biometric Society (WNAR) Student Award Committee

| 2021 - 2025 | Associate Editor of *Biometrics* |
| 2022 | WNAR Student Competition Committee Chair |

Teaching

Courses

2023	Sole instructor, Yau Mathematical Sciences Center, Tsinghua University. Course: Mathematical Statistics (Fall 2023)
2023-2024	Sole instructor, Yau Mathematical Sciences Center, Tsinghua University. Course: Statistical Inference (Spring 2023, 2024)
2019, 2022	Sole instructor, Department of Biostatistics and Informatics, University of Colorado Anschutz Medical Campus. Course: BIOS 6641 Causal Analytics in Public Health (Spring 2019; Spring 2022).
2020 - 2021	Sole instructor, Department of Biostatistics and Informatics, University of Colorado Anschutz Medical Campus. Course: BIOS 6632 Master's Statistical Theory II (Spring 2020; Spring 2021).
2019	Co-instructor, Department of Biostatistics and Informatics, University of Colorado Anschutz Medical Campus. Course: BIOS 6990 MPH Capstone Preparation (Fall 2019).
2016 - 2017	Co-instructor, Department of Public Health Sciences, The University of Chicago. Course: CHDV 30102 Causal Inference (Winter 2016; Winter 2017).
2015 - 2017	Sole instructor, Department of Public Health Sciences, The University of Chicago. Course: PBHS 32700 Biostatistical Methods (Winter 2015; Winter 2016; Winter 2017).
2013	Sole instructor, Department of Statistics, The Wharton School, University of Pennsylvania. Course: Introductory Statistics (Summer 2013).

Guest Lectures

| May 2017 | "Truncation by death", Advanced Topics in Causal Inference, The University of Chicago |
| April 2018 | "Truncation by death", Survival Analysis, University of Colorado Anschutz Medical Campus |

Advising
MPH

| 2016 | Poushali Bhattacharjee, "Detecting sepsis: Are two opinions better than one?" My role: Academic advisor in Biostatistics |
| 2017 | Kiran Pandey, "The impact of marital status on health care utilization among Medicare beneficiaries." My role: Academic advisor in Biostatistics |

MS

| 2021 Expected | Stella Veazey, "Counterfactual approximation using local classification boundaries." My role: Thesis committee member |

2020	Vanessa Richardson, "Intermittent data in marginal structural models". My role: Thesis committee member

PhD

2020	Kevin Josey (PhD in Biostatistics). "Doubly-robust covariate balancing weights for causal inference and transportability". My role: Dissertation committee member
2020	Kristen Polinski (PhD in Epidemiology), "The role of lipid mediators in preclinical rheumatoid arthritis". My role: Dissertation committee member
2020	Kevin Gleason (PhD in Biostatistics, University of Chicago), "Statistical methods for analyzing omics and multi-omics quantitative trait loci (QTLs) in multiple conditions". My role: Dissertation committee member
2025 Expected	Shuozhi Zuo (PhD student in Biostatistics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus). My role: Advisor

Administrative Position

2019 - 2020	MPH Program Co-Director, Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus
2020 - 2022	MPH Program Director, Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus
2023 - present	Deputy Director of the Yau Mathematical Sciences Center, Tsinghua University

Research

Grants

- R305D200031 "Causal moderation and mediation analyses with noncompliance in single-site and multi-site randomized trials." Principal Investigator: Hong. IES, 09/01/2020-08/31/2023. My role: Co-PI of grant.

  In this project, we develop analytic strategies that innovatively incorporate an instrumental variable (IV) for identifying and estimating the ATE, the moderated ATE, the mediated ATE, and the moderated mediation effects in analyzing single-site or multi-site experimental data.

- R01GM108711 "Integrative multivariate association and genomic analyses." Principal Investigators: Chen and Yang. NIH, 09/01/2019-05/31/2023. My role: MPI of grant.

  This project is to develop new statistical methods and computational tools in order to characterize the effects of genetic regulation on different molecular phenotypes in different cell contexts, to integrate and recapitalize on existing large data resources, and to create a more comprehensive understanding of how genetic variation impacts biological processes related to complex diseases.

- SES-1659935 "Methods for investigating causal mechanisms in multi-site experimental and quasi-experimental studies." Principal Investigator: Hong. NSF/SES, 07/01/2017-06/30/2021. My role: Co-PI of grant.

