Fan Li, PhD

CONTACT INFORMATION

Department of Biostatistics, YSPH
Room 229, 135 College Street
New Haven, CT 06510 USA

Phone: (203) 764-9835
Email: fan.f.li@yale.edu
Website: http://campuspress.yale.edu/fl426/

RESEARCH INTERESTS

Cluster randomized trials; Stepped wedge designs; Observational studies; Causal inference; Estimands; Multilevel data; Epidemiologic methods; Study designs; Semiparametric methods; Mediation analysis, Missing data; Survival analysis; Longitudinal data analysis; Bayesian inference

EDUCATION

Duke University, Durham, North Carolina USA

Ph.D., Biostatistics, May 2019

— Dissertation: Topics and applications of weighting methods in case-control and observational studies.

— Advisors: Fan Li, Ph.D. and Andrew S. Allen, Ph.D.

M.S., Biostatistics, May 2014

ACADEMIC APPOINTMENTS

Department of Biostatistics, Yale University School of Public Health (YSPH)
New Haven, Connecticut USA
Assistant Professor July 2019 - Present

Center for Methods in Implementation and Preventive Science (CMIPS)
New Haven, Connecticut USA
Faculty Member July 2019 - Present

Yale Center for Analytical Sciences (YCAS)
New Haven, Connecticut USA
Affiliated Faculty Member January 2020 - Present

Center for Interdisciplinary Research on AIDS (CIRA)
New Haven, Connecticut USA
Affiliated Faculty Member August 2022 - Present

Data Management and Statistics Core, Alzheimer's Disease Research Center (ADRC)
New Haven, Connecticut USA
Affiliated Faculty Member January 2021 - Present

Yale Research Initiative on Innovation and Scale (Y-RISE)
New Haven, Connecticut USA
Affiliated Faculty Member March 2020 - Present

Outcomes Research Group, Duke Clinical Research Institute (DCRI)
Durham, North Carolina USA
Graduate Research Assistant September 2016 - May 2019

PEER-REVIEWED PUBLICATIONS

"$\star$" indicates mentored graduate student or postdoctoral associate

1. **Li F**, Kasza J, Turner EL, Rathouz PJ, Forbes AB, Preisser JS (2022). Generalizing the information content for stepped wedge designs: A marginal modeling approach. *Scandinavian Journal of Statistics*. Accepted.

2. **Li F**, Tian Z*, Tian Z*, Li F (2022). A note on identification of causal effects in cluster randomized trials with post-randomization selection bias. *Communications in Statistics - Theory and Methods*. Accepted.

3. Blaha O*, Esserman D, **Li F** (2022). Design and analysis of cluster randomized trials with survival outcomes under the additive hazards mixed model. *Statistics in Medicine*. http://doi.org/10.1002/sim.9541.

4. **Li F**, Chen X, Tian Z*, Esserman DA, Heagerty PJ, Wang R (2022). Designing three-level cluster randomized trials to assess treatment effect heterogeneity. *Biostatistics*. https://doi.org/10.1093/biostatistics/kxac026.

5. Kahan B, **Li F**, Copas AJ, Harhay MO (2022). Estimands in cluster-randomised trials: choosing analyses which answer the right question. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyac131.

6. Gallis J, Wang X*, Rathouz PJ, Preisser JS, **Li F**, Turner EL (2022). `power swgee`: GEE-based power calculation in stepped wedge cluster randomized trials. *The Stata Journal*. Accepted.

7. Ouyang Y, **Li F**, Preisser JS, Taljaard M (2022). Sample size calculators for planning stepped-wedge cluster randomized trials: A review and comparison. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyac123.

8. Yang S*, Moerbeek M, Taljaard M, **Li F** (2022). Power analysis for cluster randomized trials with continuous co-primary endpoints. *Biometrics*. https://doi.org/10.1111/biom.13692 (Winner of 2022 ENAR Distinguished Student Paper Award)

9. Gettel CJ, Yiadom MYAB, Bernstein SL, Grudzen C, Nath B, **Li F**, Hwang U, Hess EP, Melnick ER (2022). Pragmatic clinical trial design in emergency medicine: Study considerations and design types. *Academic Emergency Medicine*. https://doi.org/10.1111/acem.14513.

10. Chen X, Harhay MO, **Li F** (2022). Clustered restricted mean survival time regression. *Biometrical Journal*. https://doi.org/10.1002/bimj.202200002

11. Wang Y*, **Li F**, Blaha O*, Meng C, Esserman DA (2022). Design and analysis of partially randomized preference trials with propensity score stratification. *Statistical Methods in Medical Research*. https://doi.org/10.1177/09622802221095673.

12. Nicholls SG, Carroll K, Nix HP, **Li F**, Hey SP, Mitchell SL, Weijer C, Taljaard M (2022). Ethical considerations within pragmatic randomized controlled trials in dementia: Results from a literature survey. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*. 8(1), e12287. https://doi.org/10.1002/trc2.12287.

13. Ghazi L*, **Li F**, Chen X, Simonov M, Yamamoto Y, Biswas A, Hanna J, Shah T, Peixoto A, Wilson FP (2022). Blood pressure response to commonly administered antihypertensives for severe inpatient hypertension. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0265497.

14. Chen X, **Li F** (2022). Finite-sample adjustments in variance estimators for clustered competing risks regression. *Statistics in Medicine*. 41(14), 2645-2664. https://doi.org/10.1002/sim.9375.

15. **Li F**, Lu W*, Wang Y*, Pan Z*, Greene EJ, Meng G*, Meng C, Blaha O*, Zhao Y, Peduzzi PN, Esserman DA (2022). A comparison of analytical strategies for cluster randomized trials with

survival outcomes in the presence of competing risks. *Statistical Methods in Medical Research*. 31(7), 1224-1241. https://doi.org/10.1177/09622802221085080.

