

Danyu Lin, Ph.D.

Department of Biostatistics
 Gillings School of Global Public Health
 University of North Carolina at Chapel Hill
 3101E McGavran-Greenberg Hall
 Chapel Hill, NC 27599-7420
 Phone: (919) 843-5134
 Fax: (919) 966-3804
 E-Mail: lin@bios.unc.edu
 Website: <https://dlin.web.unc.edu/>

PERSONAL

Date of birth: February 4, 1963
 Citizenship: United States

EDUCATION

1989	Ph.D., Biostatistics, University of Michigan, Ann Arbor, Michigan
1986	M.S., Biostatistics, University of Michigan, Ann Arbor, Michigan
1983	B.S., East China Normal University, Shanghai, China

PROFESSIONAL EXPERIENCE

1987–1988	Research Associate, Department of Biostatistics, University of Michigan
1988–1989	Research Intern, Biostatistics Center, University of Wisconsin
1989–1990	Research Fellow, Department of Biostatistics, Harvard University
1990–1991	Research Assistant Professor, Department of Biostatistics, University of Washington
1991–1994	Assistant Professor, Department of Biostatistics, University of Washington
1992–1994	Assistant Member, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center
1994–1998	Associate Professor, Department of Biostatistics, University of Washington
1994–1998	Associate Member, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center
1997–1998	Genentech Distinguished Professor in Biostatistics, University of Washington
1998–2000	Professor, Department of Biostatistics, University of Washington
1998–2000	Member, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center
1998–present	Special Government Employee (Consultant), Food and Drug Administration, U.S. Department of Health and Human Services
2001–2007	Affiliate Professor, Department of Biostatistics, University of Washington
2001–present	Dennis Gillings Distinguished Professor, Department of Biostatistics, University of North Carolina
2002–present	Member, Center for AIDS Research, University of North Carolina
2003–present	Member, Lineberger Comprehensive Cancer Center, University of North Carolina

AWARDS AND HONORS

1988–1989	Rackham Predoctoral Fellowship, University of Michigan
1989–1990	Fogarty Fellowship, National Institutes of Health
1992–1997	First Independent Research Award, National Institutes of Health
1999	Fellow, Institute of Mathematical Statistics
1999	Mortimer Spiegelman Award, American Public Health Association
2000	Fellow, American Statistical Association
2000	<i>Journal of the American Statistical Association–Theory and Methods</i> Discussion Paper
2002	Myrto Lefkopoulou Distinguished Lecturer, Harvard University
2003	Odoroff Memorial Lecturer, University of Rochester
2003	Highly Cited Researcher in Mathematics, Thomson ISI
2005–2015	National Institutes of Health MERIT Award
2006	<i>Journal of the American Statistical Association–Theory and Methods</i> Discussion Paper
2007	<i>Journal of the Royal Statistical Society, Series B</i> Discussion Paper
2011	Keynote address, ICSA (International Chinese Statistical Association) Applied Statistics Symposium
2015	COPSS (Committee of Presidents of Statistical Societies) George W. Snedecor Award
2015	ICSA (International Chinese Statistical Association) Distinguished Achievement Award
2019	Keynote address, The Second Conference on Lifetime Data Science, Pittsburgh, PA
2020	Best Publication in <i>Genetic Epidemiology</i>
2023	Best Mathematics Scientists (National Ranking: 95; World Ranking: 172), Research.com

PROFESSIONAL MEMBERSHIPS

- American Statistical Association
- International Biometric Society
- Institute of Mathematical Statistics
- International Chinese Statistical Association
- American Society of Human Genetics
- International Genetic Epidemiology Society

EDITORIAL SERVICE

1997–2000	Associate Editor, <i>Biometrics</i>
1998–2001	Associate Editor, <i>Probability and Statistics Letters</i>
1999–2009	Associate Editor, <i>Statistica Sinica</i>
1999–2023	Associate Editor, <i>Biometrika</i>
2006–2009	Associate Editor, <i>Scandinavian Journal of Statistics</i>
2012–2018	Associate Editor, <i>Journal of the American Statistical Association–Applications and Case Studies</i>
2014–present	Editorial Board, <i>Genetic Epidemiology</i>
2014–2020	Associate Editor, <i>Journal of the American Statistical Association–Theory and Methods</i>
2017–2022	Editor, <i>Statistical Theory and Related Fields</i>
2019–2023	Associate Editor, <i>Journal of the American Statistical Association–Reviews</i>