  This project is to develop new strategies for investigating causal mechanisms in multi-site experimental and quasi-experimental evaluations of intervention programs.

- CIN13-402 "The Health Services Research and Development Denver-Seattle Center of Innovation for Veteran-Centered and Value-Driven Care." Principal Investigator: Ho. Veterans Administration, 03/01/2019-01/31/2022. My role: Statistician.

  This project is to conduct health services research that advances Veteran-centered and value-driven healthcare.

- Small Research Grant "Sensitivity analysis for causal mediation analysis in the presence of unmeasured pretreatment or posttreatment confounding." Principal Investigator: Qin. Spencer Foundation, 01/01/2019-06/30/2020. My role: Co-PI of grant.

  This project is to develop simulation-based sensitivity analysis strategies to numerically and graphically evaluate the sensitivity of causal mediation analysis results to the presence of an unmeasured pretreatment or posttreatment confounder of the mediator-outcome relationship under treatment randomization.

- R03 CA208387 "Understanding causal effects of a treatment on survival in observational studies with unmeasured confounding and an application to effect of PMRT on breast cancer patients." Principal Investigator: Yang. NIH/NCI, 09/15/2017-02/29/2020. My role: PI of grant.

  This project is to develop a novel instrumental variable method to obtain consistent estimates of causal treatment effects for survival outcomes and to evaluate the effectiveness of PMRT on survival of the AJCC pT1-2pN1 breast cancer patients.

- P30CA046934-30 "University of Colorado Cancer Center Support Grant." Principal Investigator: Schulick. NIH, 02/01/2018-01/31/2019. My role: Statistician.

  The UCCC mission is to discover, develop, and deliver breakthroughs in cancer science through interdisciplinary collaborative research.

Publications
[† Indicates co-first authors contributed equally][‡ Indicates corresponding authors]

1. **Yang, F.**, Lorch, S. A., and Small, D. S. (2014). "Estimation of causal effects using instrumental variables with nonignorable missing covariates: Application to effect of type of delivery hospital on premature infants." *Annals of Applied Statistics*, Volume 8, Number 1, pp. 48-73.

2. **Yang, F.**, Zubizarreta, J., Small, D. S., Lorch, S. A., and Rosenbaum, P. (2014). "Dissonant conclusions when testing the validity of an instrumental variable." *The American Statistician*, Volume 68, Issue 4, pp. 253-263.

3. **Yang, F.** and Small, D. S. (2016). "Using post-outcome measurement information in censoring-by-death problems." *Journal of the Royal Statistical Society, Series B*, Volume 78, Number 1, pp. 299-318.

4. **Yang, F.**, Wang, J., the GTEx Consortium, Pierce, B. L, and Chen, L. S. (2017). "Identifying cis-mediators for trans-eQTLs across many human tissues using genomic mediation analysis." *Genome Research*, Volume 27, pp. 1859-1871.

5. Hong, G., Qin, X., and **Yang, F.** (2018). "Weighting-based sensitivity analysis in causal mediation studies." *Journal of Educational and Behavioral Statistics*, Volume 43, Issue 1, pp. 32 - 56.

6. **Yang, F.** and Ding, P. (2018). "Using survival information in truncation by death problems without the monotonicity assumption." *Biometrics*, Volume 74, Issue 4, pp. 1232-1239.

7. Konetzka, R. T., **Yang, F.**, and Werner, R. M. (2019). "Use of instrumental variables for endogenous treatment at the provider level." *Health Economics*, Volume 28, pp. 710-716.

8. Pandey, K. R., **Yang, F.**, Cagney, K. A., Smieliauskas, F., Meltzer, D. O., and Ruhnke, G. W. (2019). "The impact of marital status on health care utilization among Medicare beneficiaries." *Medicine*, Volume 98, No.12, e14871.

9. Song, X., Ji, J., Gleason, K. J., **Yang, F.**, Martignetti, J. A., Chen, L. S., and Wang, P. (2019). "Insights into impact of DNA copy number alteration and methylation on the proteogenomic landscape of human ovarian cancer via a multi-omics integrative analysis." *Molecular & Cellular Proteomics*, mpc-RA118.

10. **Yang, F.**, Cheng, J., and Huo, D. (2019). "Instrumental variable approach for estimating a causal hazard ratio: application to the effect of postmastectomy radiotherapy on breast cancer patients." *Observational Studies*. Volume 5, pp.141-162.