16. Zhou T*, Tong G, Li F, Thomas LE, **Li F** (2022). `PSweight`: An R package for propensity score weighting analysis. *The R Journal*. https://arxiv.org/abs/2010.08893

17. **Li F**, Buchanan AB, Cole SR (2022). Generalizing trial evidence to target populations in non-nested designs: Applications to AIDS clinical trials. *Journal of the Royal Statistical Society: Series C*. 71(3), 669-697. https://doi.org/10.1111/rssc.12550

18. Cheng C*, Spiegelman D, **Li F** (2022). Is the product method more efficient than the difference method for mediation analysis? *American Journal of Epidemiology*. Accepted.

19. Wang X*, Turner EL, **Li F**, Wang R, Moyer J, Cook AJ, Murray DM, Heagerty PJ (2022). Two weights make a wrong: Cluster randomized trials with variable cluster sizes and heterogeneous treatment effects. *Contemporary Clinical Trials*. https://doi.org/10.1016/j.cct.2022.106702

20. Zhou Y, Turner EL, Simmons R, **Li F** (2022). Constrained randomization and statistical inference for multi-arm parallel cluster randomized controlled trials. *Statistics in Medicine*. 40(10), 1862-1883. https://doi.org/10.1002/sim.9333

21. Tian Z*, Esserman DA, Tong G, Blaha O*, Dziura J, Peduzzi PN, **Li F** (2022). Sample size estimation in hierarchical 2×2 factorial trials accounting for unequal cluster sizes. *Statistics in Medicine*. https://doi.org/10.1002/sim.9284

22. Ghazi L*, **Li F**, Chen X, Simonov M, Yamamoto Y, Biswas A, Hanna J, Shah T, Peixoto A, Wilson FP (2022). Severe inpatient hypertension prevalence and blood pressure response to antihypertensive treatment. *The Journal of Clinical Hypertension*. 24(3), 339-349. https://doi.org/10.1111/jch.14431

23. Chen X*, Chang J, Spielgelman D, **Li F** (2022). A Bayesian approach for estimating the partial population impact fraction with exposure measurement error under a main study/internal validation design. *Statistical Methods in Medical Research*. 31(3), 404-418. https://doi.org/10.1177/09622802211060514

24. Wang X*, Turner EL, Preisser JS, **Li F** (2022). Power considerations for generalized estimating equations analyses of four-level cluster randomized trials. *Biometrical Journal*. 64(4), 663-680. https://doi.org/10.1002/bimj.202100081

25. Chen J*, Xin Z, **Li F**, Spiegelman D (2022). `swdpwr`: A SAS macro and an R package for power calculation in stepped wedge cluster randomized trials. *Computer Methods and Programs in Biomedicine*. 213(106522), 1-19. https://www.sciencedirect.com/science/article/pii/S0169260721005964

26. Tian Z*, Preisser JS, Esserman DA, Turner EL, Rathouz PJ, **Li F** (2022). Impact of unequal cluster sizes for GEE analyses of stepped wedge cluster randomized trials with binary outcomes. *Biometrical Journal*. 64(3), 419-439. https://doi.org/10.1002/bimj.202100112

27. Tong G, Esserman DA, **Li F** (2022). Accounting for unequal cluster sizes in designing cluster randomized trials to detect treatment effect heterogeneity. *Statistics in Medicine*. 41(8), 1376-1396. https://doi.org/10.1002/sim.9283

28. **Li F**, Yu H*, Rathouz PJ, Turner EL, Preisser JS (2021). Marginal modeling of cluster-period means and intraclass correlations in stepped wedge designs with binary outcomes. *Biostatistics*.

23(3), 772-788. https://doi.org/10.1093/biostatistics/kxaa056

29. Zeng S*, Li F, Hu L, **Li F** (2021). Propensity score weighting analysis of survival outcomes using pseudo-observations. *Statistica Sinica*. https://arxiv.org/abs/2103.00605

30. **Li F**, Hong H, Stuart EA (2021). A note on semiparametric efficient generalization of causal effects from randomized trials to target populations. *Communications in Statistics - Theory and Methods*. https://doi.org/10.1080/03610926.2021.2020291

31. Cheng C*, Li F, Thomas LE, **Li F** (2021). Addressing extreme propensity scores in estimating counterfactual survival functions via the overlap weights. *American Journal of Epidemiology*. https://arxiv.org/abs/2108.04394

32. Luo J, Tang X, **Li F**, Wen H, Wang L, Ge S, Tang C, Xu N, Lu L (2021). Cigarette smoking and risk of different pathologic types of stroke: A systematic review and dose-response meta-analysis. *Frontiers in Neurology*.

33. Davis-Plourde K*, Taljaard M, **Li F** (2021). Sample size considerations for stepped wedge designs with subclusters. *Biometrics*. https://doi.org/10.1111/biom.13596 (Senior-author Paper Recognized by the 2022 Early Career Investigator Research Award at Yale School of Public Health)

34. **Li F**, Wang R (2021). Stepped wedge cluster randomized trials: A methodological overview. *World Neurosurgery*. 161, 323-330. https://doi.org/10.1016/j.wneu.2021.10.136

35. Li F, Tian Z*, Bobb JF, Papadogeorgou G, **Li F** (2021). Clarifying selection bias in cluster randomized trials. *Clinical Trials*. https://doi.org/10.1177/17407745211056875

36. Cheng C*, Spiegelman D, **Li F** (2021). Estimating natural indirect effect and mediation proportion via the product method. *BMC Medical Research Methodology*. 21, 1-20. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8606099/

37. Tong G, Seal K, Becker W, **Li F**, Dziura J, Peduzzi PN, Esserman DA (2021). Impact of complex, partially nested clustering in a three-arm individually randomized group treatment trial: A case study with the wHOPE trial. *Clinical Trials*. https://journals.sagepub.com/doi/10.1177/17407745211051288