ADVISORY BOARDS AND COMMITTEES

1994–1995	Organizing Committee (Chair), First Seattle Symposium in Biostatistics: Survival Analysis
1997	Biostatistics Study Section, National Institutes of Health
1998	Program Committee, The Fourth ICSA Statistical Conference
1999	IMS Program Committee, Joint Statistical Meetings
1999	AIDS & Related Research Review Group, National Institutes of Health
1999–2000	Organizing Committee (Chair), Second Seattle Symposium in Biostatistics: Analysis of Correlated Data
2000–2010	Visiting Lecturer Program in Statistical Science, Committee of Presidents of the Statistical Societies (COPSS)
2000–2002	Mortimer Spiegelman Award Committee (Chair, 2002), American Public Health Association
2001–2003	Social Sciences, Nursing, Epidemiology and Methods (SNEM-5) Study Section, National Institutes of Health
2002	Site Visit Team, Biostatistics Branch, National Cancer Institute
2002	Review Panel for SNEM-5 Member Applications, National Institutes of Health
2003–2004	Program Committee, The International Biometric Society ENAR 2004 Spring Meetings
2004–2006	Review Panel for BMRD Member Applications, National Institutes of Health
2004–2013	National Heart, Lung and Blood Institute's Protocol Review and DSMB Committees for the Cardiovascular Outcomes in Renal Atherosclerotic Lesions Trial
2004	Independent Expert Panel on Statistical Methods for the Analysis of Organ Transplantation Data, Health Resources and Services Administration, U.S. Department of Health and Human Services
2006	National Cancer Institute P01 Cluster Review
2006–2007	Program Committee (Chair) and Organizing Committee (Co-Chair), ICSA 16th Applied Statistics Symposium
2007	External Advisory Committee, Public Health Sciences Division, Fred Hutchinson Cancer Research Center
2004, 2008, 2010	Tenure Review Committees, Harvard University
2015–2019	Data Monitoring Committee, Celgene ACE-536-B-THAL-001 Study
2018–present	Data Monitoring Committee, Celgene ACE-536-B-THAL-002 Study
2019–present	Data Monitoring Committee, Celgene ACE-536-B-THAL-004 Study
2017–present	Statistical Science Advisory Board on Biosimilars, AMGEN
2019	Special Emphasis Panel for Member Conflict: Epidemiology and Genetic Epidemiology, National Institutes of Health
2020	Biostatistics Methods and Research Design (BMRD) Study Section, National Institutes of Health
2020–2021	COPSS (Committee of Presidents of Statistical Societies) George W Snedecor Award Committee
2021–2022	Program Committee for the 12th ICSA (International Chinese Statistical Association) International Conference
2022–present	European Federation of Statisticians in the Pharmaceutical Industry Vaccine Steering Committee