11. Prouse, A., Gunzburger, E., Morrison, J., **Yang, F.**, Valle, J. A., Armstrong, E., and Waldo, S. W. (2020). "Contemporary utilization and outcomes of arterial closure devices after percutaneous coronary intervention: Insights from the VA Clinical Assessment Reporting and Tracking (CART) program." *Journal of the American Heart Association*. 2020 Feb 18, Volume 9, e015223.

12. Josey, K., Juarez-Colunga, E., **Yang, F.**, and Ghosh, D. (2020). "A framework for covariate balance using Bregman distances." *Scandinavian Journal of Statistics*. pp. 1-27. https://doi.org/10.1111/sjos.12457.

13. Althoff, M. D., Holguin, F., **Yang, F.**, Grunwald, G., Moss, M., Vandivier, R. W., Ho, P. M., Kiser, T. H., and Brunham, E. L. (2020). "Noninvasive ventilation use in critically ill patients with acute asthma exacerbations." *American Journal of Respiratory and Critical Care Medicine*. doi:10.1164/rccm.201910-2021OC.

14. Gleason, K. J.†, **Yang, F.**†, Pierce, B. L., He, X., and Chen, L. S. (2020). "Primo: integration of multiple GWAS and omics QTL summary statistics for elucidation of molecular mechanisms of trait-associated SNPs and detection of pleiotropy in complex traits." *Genome Biology*. Volume 21, 236.

15. Hong, G., **Yang, F.**, and Qin, X. (2021). "Did you conduct a sensitivity analysis? A new weighting-based approach for evaluations of the average treatment effect for the treated." *Journal of the Royal Statistical Society, Series A*. Volume 184, Issue 1, pp. 227-254.

16. Polinski, K. J., Bemis, E., **Yang, F.**, Crume, T., Demoruelle, K., Feser, M., Seifert, J., O'Dell, J. R., Mikuls, T., Weisman, M., Gregersen, P., Keating, R., Buckner, J., Reisdorph, N., Deane, K. D., Clare-Salzler, M., Holers, V., Norris, J. (2021) "Association of circulating lipid mediators and development of future incident inflammatory arthritis in an anti-citrullinated protein antibody positive population." *Arthritis & Rheumatology*. Volume 73, Issue 6, pp. 955-962

17. Gleason, K. J., **Yang, F.**‡, and Chen, L. S.‡ (2021). "A robust two-sample transcriptome-wide Mendelian Randomization method integrating GWAS with multi-tissue eQTL summary statistics." *Genetic Epidemiology*. Volume 45, Issue 4, pp. 353-371.

18. Polinski, K. J., Armstrong, M, Manke, J., Seifert, J., Crume, T., **Yang, F.**, Clare-Salzler, M., Holers, V. M., Reisdorph, N., Norris, J. M.. "Collection and storage of human plasma for measurement of oxylipins." (Accepted by *Metabolites*)

19. **Yang, F.**‡, Gleason, K. J., Wang, J., The GTEx Consortium, Duan, J., He, X, Pierce, B. L., and Chen, L. S.‡ (2021). "CCmed: cross-condition mediation analysis for identifying replicable trans-associations mediated by cis-gene expression." *Bioinformatics*, Volume 37, Number 17, pp. 2513-2520.

20. Cruz Cortés, E., **Yang, F.**, Juárez-Colunga, E., Warsavage, T., Ghosh, D. (2021). Comment on "Statistical Modeling: the Two Cultures" by Leo Breiman. *Observational Studies*, 7(1), 41-57.

21. Qin, X., and **Yang, F.** (2022). "Sensitivity analysis for causal mediation studies: using a simulation strategy." *Psychological Methods*, 27(6), 1000-1013.

22. Gilmartin, H. M., Warsavage, T., Hines, A., Leonard, C., Kelley, L., Wills, A., Gaskin, D., Motta, L. U., Connelly, B., Plomondon, M. E., **Yang, F.**, Kaboli, P., Burke, R. R., Jones, C. D. (2022) "Effectiveness of the Rural Transitions Nurse Program for Veterans: A Multicenter Implementation Study." *Journal of Hospital Medicine*, 17(3):149-157.