38. Taljaard M, **Li F**, Qin B*, Cui C*, Zhang L*, Nicholls SG, Carroll K, Mitchell SL (2021). Methodological challenges in pragmatic trials in Alzheimer's disease and related dementias: Opportunities for improvement. *Clinical Trials*. https://doi.org/10.1177/17407745211046672

39. Yang S*, Li F, Thomas LE, **Li F** (2021). Covariate adjustment in subgroup analyses of randomized clinical trials: A propensity score approach. *Clinical Trials*. 18(5), 570-581. https://doi.org/10.1177/17407745211028588. (Finalist of 2021 Society of Clinical Trials Thomas C. Chalmers Student Scholarship)

40. Hu L, Ji J, **Li F** (2021). Estimating heterogeneous survival treatment effect in observational data using machine learning. *Statistics in Medicine*. 40(21), 4691-4713. https://doi.org/10.1002/sim.9090

41. **Li F**, Tong G (2021). Sample size estimation for modified Poisson analysis of cluster randomized trials with a binary outcome. *Statistical Methods in Medical Research*. 30(5), 1288-1305. https://doi.org/10.1177/0962280221990415

42. Wang Y, Jiang Y, Wei D, Singh P, Yu Y, Lee T, Zhang L, Mandl HK, Piotrowski-Daspit AS, Chen X*, **Li F**, Li X, Cheng Y, Josowitz A, Yang F, Zhao Y, Wang F, Zhao Z, Huttner A, Bindra RS, Xiao H, Saltzman WM (2021). Nanoparticle-mediated convection-enhanced delivery of a DNA intercalator to gliomas circumvents temozolomide resistance. *Nature Biomedical Engineering*. 5, 1048-1058 https://doi.org/10.1038/s41551-021-00728-7

43. Nugent J, Aklilu A, Yamamoto Y, Simonov M, **Li F**, Biswas A, Ghazi L, Greenberg J, Mansour S, Moledina D, Wilson FP (2021). Assessment of acute kidney injury and longitudinal kidney function after hospital discharge among patients with and without COVID-19. *JAMA Network Open*. 4(3), e211095.

44. **Li F**, Tong G (2021). Sample size and power considerations for cluster randomized trials with count outcomes subject to right truncation. *Biometrical Journal*. 63(5), 1052-1071. https://doi.org/10.1002/bimj.202000230

45. Wegienka G, Stewart EA, Nicholson WK, Zhang S, **Li F**, Thomas L, Spies JB, Venable S, Laughlin-Tommaso S, Diamond MP, Anchan RM, Maxwell GL, Marsh EE, Myers ER, Vines AI, Wise LA, Wallace K, Jacoby VL (2021). Black women are more likely than White women to schedule a uterine-sparing treatment for Leiomyomas. *Journal of Women's Health*. 30(3), 355-366.

46. **Li F**, Hughes JP, Hemming K, Taljaard M, Melnick ER, Heagerty PJ (2021). Mixed-effects models for the design and analysis of stepped wedge cluster randomized trials: An overview. *Statistical Methods in Medical Research*. 30(2), 612-639. https://doi.org/10.1177/0962280220932962

47. Zeng S*, **Li F**, Wang R, Li F (2021). Propensity score weighting for covariate adjustment in randomized clinical trials. *Statistics in Medicine*. 40(4), 842-858. https://doi.org/10.1002/sim.8805

48. Shung D, Tsay C, Laine L, Chang D, **Li F**, Thomas P, Partridge C, Simonov M, Hsiao A, Tay KJ, Taylor A (2021). Early identification of patients with acute gastrointestinal bleeding using natural language processing and decision rules. *Journal of Gastroenterology and Hepatology*. 36(6), 1590-1597.

49. Wilson FP, Martin M, Yamamoto Y, Partridge C, Moreira E, Arora T, Biswas A, Feldman H, Garg AX, Greenberg JH, Hinchcliff M, Latham S, **Li F**, Lin H, Mansour S, Moledina D, Palevsky PM, Parikh CR, Simonov M, Testani J, Ugwuowo U (2021). Electronic health record alerts for acute kidney injury: A multi-center randomized clinical trial. *BMJ*. 372, m4786.

50. Yu H*, Tong G, **Li F** (2020). A note on the estimation and inference with quadratic inference functions for correlated outcomes. *Communications in Statistics - Simulation and Computation*. https://doi.org/10.1080/03610918.2020.1805463

51. Yang S*, **Li F**, Starks MA, Hernandez AF, Mentz RJ, Choudhury KR (2020). Sample size requirements for detecting treatment effect heterogeneity in cluster randomized trials. *Statistics in Medicine*. 39(28), 4218-4237. https://doi.org/10.1002/sim.8721

52. Yu H*, **Li F**, Turner EL (2020). An evaluation of quadratic inference functions for estimating intervention effects in cluster randomized trials. *Contemporary Clinical Trials Communications*. 19:100605. https://doi.org/10.1016/j.conctc.2020.100605

53. Lu L†, **Li F**†, Wen H†, Ge S, Zeng J, Luo W, Wang L, Tang C, Xu N (2020). An evidence mapping and analysis of registered COVID-19 clinical trials in China. *BMC Medicine*. 18(167), 1-10. (†Equal authorship contributions)

54. **Li F**, Allen AS (2020). Secondary analysis of case-control association studies: Insights on weighting-based inference motivate a new specification test. *Statistics in Medicine*. 39(22), 2869-2882. https://doi.org/10.1002/sim.8579

55. Allore H, Goldfeld K, Gutman R, **Li F**, Monin J, Taljaard M, Travison T (2020). Statistical considerations for embedded pragmatic clinical trials in people living with dementia. *Journal of the American Geriatrics Society*. 68(S2), S68-S73. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7396162/

56. Gallis JA, **Li F**, Turner EL (2020). `xtgeebcv`: A command for bias-corrected sandwich variance estimation for GEE analyses of cluster randomized trials. *The Stata Journal*. 20(2), 363-381. https://doi.org/10.1177/1536867X20931001