PEER-REVIEWED ARTICLES

Google Scholar: <https://scholar.google.com/citations?user=SG22hu0AAAAJ&hl=en>

Citations: >45,000

h-index: 96

1. Wei LJ, **Lin DY**, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* 1989; 84: 1065–1073.
2. **Lin DY**, Wei LJ. The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* 1989; 84: 1074–1078.
3. Wei LJ, Smythe RT, **Lin DY**, Park TS. Statistical inference with data-dependent treatment allocation rules. *Journal of the American Statistical Association* 1990; 85: 156–162.
4. **Lin DY**. MULCOX: A computer program for the Cox regression analysis of multiple failure time variables. *Computer Methods and Programs in Biomedicine* 1990; 32: 125–135.
5. Wei LJ, Ying Z, **Lin DY**. Linear regression analysis of censored survival data based on rank tests. *Biometrika* 1990; 77: 845–851.
6. **Lin DY**, Wei LJ. Global goodness-of-fit tests for the general Cox regression model. *Statistica Sinica* 1991; 1: 1–17.
7. **Lin DY**. Nonparametric sequential testing in clinical trials with incomplete multivariate observations. *Biometrika* 1991; 78: 123–131.
8. **Lin DY**, Wei LJ. Repeated confidence intervals for a scale change in a sequential survival study. *Biometrics* 1991; 47: 289–294.
9. **Lin DY**. Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators. *Journal of the American Statistical Association* 1991; 86: 725–728.
10. **Lin DY**, Wei LJ, DeMets DL. Exact statistical inference for group sequential trials. *Biometrics* 1991; 47: 1399–1408.
11. **Lin DY**, Liu PY. Nonparametric sequential tests against ordered alternatives in multiple-armed clinical trials. *Biometrika* 1992; 79: 420–425.
12. **Lin DY**, Geyer CJ. Computational methods for semiparametric linear regression with censored data. *Journal of Computational and Graphical Statistics* 1992; 1: 77–90.
13. **Lin DY**. Sequential log rank tests adjusting for covariates with the accelerated life model. *Biometrika* 1992; 79: 523–529.
14. **Lin DY**. GOF-COX: A computer program for the goodness-of-fit analysis of the Cox proportional hazards model. *Computer Methods and Programs in Biomedicine* 1992; 38: 101–105.
15. **Lin DY**, Fischl MJA, Schoenfeld DA. Evaluating the role of CD4-lymphocyte counts as surrogate endpoints in human immunodeficiency virus clinical trials. *Statistics in Medicine* 1993; 12: 835–842.
16. **Lin DY**. MULCOX2: A general computer program for the Cox regression analysis of multivariate failure time data. *Computer Methods and Programs in Biomedicine* 1993; 40: 279–293.
17. Koenig JQ, Larson TV, Hanley QS, Rebolledo V, Dumler K, Checkoway H, Wang SZ, **Lin DY**, Pierson WE. Pulmonary function changes in children associated with fine particulate matter. *Environmental Research* 1993; 63: 26–38.
18. **Lin DY**, Wei LJ, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* 1993; 80: 557–572.
19. **Lin DY**, Ying Z. A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika* 1993; 80: 573–581.
20. **Lin DY**, Ying Z. Cox regression with incomplete covariate measurements. *Journal of the American Statistical Association* 1993; 88: 1341–1349.
21. **Lin DY**, Ying Z. Semiparametric analysis of the additive risk model. *Biometrika* 1994; 81: 61–71.

22. **Lin DY**, Fleming TR, Wei LJ. Confidence bands for survival curves under the proportional hazards model. *Biometrika* 1994; 81: 73–81.
23. Guo SW, **Lin DY**. Regression analysis of multivariate grouped survival data. *Biometrics* 1994; 50: 632–639.
24. **Lin DY**. Cox regression analysis of multivariate failure time data: The marginal approach. *Statistics in Medicine* 1994; 13: 2233–2247.
25. **Lin DY**, Ying Z. Semiparametric inference for the accelerated life model with time-dependent covariates. *Journal of Statistical Planning and Inference* 1995; 44: 47–63.
26. Crawford SW, Pepe M, **Lin DY**, Benedetti F, Deeg HJ. Abnormalities of pulmonary function tests after marrow transplantation predict non-relapse mortality. *American Journal of Respiratory and Critical Care Medicine* 1995; 152: 690–695.
27. Combes B, Carithers RL, Maddrey WC, **Lin DY** et al.. A randomized, double-blind, placebo-controlled trial of ursodeoxycholic acid in primary biliary cirrhosis. *Hepatology* 1995; 22: 759–766.
28. **Lin DY**, Ying Z. Semiparametric analysis of general additive-multiplicative hazard models for counting processes. *The Annals of Statistics* 1995; 23: 1712–1734.
29. Psaty BW, Siscovick DS, Weiss NS, Koepsell TD, Rosendaal FR, **Lin DY**, Hekbert SR, Wagner EM, Furberg CD. Hypertension and outcomes research: from clinical trials to clinical epidemiology. *American Journal of Hypertension* 1996; 9: 178–183.
30. **Lin DY**, Spiekerman CF. Model checking techniques for parametric regression with censored data. *Scandinavian Journal of Statistics* 1996; 23: 157–177.
31. Yip PSF, Huggins RM, **Lin DY**. Inference for capture-recapture experiments in continuous time with variable capture rates. *Biometrika* 1996; 83: 477–483.
32. **Lin DY**, Robins JM, Wei LJ. Comparing two failure time distributions in the presence of dependent censoring. *Biometrika* 1996; 83: 381–393.
33. Spiekerman CF, **Lin DY**. Checking the marginal Cox model for correlated failure time data. *Biometrika* 1996; 83: 143–156.
34. **Lin DY**, Shen L, Ying Z, Breslow NE. Group sequential designs for monitoring survival probabilities. *Biometrics* 1996; 52: 1033–1041.
35. DeGruttola V, Fleming TR, **Lin DY**, Coombs R. Validating surrogate markers: Are we being naive? *Journal of Infectious Diseases* 1997; 175: 237–246.
36. **Lin DY**. Nonparametric inference for cumulative incidence functions in competing-risks studies. *Statistics in Medicine* 1997; 16: 901–910.
37. **Lin DY**, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997; 53: 419–434.
38. **Lin DY**, Fleming TR, DeGruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. *Statistics in Medicine* 1997; 16: 1515–1527.
39. Deeg HJ, **Lin DY**, Leisenring W, Boeckh M, Anasetti C, Appelbaum FR, Chauncey TR, Doney K, Flowers M, Martin P, Nash R, Schoch G, Sullivan KM, Whitherspoon RP, Storb R. Cyclosporine or cyclosporine plus methylprednisolone for prophylaxis of graft-versus-host disease: a prospective randomized trial. *Blood* 1997; 89: 3880–3887.
40. Hekbert SR, Weiss NS, Koepsell TD, Lemaitre RN, Smith NL, Siscovick DS, **Lin DY**, Psaty BM. Duration of estrogen replacement therapy in relation to the risk of incident myocardial infarction in postmenopausal women. *Archives of Internal Medicine* 1997; 157: 1330–1336.
41. Schwartz SM, Siscovick DS, Longstreth WT, Psaty BM, Beverly RK, Raghunathan TE, **Lin DY**, Koepsell TD. Use of low-dose oral contraceptives and stroke in young women. *Annals of Internal Medicine* 1997; 127: 596–603.
42. **Lin DY**, Oakes D, Ying Z. Additive hazards regression with current status data. *Biometrika* 1998; 85: 289–298.