23. Josey, K., **Yang, F.**, Ghosh, D., Raghavan, S. (2022) "A calibration approach to transportability with observational data." *Statistics in Medicine*, 41(23):4511-4531

24. Zuo, S., Josey, K. P., Raghavan, S., **Yang, F.**, Juaréz-Colunga, E., and Ghosh, D. (2022). Transportability Methods for Time-to-Event Outcomes: Application in Adjuvant Colon Cancer Trials. *JCO clinical cancer informatics*, 6, e2200088.

25. Hong, G. † , **Yang, F.** †, Qin, X. (2023) "Post-treatment confounding in causal mediation studies: A cutting-edge problem and a novel solution via sensitivity analysis." *Biometrics*, 79(2), 1042-1056.

26. Gao, D., Hu, J., Bradley, C. J., **Yang, F.**‡ (2023) "Instrumental variable analysis for cost outcome: Application to the effect of primary care visit on medical cost among low-income adults." *Statistics in Medicine*, 42(24): 4349-4376.

27. Johnson, R.K., Ireton, A.J., Carry, P.M., Vanderlinden, L.A., Dong, F., Romero, A., Johnson, D.R., Ghosh, D., **Yang, F.**, Frohnert, B. and Yang, I.V., Kechris, K., Rewers, M., and Norris, J. M. (2023). "DNA Methylation Near DLGAP2 May Mediate the Relationship between Family History of Type 1 Diabetes and Type 1 Diabetes Risk." *Pediatric Diabetes*. Volume 2023, Article ID 5367637.

28. **Yang, F.** †‡, Chen, L. S. †‡, Oveisgharan, S., Barbar, D., and Bennett, D. A. (2024) "Integrating Mendelian randomization with causal mediation analysis for characterizing direct and indirect exposure-to-outcome effects." Accepted by *Annals of Applied Statistics*.

29. Zuo, S., Ghosh, D., Ding, P.‡, and **Yang, F.**‡ (2024) "Mediation analysis with the mediator and outcome missing not at random" Accepted by *Journal of the American Statistical Association*.

30. Lu, Y., Xu, K., Maydanchik, N., Kang, B., Pierce, B. L., **Yang, F.**‡, Chen, L. S.‡ (2024) "An integrative multi-context Mendelian randomization method for identifying risk genes across human tissues" Accepted by *American Journal of Human Genetics*.

Manuscripts Under Revision/Review

1. Lu, Y., Pierce, B. L., Wang, P., **Yang, F.**, Chen, L. S. "Alternative splicing induces sample-level variation in gene-gene correlations." (Under Revision at *BMC Bioinformatics*)

2. Cruz Cortés, E., Josey, K., **Yang, F.**, Ghosh, D. "An empirical process framework for covariate balance in causal inference." (Under Review)

R software packages

1. "GMAC": Genomic mediation analysis with adaptive confounding adjustment. Authors: **Yang, F.**, Wang, J., and Chen, L.

2. "primo": Package in R for integrative multi-omics association analysis. Authors: Gleason, K., **Yang, F.**, and Chen, L.

3. "rmpw": Causal mediation analysis using weighting approach. Authors: Qin, X., Hong, G., and **Yang, F.**

4. "CCmed": Cross-condition mediation analysis. Authors: Gleason, K., **Yang, F.**, and Chen, L.

5. "MrRobin": Two-sample Mendelian Randomization method ROBust to correlated and some INvalid instruments. Authors: Gleason, K., **Yang, F.**, and Chen, L.

6. "mediationsens": Simulation-based sensitivity analysis for causal mediation studies. Authors: Qin, X. and **Yang, F.**

Invited Presentations and Presentations at Conferences

1. "Estimation of causal effects using instrumental variables with nonignorable missing covariates: application to effect of type of delivery hospital on premature infants," Joint Statistical Meetings, Montreal, Canada (Topic Contributed Session, Oral), August 2013

2. "Using post-quality of life measurement information in censoring by death problems," Causal Reading Group, University of Pennsylvania, January 2014

3. "Using post-quality of life measurement information in censoring by death problems," Joint Statistical Meetings, Boston, MA (Topic Contributed Session, Oral), August 2014

4. "Do C-sections protect very premature babies? Aporetic conclusions when testing the validity of an instrumental variable," Workshop on Quantitative Methods in Education, Health and the Social Sciences, The University of Chicago, November 2014

5. "Using post-quality of life measurement information in censoring by death problems," Department Seminar, Division of Epidemiology and Biostatistics, University of Illinois at Chicago, January 2015