57. **Li F** (2020). Design and analysis considerations for cohort stepped wedge cluster randomized trials with a decay correlation structure. *Statistics in Medicine*. 39(4), 438-455. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7027591/

58. Majoros WH, Barrera A, Kim Y-S, **Li F**, Wang X, Cunningham SJ, Johnson GD, Guo C, Lowe WL, Scholtens DM, Hayes MG, Reddy TE, Allen AS (2020). Bayesian estimation of genetic regulatory effects in high-throughput reporter assays. *Bioinformatics*. 36(2), 331-338. https://doi.org/10.1093/bioinformatics/btz545

59. Turner EL, Yao L, **Li F**, Prague M (2020). Properties and pitfalls of weighting as an alternative to multilevel multiple imputation in cluster randomized trials with missing binary outcomes under covariate-dependent missingness. *Statistical Methods in Medical Research*. 29(5), 1338-1353. https://doi.org/10.1177/0962280219859915

60. Logan J, Bauer C, Ke J, Xu H, **Li F** (2020). Models for small area estimation for Census Tracts. *Geographical Analysis*. 52(3), 325-350. https://doi.org/10.1111/gean.12215

61. **Li F**, Li F (2019). Propensity score weighting for causal inference with multiple treatments. *The Annals of Applied Statistics*. 13(4), 2389-2415. https://projecteuclid.org/euclid.aoas/1574910049 (Winner of 2019 JSM Student Paper Award, ASA Biometrics Section)

62. **Li F**, Li K, C Li, Luo S (2019). Predicting the risk of Huntington's disease with multiple longitudinal biomarkers. *Journal of Huntington's Disease*. 8(3), 323-332. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6718328/

63. Yu H, **Li F**, Gallis JA, Turner EL (2019). `cvcrand`: A package for covariate-constrained randomization and the clustered permutation test for cluster randomization trials. *The R Journal*. 11(2), 191-204.

64. **Li F**, Li F (2019). Double-robust estimation in difference-in-differences with an application to traffic safety evaluation. *Observational Studies*. 5, 1-20.

65. **Li F**, Forbes AB, Turner EL, Preisser JS (2019). Power and sample size requirements for GEE analyses of cluster randomized crossover trials. *Statistics in Medicine*. 38(4), 636-649. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6461037/

66. **Li F**, Thomas LE, Li F (2019). Addressing extreme propensity scores via the overlap weights. *American Journal of Epidemiology*. 188(1), 250-257. https://doi.org/10.1093/aje/kwy201

67. **Li F**, Turner EL, Preisser JS (2018). Sample size determination for GEE analyses of stepped wedge cluster randomized trials. *Biometrics*. 74(4), 1450-1458. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6461045/ (Winner of 2018 ENAR Distinguished Student Paper Award)

68. Logan J, Foster A, Ke J, **Li F** (2018). The uptick in income segregation: Real trend or random sampling variation? *American Journal of Sociology*. 124(1), 185-222. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7382782/

69. Gallis JA, **Li F**, Yu H, Turner EL (2018). `cvcrand` and `cptest`: Commands for efficient design and analysis of cluster randomized trials using constrained randomization and permutation tests. *The Stata Journal*. 18(2), 357-378. https://doi.org/10.1177/1536867X1801800204

70. **Li F**, Turner EL, Preisser JS (2018). Optimal allocation of clusters in cohort stepped wedge designs. *Statistics and Probability Letters*. 137, 257-263. https://doi.org/10.1016/j.spl.2018.02.002

71. Zinszer B, Rolotti S, **Li F**, Li P (2017). Bayesian word learning in multiple language environments. *Cognitive Science*. 42(S2), 439-462. https://doi.org/10.1111/cogs.12567

72. Assel MJ†, **Li F**†, Wang Y†, Allen AS, Baggerly KA, Vickers AJ (2017). Genetic polymorphisms of CFH and ARMS2 do not predict response to antioxidants and zinc in patients with Age-Related Macular Degeneration: Independent statistical evaluations of data from the Age-Related Eye Disease Study. *Ophthalmology*. 125(3), 391-397. (†Co-first authors)

73. **Li F**, Turner EL, Heagerty PJ, Murray DM, Vollmer W, DeLong ER (2017). An evaluation of constrained randomization for the design and analysis of group-randomized trials with binary outcomes. *Statistics in Medicine*. 36(24), 3791-3806. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5624845/

74. Brennan JM, Thomas L, Cohen DJ, Shahian S, Wang A, Mack MJ, Holmes DR, Edwards FH, Frankel NZ, Baron SJ, Carroll J, Thourani V, Tuzcu EM, Arnold SV, Cohn R, Maser T, Schawe B, Strong S, Stickfort A, Patrick-Lake E, Graham FL, Dai D, **Li F**, Matsouaka RA, O'Brien S, Li F, Pencina MJ, Peterson ED (2017). Transcatheter versus surgical aortic valve replacement: propensity-matched comparison. *Journal of the American College of Cardiology*. 70(4), 439-450.

75. Turner EL, Prague M, Gallis J, **Li F**, Murray DM (2017). Review of recent methodological developments in group-randomized trials: Part 2–Analysis. *American Journal of Public Health*. 107(7), 1078-1086. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5463203/

76. Turner EL, **Li F**, Gallis J, Prague M, Murray DM (2017). Review of recent methodological developments in group-randomized trials: Part 1–Design. *American Journal of Public Health*. 107(6), 907-915. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5425852/

77. Benjamin Neelon SE, Mayhew M, O'Neill JR, Neelon B, **Li F**, Pate RR (2016). Comparative evaluation of a South Carolina policy to improve nutrition in child care. *Journal of the Academy of Nutrition and Dietetics*. 116(6), 949-956.