43. **Lin DY**, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* 1998; 54: 948–963.
44. **Lin DY**, Wei LJ, Ying Z. Accelerated failure time models for counting processes. *Biometrika* 1998; 85: 605–618.
45. Spiekerman CF, **Lin DY**. Marginal regression models for multivariate failure time data. *Journal of the American Statistical Association* 1998; 93: 1164–1175.
46. Albert JM, Ioannidis JPA, Reichelderfer P, Conway B, Coombs RW, Crane L, DeMasi R, Dixon DO, Flandre P, Hughes MD, Kalish L, Larntz K, **Lin DY**, et al. Statistical issues for HIV surrogate endpoints: Point/counterpoint. *Statistics in Medicine* 1998; 17: 2435–2462.
47. **Lin DY**, Arbogast P, Siscovick DS, Lemaitre RN. Poisson regression with missing durations of exposure. *Biometrics* 1999; 55: 252–257.
48. **Lin DY**, Yip PSF. Parametric regression models for continuous-time removal and recapture studies. *Journal of the Royal Statistical Society, Series B* 1999; 61: 401–411.
49. **Lin DY**, Sun W, Ying Z. Nonparametric estimation of the gap time distributions for serial events with censored data. *Biometrika* 1999; 86: 59–70.
50. Fisher LD, **Lin DY**. Time-dependent covariates in the Cox proportional hazards regression model. *Annual Review of Public Health* 1999; 20: 145–157.
51. Diehr P, Yanez D, Ash A, Hombrook M, **Lin DY**. Methods for analyzing health care utilization and costs. *Annual Review of Public Health* 1999; 20: 125–144.
52. Psaty BM, Koepsell TD, **Lin DY**, Weiss NS, Siscovick DS, Rosendaal FR, Pahor M, Furberg CD. Assessment and control for confounding by indication in observational studies. *Journal of the American Geriatrics Society* 1999; 47: 749–754.
53. Lemaitre RN, Siscovick DS, Raghunathan TE, Weinmann S, Arbogast P, **Lin DY**. Leisure-time physical activity and the risk of primary cardiac arrest. *Archives of Internal Medicine* 1999; 159: 686–690.
54. **Lin DY**, Yao Q, Ying Z. A general theory on stochastic curtailment for censored survival data. *Journal of the American Statistical Association* 1999; 94: 510–521.
55. Psaty BM, Weiss NS, Furberg CD, Koepsell TD, Siscovich DS, Rosendaal FR, Smith NL, Heckbert SR, Kaplan RC, **Lin DY**, Fleming TR, Wagner EH. Surrogate endpoints, health outcomes, and the drug-approval process for the treatment of risk factors for cardiovascular diseases. *Journal of the American Medical Association* 1999; 282: 786–790.
56. Etzioni RD, Feuer EJ, Sullivan SD, **Lin DY**, Hu C, Ramsey SD. On the use of survival analysis techniques to estimate medical care costs. *Journal of Health Economics* 1999; 18: 365–380.
57. Yip PSF, Zhou Y, **Lin DY**, Fang XZ. Estimation of population size based on additive hazards models for continuous-time recapture experiments. *Biometrics* 1999; 55: 904–908.
58. Kulich M, **Lin DY**. Additive hazards regression with covariate measurement error. *Journal of the American Statistical Association* 2000; 95: 238–248.
59. **Lin DY**. Linear regression analysis of censored medical costs. *Biostatistics* 2000; 1: 35–47.
60. **Lin DY**. On fitting Cox’s proportional hazards models to survey data. *Biometrika* 2000; 87: 37–47.
61. Kulich M, **Lin DY**. Additive hazards regression for case-cohort studies. *Biometrika* 2000; 87: 73–87.
62. Ghosh D, **Lin DY**. Nonparametric analysis of recurrent events and death. *Biometrics* 2000; 56: 554–562.
63. **Lin DY**. Proportional means regression for censored medical costs. *Biometrics* 2000; 56: 775–778.
64. **Lin DY**, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, Series B* 2000; 62: 711–730.
65. Fleming TR, **Lin DY**. Survival analysis in clinical trials: past developments and future directions. *Biometrics* 2000; 56: 971–983 (Editors’ Invited Paper).
66. Siscovick DS, Raghunathan TE, **Lin DY**, Weinmann S, Arbogast P, Lemaitre RV, Psaty BM, Alexander R, Cobb L. Influenza vaccination and the risk of primary cardiac arrest. *American Journal of*