6. "Using post-quality of life measurement information in censoring by death problems," Department Seminar, Department of Preventive Medicine, Northwestern University, March 2015

7. "Understanding causal effects of a treatment on survival in observational studies with unmeasured confounding," Department Seminar, Division of Biostatistics, Medical College of Wisconsin, November 2016

8. "Using survival information in truncation by death problems without monotonicity assumption," Department Seminar, Department of Statistics, University of Notre Dame, April 2017

9. "Using survival information in truncation by death problems without monotonicity assumption," International Chinese Statistical Association Symposium, Chicago, IL (Invited Session, Oral), June 2017

10. "Estimation of causal effects using instrumental variables with nonignorable missing covariates: application to effect of type of delivery hospital on premature infants," ACCORDS Seminar, University of Colorado Anschutz Medical Campus, November 2017

11. "Using post-quality of life measurement information in censoring by death problems," Data Science to Patient Value Seminar, University of Colorado Anschutz Medical Campus, February 2018

12. "Using survival information in truncation by death problems without monotonicity assumption," Department Seminar, Division of Statistics, Northern Illinois University, April 2018

13. "Did you conduct a sensitivity analysis? A new weighting-based approach in evaluations of the average treatment effect for the treated," Department Seminar, Department of Biostatistics and Informatics, University of Colorado Anschutz Medical Campus, May 2018

14. "Using survival information in truncation by death problems without monotonicity assumption," Department Seminar, Department of Applied Mathematics, University of Colorado Boulder, March 2019

15. "Causal mediation analysis under partial compliance in randomized trials," Society for Research on Educational Effectiveness Spring Conference, Washington, DC (Symposium, Oral), March 2019

16. "Causal mediation analysis under partial compliance in randomized trials," International Chinese Statistical Association Symposium, Raleigh, NC (Invited Session, Oral), June 2019

17. "Integrative genomic association and mediation analysis," Department Seminar, Department of Biostatistics and Bioinformatics, Emory University, December 2019

18. "Integrative genomic association analysis," Joint Statistical Meetings, Virtual (Topic Contributed, Oral), August 2020

19. "Integrative genomic mediation analysis," International Chinese Statistical Association Symposium, Virtual (Invited Session, Oral), December 2020

20. "Post-treatment confounding in causal mediation studies: a cutting-edge problem and a novel solution via sensitivity snalysis," Department Seminar, Department of Biostatistics, University of Pittsburgh, March 2021.

21. "Cross-condition mediation analyses for identifying robust trans-associations mediated by cis-gene expression levels in multiple human brain tissues," ENAR, Baltimore, MD (Invited Session, Oral), March 2021

22. "Cross-condition mediation analyses for identifying robust trans-associations mediated by cis-gene expression levels in multiple human brain tissues," WNAR, Virtual (Invited Session, Oral), June 2021

23. "Post-treatment confounding in causal mediation studies: a cutting-edge problem and a novel solution via sensitivity analysis," JSM, Virtual (Invited Session, Oral), August 2021

24. "Mediation analysis with the mediator and outcome missing not at random," Invited presentation at Banff International Research Station (BIRS) 5-day workshop on "Emerging Challenges for Statistics and Data Sciences: Complex Data with Missingness, Measurement Errors, and High Dimensionality", May 2022.

25. "Mediation analysis with the mediator and outcome missing not at random," Invited presentation at Frontier in Mathematical Science at Tsinghua-Sanya International Mathematics Forum (TSIMF), Dec 2022.

26. "Mediation analysis with the mediator and outcome missing not at random," Department Seminar, Department of Biostatistics, Peking University, April 2023.

27. "Mediation analysis with the mediator and outcome missing not at random," ICBS, Beijing, China, July 2023.

28. "Mediation analysis with the mediator and outcome missing not at random," Joint Statistical Meetings, Toronto, Canada (Topic Contributed Session), August 2023.

29. "Mediation analysis with the mediator and outcome missing not at random," Stanford Online Causal Inference Seminar, Stanford University, February 2024.

30. "Mediation analysis with the mediator and outcome missing not at random," Stanford Online Causal Inference Seminar, Stanford University, February 2024.

31. "Integrating Mendelian randomization with causal mediation anlaysis for characterizing direct and indirect exposure-to-outcome effects," Chinese Association for Applied Statistics Section on Causal Inference, Wuhan, China, March 2024.