78. **Li F**, Lokhnygina Y, Murray DM, Heagerty PJ, DeLong ER (2016). An evaluation of constrained randomization for the design and analysis of group-randomized trials. *Statistics in Medicine*. 35(10), 1565-1579. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4826850/

79. Neelon B, **Li F**, Burgette LF, Benjamin Neelon SE (2015). A spatiotemporal quantile regression model for emergency department expenditures. *Statistics in Medicine*. 34(17), 2559-2575. https:

//doi.org/10.1002/sim.6480 (Winner of 2015 ICHPS Student Travel Award)

Commentaries

80. Li F, **Li F** (2022). Invited Commentary: Using propensity scores for racial disparity analysis. *Observational Studies*.

81. **Li F**, Harhay MO (2020). Commentary: Right truncation in cluster randomized trials can attenuate the power of a marginal analysis. *International Journal of Epidemiology*. 49(3), 964-967. https://doi.org/10.1093/ije/dyaa037

82. **Li F** (2020). Comment: Stabilizing the doubly-robust estimators of the average treatment effect under positivity violations. *Statistical Science*. 35(3), 503-510. https://projecteuclid.org/euclid.ss/1599789710

Book Chapters

83. Tong G, **Li F**, Allen AS (2020). Missing data. In Steven Piantadosi and Curtis Meinert (eds.) *Principles and Practice of Clinical Trials*. Chapter 6, Section 7. Springer Nature Switzerland.

Other Publications

84. Wang R, **Li F**, Cook AJ, DeLong ER (2019). Missing data and intention-to-treat analyses. *NIH Collaboratory Living Textbook of Pragmatic Clinical Trials.*

85. Cook AJ, **Li F**, Murray DM, DeLong ER (2018). Alternative cluster randomized designs. *NIH Collaboratory Living Textbook of Pragmatic Clinical Trials.*

Statistical Software

1. `CoxBcv`: R package to implement bias-corrected sandwich variance estimators under the marginal Cox proportional hazards regression model for clustered time-to-event analysis (on CRAN)

2. `crrcbcv`: R package to implement bias-corrected sandwich variance estimators under the marginal proportional subdistribution hazards regression model for clustered competing risks regression (on CRAN)

3. `power swgee`: Embedded command for power calculation in stepped wedge cluster randomized trials based on marginal models and multilevel correlation structures (on SSC archive)

4. `H2x2Factorial`: R package to calculate required sample size and power for hierarchical 2×2 factorial trials with a cluster-level intervention and an individual-level intervention (on CRAN)

5. `mediateP`: R package to calculate the point and interval estimates of the natural indirect effect, total effect, and mediation proportion, based on the product approach (on CRAN)

6. `geeCRT`: R package to perform bias-corrected estimation and inference on mean and correlation parameters using generalized estimating equations in parallel and stepped wedge cluster randomized trials (on CRAN)

7. `swdpwr`: R package to perform power calculation for stepped wedge cluster randomized trials based on conditional and marginal models with continuous and binary outcomes (on CRAN)

8. `PSweight`: R package to implement a class of propensity score weighting analyses for randomized trials and observational studies (on CRAN)

9. `cvcrand`: R package to implement covariate-constrained randomization and permutation test for cluster randomized trials (on CRAN)

10. `cvcrand` and `cptest`: Stata module to implement covariate-constrained randomization and permutation inference in cluster randomized trials (on SSC archive)

11. `xtgeebcv`: Stata module to compute finite-sample bias-corrected sandwich standard errors for generalized estimating equations with clustered outcomes (on SSC archive)

Active Funding

1R01HL159077-01A1 (PI: Hu) 04/01/2022-3/31/2027
National Heart, Lung, and Blood Institute (NHLBI)
Bayesian Machine Learning for Causal Inference with Incomplete Longitudinal Covariates and Censored Survival Outcomes
The major goals of this project are to develop a suite of novel Bayesian machine learning methods to address challenges posed by complex longitudinal data with censored survival outcomes, including methods accounting for missing exposure data, causal inference with dynamic treatment regimens and developing strategies for assessing the impact of violations of sequential ignorability assumption.
Role: Subaward PI

ME-2021C2-23685 (PI: Hu) 07/01/2022-06/30/2025
Patient-Centered Outcomes Research Institute (PCORI)
Robust Longitudinal Causal Inference Methods with Machine Learning
This project seeks to fill in critical methodological gaps in longitudinal causal inference research and to address the important patient-centered outcome research questions. We propose new methods to estimate the causal effects of time-varying treatments on different quantiles of outcomes and on censored survival outcomes, and further contribute a novel sensitivity analysis approach for addressing longitudinal unmeasured confounding.
Role: Subaward PI

ME-2020C3-21072 (PI: Li) 11/01/2021-10/31/2024
Patient-Centered Outcomes Research Institute (PCORI)
New Methods for Planning Cluster Randomized Trials to Detect Treatment Effect Heterogeneity
The objective of this research is to develop rigorous statistical methods, guidance, and software for planning cluster randomized trials to ensure sufficient power for confirmatory heterogeneous treatment effect analysis, addressing a critical methodological gap in designing definitive pragmatic trials.
Project Budget: $1,069,312
Role: Principal Investigator

UL1 TR001863 (PI: Smith) 06/01/2021-05/31/2026
NCATS
Yale Clinical and Translational Science Award
The Clinical and Translational Science Awards (CTSA) Program supports a national network of medical research institutions — called hubs — that work together to improve the translational research process to get treatments to patients more quickly. The hubs collaborate locally and regionally to catalyze innovation in training, research tools and processes. This award funds the Yale Center for Clinical Investigation (YCCI) as a member of the consortium.
Role: Methodology Core Member