- Epidemiology* 2000; 152: 674–677.
67. **Lin DY**, Ying Z. Semiparametric and nonparametric regression analysis of longitudinal data (with discussion). *Journal of the American Statistical Association* 2001; 96: 103–126 (Editor’s Invited Paper for the 2000 Joint Statistical Meetings).
 68. **Lin DY**, Ying Z. Nonparametric tests for the gap time distributions of serial events based on censored data. *Biometrics* 2001; 57: 369–375.
 69. **Lin DY**, Wei LJ, Ying Z. Semiparametric transformation models for point processes. *Journal of the American Statistical Association* 2001; 96: 620–628.
 70. Willan AR, **Lin DY**. Incremental net benefit in randomized clinical trials. *Statistics in Medicine* 2001; 20: 1563–1574.
 71. Whitsel EA, Raghunathan TE, Pearce RM, **Lin DY**, Rautaharju PM, Lemaitre R, Siscovick DS. RR interval variation, the QT interval index and risk of primary cardiac arrest among patients without clinically recognized heart disease. *European Heart Journal* 2001; 22: 165–173.
 72. Heckbert SR, Kaplan RC, Weiss NS, Psaty BM, **Lin DY**, Furberg CD, Starr JR, Anderson GD, LaCroix AZ. Risk of recurrent coronary events in relation to use and recent initiation of postmenopausal hormone therapy. *Archives of Internal Medicine* 2001; 161: 1709–1713.
 73. Rea TD, Heckbert SR, Kaplan RC, Psaty BM, Smith NL, Lemaitre RN, **Lin DY**. Body mass index and the risk of recurrent coronary events following acute myocardial infarction. *American Journal of Cardiology* 2001; 88: 467–472.
 74. **Lin DY**, Wei LJ, Ying Z. Model-checking techniques based on cumulative residuals. *Biometrics* 2002; 58: 1–12.
 75. Arbogast PG, **Lin DY**, Siscovick DS, Schwartz SM. Estimating incidence rates from population-based case-control studies in the presence of nonrespondents. *Biometric Journal* 2002; 44: 227–239.
 76. Ghosh D, **Lin DY**. Marginal regression models for recurrent and terminal events. *Statistica Sinica* 2002; 12: 663–688.
 77. Hu C, **Lin DY**. Cox regression with covariate measurement error. *Scandinavian Journal of Statistics* 2002; 29: 637–655.
 78. Willan AR, **Lin DY**, Cook RJ, Chen EB. Using inverse-weighting in cost-effectiveness analysis with censored data. *Statistical Methods in Medical Research* 2002; 11: 539–551.
 79. Lemaitre RN, Siscovick DS, Psaty BM, Pearce RM, Raghunathan TE, Whitsel EA, Weinmann SA, Anderson GD, **Lin DY**. Inhaled Beta-2 adrenergic receptor agonists and primary cardiac arrest. *American Journal of Medicine* 2002; 113: 711–716.
 80. Willan AR, Chen EB, Cook RJ, **Lin DY**. Incremental net benefit in randomized clinical trials with quality-adjusted survival. *Statistics in Medicine* 2003; 22: 353–363.
 81. **Lin DY**. Regression analysis of incomplete medical cost data. *Statistics in Medicine* 2003; 22: 1181–1200.
 82. **Lin DY**, Ying Z. Semiparametric regression analysis of longitudinal data with informative drop-outs. *Biostatistics* 2003; 4: 385–398.
 83. Jin Z, **Lin DY**, Wei LJ, Ying Z. Rank-based inference for the accelerated failure time model. *Biometrika* 2003; 90: 341–353.
 84. Ghosh D, **Lin DY**. Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* 2003; 59: 877–885.
 85. Rea TD, Siscovick DS, Psaty BM, Pearce RM, Raghunathan TE, Whitsel EA, Cobb LA, Weinmann S, Anderson GD, Arbogast P, **Lin DY**. Digoxin therapy and the risk of primary cardiac arrest in patients with congestive heart failure: Effect of mild-moderate renal impairment. *Journal of Clinical Epidemiology* 2003; 56: 646–650.
 86. Combes B, Luketic VA, Peters MG, Zetterman RK, Garcia-Tsao G, Munoz SJ, **Lin DY**, Flye N, Carithers RL. Prolonged follow-up of patients in the U.S. multicenter trial of ursodeoxycholic acid for