ME-2020C1-19220 (PI: Harhay) 04/01/2021-03/31/2024
Patient-Centered Outcomes Research Institute (PCORI)
Developing Methods to Analyze Data Missing Due to Death in Cluster Randomized Trials
The objective of this project is to create new statistical methods, user-friendly statistical software, and methodological guidance that incorporates stakeholder views of desirable qualities of competing approaches to deal with patient-centered outcome data that are missing due to death in cluster-

randomized trials.
Role: Subaward PI

ME-2019C1-16196 (PI: Preisser) 04/01/2020-09/30/2022
Patient-Centered Outcomes Research Institute (PCORI)
Developing Methods for Conducting Incomplete Stepped-Wedge Cluster Randomized Trials
The objective of this project is to formally quantify trade-offs involved when logistical, resource and patient-centered considerations are balanced against methodological implications in cluster-randomized trials employing incomplete stepped-wedge designs in health intervention research.
Role: Subaward PI

ME-2019C1-16146 (PI: Li) 01/01/2020-06/30/2022
Patient-Centered Outcomes Research Institute (PCORI)
Improving Methods for Addressing Post-Randomization Selection Bias in Cluster Randomized Trials
The objective of this proposal is to develop new statistical methods based on principal stratification to handle bias from intermediate variables in cluster randomized trials.
Role: Subaward PI

U54AG063546-01 (MPIs: Mor and Mitchell) 09/01/2019-06/30/2024
National Institute of Aging (NIA)
NIA AD/ADRD Health Care Systems Research Collaboratory
The NIA AD/ADRD Collaboratory will provide the national infrastructure necessary to catalyze and support embedded pragmatic clinical trials of non-pharmacological interventions for persons with dementia. By convening national experts to provide consultation and guidance to Collaboratory-funded pilot projects and NIA-funded trials, the Collaboratory has the potential to transform care delivery, quality, and outcomes for millions of Americans suffering with AD/ADRD.
Role: Design and Statistics Core, Executive Committee Member

Past Funding

Final Research Report Peer Review Management (PI: Helfand) 01/01/2021-12/31/2021
Patient-Centered Outcomes Research Institute (PCORI)/Oregon Health and Science University
PCORI Peer Review Project
The objective is to serve as an internal statistician and statistical peer reviewer in the review process for PCORI Primary Research Projects.
Role: Subaward PI

NIA IMPACT Collaboratory Pilot Award (PI: Juthani-Mehta) 04/01/2021-06/31/2021
National Institute of Aging (NIA)/Brown University
Electronic Consultation for AD/ADRD Residents Experiencing Infectious Diseases (eCARE-ID)
The objective of the study is to pilot test and establish the feasibility of an embedded pragmatic cluster randomized controlled trial to reduce the duration of antibiotic therapy and number of antibiotic prescriptions among nursing home residents with Alzheimer's disease (AD) and AD-related dementias (ADRD).
Role: Biostatistician

UL1 TR000142 (PI: Smith) 01/01/2020-05/31/2021
NCATS
Yale Clinical and Translational Science Award
The Clinical and Translational Science Awards (CTSA) Program supports a national network of medical research institutions — called hubs — that work together to improve the translational research process to get treatments to patients more quickly. The hubs collaborate locally and regionally to catalyze innovation in training, research tools and processes. This award funds the

Yale Center for Clinical Investigation (YCCI) as a member of the consortium.
Role: Methodology Core Member

3-UH3-DA047003-02S2 (MPIs: D'Onofrio and Melnick) 07/01/2019-03/31/2020
National Institute on Drug Abuse (NIDA)
Administrative Supplement to Pragmatic trial of user-centered clinical decision support to implement EMergency department-initiated BuprenorphinE for opioid use Disorder (EMBED)
The objective of this project is to review and compare analytical approaches based on mixed effects models for the design and analysis of pragmatic stepped wedge cluster randomized trials.
Role: Co-Investigator

Invited Talks

1. "Clarifying Selection Bias in Cluster Randomized Trials". Joint Statistical Meetings. Washington, DC, August 2022.

2. "Impact of Unequal Cluster Sizes for GEE Analyses of Stepped Wedge Cluster Randomized Trials with Binary Outcomes". Society for Clinical Trials 43th Annual Meeting, San Diego, US, May 2022.

3. "Power Analysis for Cluster Randomized Trials with Multiple Continuous Co-Primary Endpoints". Australian Trials Methodology Conference, December 2021.

4. "Propensity Score Weighting Analysis of Survival Outcomes using Pseudo-Observations". Biostatistics Seminar, Department of Epidemiology, Biostatistics and Occupational Health, McGill University, November 2021.

5. "Methods for Designing Cluster Randomized Trials to Detect Treatment Effect Heterogeneity". (Invited Plenary Talk) Current Developments in Cluster Randomised Trials and Stepped Wedge Designs, November 2021.

6. "Propensity Score Weighting Analysis of Survival Outcomes using Pseudo-Observations". ICSA Applied Statistics Symposium, September 2021.

7. "Recent Developments in Statistical Methods for Pragmatic, Stepped Wedge Cluster Randomized Trials". (Invited Plenary Talk) Colorado Pragmatic Research in Health Conference, May 2021.

8. "Marginal Modeling of Cluster-Period Means and Intraclass Correlations in Stepped Wedge Designs with Binary Outcomes". Third International Conference on Stepped Wedge Trial Design. Online Conference, March 2021.

9. "Propensity Score Weighting for Covariate Adjustment in Randomized Clinical Trials". Online Causal Inference Seminar, November 2020.

10. "Propensity Score Weighting for Covariate Adjustment in Randomized Clinical Trials". Center for Causal Inference, University of Pennsylvania, PA, October 2020.

11. "Overview of Statistical Models for the Design and Analysis of Stepped Wedge Cluster Randomized Trials." Methods: Mind the Gap Webinar Series, NIH Office of Disease Prevention, July 2020.

12. "Propensity Score Weighting for Causal Inference with Multiple Treatments - Concept, Extension and Implementation". ENAR Virtual Spring Meeting, March 2020.

13. "Sample Size Determination for GEE Analyses of Stepped Wedge Cluster Randomized Trials". The International Biometric Society Journal Club, April 2019.

14. "Augmented Weighting Estimators for Difference-in-Differences". ENAR Spring Meeting, Philadelphia, PA, March 2019.

15. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics and Bioinformatics, Duke University, NC, February 2019.

16. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics and Informatics, University of Colorado Denver, CO, January 2019.

17. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics, University of Michigan, MI, January 2019.

18. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics, Yale University, CT, January 2019.

19. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Health Policy and Management, Johns Hopkins University, MD, January 2019.

20. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics, University of Florida, FL, December 2019.

21. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics and Epidemiology, University of Massachusetts Amherst, MA, December 2018.

22. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Population Medicine, Harvard Medical School, MA, December 2018.

23. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, PA, December 2018.

24. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Division of Biostatistics and Bioinformatics, University of California San Diego, CA, November 2018.

25. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Department of Biostatistics and Epidemiology, Rutgers School of Public Health, NJ, November 2018.

26. "Sample Size Determination for GEE Analyses of Stepped Wedge Cluster Randomized Trials". The 32nd New England Statistics Symposium, Amherst, MA, April 2018.

27. "Linking Design to Analysis: Recent Developments in Balancing Strategies for Cluster Randomized Trials." NIH Collaboratory Grand Rounds. February 2018.

28. "Linking Design to Analysis: Recent Developments in Balancing Strategies for Cluster Randomized Trials." Society for Clinical Trials 38th Annual Meeting, Liverpool, UK, May 2017.

Contributed Talks

1. "Propensity Score Weighting for Causal Inference with Multiple Treatments". Joint Statistical Meetings. Denver, CO, July 2019.
⋆*Received Student Paper Award from ASA Biometrics Section.*

2. "Power and Sample Size Requirements for GEE Analyses of Cluster Randomized Crossover Trials". Joint Statistical Meetings. Vancouver, British Columbia, Canada, July 2018.

3. "Estimating the Causal Effect of Traffic Safety Countermeasures on Crash Frequency: A Difference-in-Differences Approach" (Poster Presentation). Atlantic Causal Inference Conference. Pittsburgh, PA, May 2018.
⋆*Received NSF Travel Award from ACIC*.

4. "Sample Size Determination for GEE Analyses of Stepped Wedge Cluster Randomized Trials". ENAR Spring Meeting, Atlanta, GA, March 2018.
⋆*Received Distinguished Student Paper Award from ENAR*.

5. "Sample Size Determination for GEE Analyses of Stepped Wedge Cluster Randomized Trials". 12th International Conference on Health Policy Statistics, Charleston, SC, January 2018.

6. "An Evaluation of Constrained Randomization for the Design and Analysis of Group-randomized Trials with Binary Outcomes". Society for Clinical Trials 38th Annual Meeting, Liverpool, UK, May 2017.

7. "An Evaluation of Constrained Randomization for the Design and Analysis of Group-randomized Trials" (Poster Presentation). ENAR Spring Meeting, Austin, TX, March 2016.

8. "A Spatiotemporal Quantile Regression Model for Emergency Department Expenditures". 11th International Conference on Health Policy Statistics, Providence, RI, October 2015.
⋆*Received Student Travel Award from ICHPS*.

Honors and Awards

YSPH Early Career Investigator Research Award, Yale School of Public Health, May 2022

JSM Student Paper Award, Biometrics Section, American Statistical Association, July 2019

National Award for Outstanding Students Overseas, China Scholarship Council, April 2019

Chancellor's Award for Research Excellence, Duke University School of Medicine, October 2018

NSF Student Travel Award, Atlantic Causal Inference Conference, May 2018

Distinguished Student Paper Award, International Biometric Society (ENAR), March 2018

Student Travel Award, 11th International Conference on Health Policy Statistics, October 2015

Chancellor's Scholarship, Duke University School of Medicine, 2015

Chair's Academic Recognition Award, Duke Biostatistics & Bioinformatics, May 2014

Overall Academic Excellence Award, Duke Biostatistics & Bioinformatics, May 2014

Scholarship for Outstanding Academic Performance, Duke Biostatistics & Bioinformatics, 2013

Scholarship for Outstanding Academic Performance, Ministry of Education in China, 2009

Teaching

Instructor, Statistical Methods for Causal Inference (BIS 537), Yale, Fall 2022

Instructor, Statistical Methods for Causal Inference (BIS 537), Yale, Fall 2021

Instructor, Statistical Methods for Causal Inference (BIS 537), Yale, Fall 2020

Instructor, Seminar in Biostatistics (BIS 525), Yale, Spring 2021

Instructor, Seminar in Biostatistics (BIS 525), Yale, Fall 2020

Guest Lecturer, Case Studies in Bayesian Statistics (STA 723), Duke, Spring 2019

Guest Lecturer, Causal Inference (STA 640), Duke, Fall 2018

Instructor, Biostatistics Review for M.S Qualify Exam, Duke, Summer 2018, 2017, 2016

Teaching Assistant, Introduction to Biostatistics (STA 102), Duke, Spring 2016 and Spring 2013

Teaching Assistant, Mathematical Statistics II (BIOSTAT 704), Duke, Spring 2015

Teaching Assistant, Statistical Analysis (CRP 245), Duke, Fall 2012

Mentoring

Postdoctoral Mentees

Xueqi Wang, PhD (co-advise with Keith Goldfeld) Current

Bingkai Wang, PhD (K99 award co-mentor with Dylan Small and Nick Jewell) Current

Mary Ryan, PhD (co-advise with Denise Esserman) Current

Lama Ghazi, PhD (YCCI KL-2/Scholar Award co-mentor with Francis P. Wilson and Aldo Peixoto) 2020 - 2022
Currently Assistant Professor, Department of Epidemiology, University of Alabama at Birmingham School of Public Health

Kendra Davis-Plourde, PhD (co-advise with Monica Taljaard) 2020 - 2022
Currently Assistant Professor, Department of Biostatistics, Yale School of Public Health

Ondrej Blaha, PhD (co-advise with Denise Esserman) 2019 - 2020
Currently Associate Research Scientist, Department of Biostatistics & Yale Center for Analytical Sciences, Yale School of Public Health

PhD Dissertation Committee/Reader

Chao Cheng, Yale University, Department of Biostatistics Current

Jinghao Sun, Yale University, Department of Biostatistics Current

Lara Maleyeff, Harvard University, Department of Biostatistics Current

Cenjing Zhu, Yale University, Department of Chronic Disease Epidemiology Current

Margret Erlendsdottir, Yale University, Department of Biostatistics 2021

Ahmed Al-Jaishi, McMaster University, Department of Health Research Methods, Evidence and Impact (thesis external examiner) 2021

Xinyuan Chen, Yale University, Department of Statistics and Data Science (PhD dissertation reader) 2021

Xiaoxuan Cai, Yale University, Department of Biostatistics (PhD dissertation reader) 2020

PhD Qualifying Oral Exam/Prospectus Committee

Yichi Zhang, Yale University, Department of Biostatistics 2022

Samantha Dean, Yale University, Department of Biostatistics 2022

Melody Owen, Yale University, Department of Biostatistics 2022

Margret Erlendsdottir, Yale University, Department of Biostatistics 2022

MS Student Research Mentees

Jiaqi Tong, Yale University, Department of Biostatistics Current

Ruyi Liu, Yale University, Department of Biostatistics Current

Rui Li, Yale University, Department of Biostatistics Current

Wenxing Wang, Yale University, Department of Biostatistics Current

Xinyuan Tian, Yale University, Department of Biostatistics 2021 - 2022
Currently PhD student in Biostatistics, Department of Biostatistics, Yale University

Fangzhou Wei, Yale University, Department of Biostatistics 2021 - 2022
Currently Analyst at Analysis Group, California

Yumin Wang, Yale University, Department of Biostatistics 2020 - 2021
Currently MS student in Biostatistics, Department of Biomedical Informatics, Harvard Medical School

Zibo Tian, Yale University, Department of Biostatistics 2020 - 2021
Currently PhD student in Biostatistics, Department of Biostatistics, University of Florida

Zizhong Tian, Yale University, Department of Biostatistics 2020 - 2021
Currently PhD student in Biostatistics, Department of Public Health Sciences, Pennsylvania State University College of Medicine

Guanqun Meng, Yale University, Department of Biostatistics 2019 - 2020
Currently PhD student in Epidemiology and Biostatistics, Department of Population and Quantitative Health Sciences, Case Western Reserve University

University and School Service

Member, YSPH Global Health Concentration Faculty Advisory Committee 2020 - 2022

Organizer, "Statistical Methods for Pragmatic, Prevention and Implementation Trials", monthly working group between CMIPS and YCAS at YSPH 2021 - Present

Professional Activities

Editorial Boards

Associate Editor, *Statistics in Medicine* 2020 - Present

Associate Editor, *Clinical Trials* 2020 - Present

Editorial Board, *Epidemiologic Methods* 2021 - Present

Editorial Board, *Implementation Science* 2022 - Present

Peer Review Activities for Journals

American Journal of Epidemiology, American Journal of Public Health, Annals of Applied Statistics, Biometrika, Biometrics, Biometrical Journal, Biostatistics, BMC Medical Research Methodology, BMJ Open, Canadian Journal of Statistics, Communications in Statistics - Simulation and Computation, Clinical Trials, Epidemiologic Methods, Engineering, Health Services and Research Outcome Methodology, International Journal of Biostatistics, International Journal of

Epidemiology, Implementation Science, JAMA Cardiology, Journal of the American Statistical Association, Journal of Biopharmaceutical Statistics, Journal of Causal Inference, Journal of Clinical Epidemiology, Journal of Nonparametric Statistics, Journal of Statistical Computation and Simulation, Journal of the Royal Statistical Society (Series A and Series C), Pharmaceutical Statistics, Research Synthesis Methods, Statistics in Biosciences, Statistics in Medicine, Statistical Methods and Applications, Statistical Methods in Medical Research, Statistica Sinica, The American Statistician, Trials

Ad-hoc Reviewer of Grants

Patient-Centered Outcomes Research Institute (PCORI), 2020, 2019, 2018

PCORI Methods Consultant, 2022

National Institute on Minority Health and Health Disparities, 2020

The Wellcome Trust, UK, 2020

Paper Award Committee

ICSA Student Paper Award Committee, International Chinese Statistical Association, 2022

NESS Student Research Awards Committee, New England Statistics Symposium, 2022

Reviewer, ASA Section on Statistics in Epidemiology Young Investigator Awards, 2022

Invited session organizer, "Design and Analysis of Stepped Wedge Cluster Randomized Trials Based on Marginal Models for Discrete and Continuous Outcomes", Society for Clinical trials 43rd Annual Meeting, May 2022, San Diego, CA

Moderator, Design & Analysis of Embedded Pragmatic Clinical Trials Workshop, National Institute of Health (NIH), May 2019

Data Safety and Monitoring Committee, Implementation of Teleophthalmology in Rural Health Systems (I-TRUST) trial, PI: Yao Liu (funded by National Eye Institute), September 2021 – present

Other Professional Working Groups

Member, Society of Clinical Trials (SCT) Program Committee, 2022 – present

Member, ENAR Council for Emerging and New Statisticians (CENS), 2018 – 2020

Member, Biostatistics and Study Design Core, NIH Pragmatic Clinical Trials Collaboratory, 2013 – present

Executive Committee Member, Design and Statistics Core, NIA IMPACT Collaboratory, 2019 – present

Professional Societies

American Statistical Association (ASA), 2012 – present

Eastern North American Region (ENAR), 2016 – present

Society for Clinical Trials (SCT), 2016 – present