- primary biliary cirrhosis. *American Journal of Gastroenterology* 2004; 99: 264–268.
87. **Lin DY**. Haplotype-based association analysis in cohort studies of unrelated individuals. *Genetic Epidemiology* 2004; 26: 255–264.
 88. Hu C, **Lin DY**. Semiparametric failure time regression with replicates of mismeasured covariates. *Journal of the American Statistical Association* 2004; 99: 105–118.
 89. Kulich M, **Lin DY**. Improving the efficiency of relative-risk estimation in case-cohort studies. *Journal of the American Statistical Association* 2004; 99: 832–844.
 90. **Lin DY**, Zou F. Assessing genomewide statistical significance in linkage studies. *Genetic Epidemiology* 2004; 27: 202–214.
 91. Diao G, **Lin DY**, Zou F. Mapping quantitative trait loci with censored observations. *Genetics* 2004; 168: 1689–1698. PMID: PMC1448809.
 92. Zou F, Fine JP, Hu J, **Lin DY**. An efficient resampling method for assessing genomewide statistical significance in mapping quantitative trait loci. *Genetics* 2004; 168: 2307–2316. PMID: PMC1448705.
 93. Arbogast PG, **Lin DY**. Goodness-of-fit methods for matched case-control studies. *Canadian Journal of Statistics* 2004; 32: 373–386.
 94. Arbogast PG, **Lin DY**. Model-checking techniques for stratified case-control studies. *Statistics in Medicine* 2005; 24: 229–247.
 95. Zeng D, **Lin DY**. Estimating haplotype-disease associations with pooled genotype data. *Genetic Epidemiology* 2005; 28: 70–82.
 96. Willan AR, **Lin DY**, Manca A. Regression methods for cost-effectiveness analysis with censored data. *Statistics in Medicine* 2005; 24: 131–145.
 97. Diao G, **Lin DY**. A powerful and robust method for mapping quantitative trait loci in general pedigrees. *American Journal of Human Genetics* 2005; 77: 97–111.
 98. Zeng D, **Lin DY**, Yin G. Maximum likelihood estimation for the proportional odds model with random effects. *Journal of the American Statistical Association* 2005; 100: 470–483.
 99. **Lin DY**. An efficient Monte Carlo approach to assessing statistical significance in genomic studies. *Bioinformatics* 2005; 21: 781–787.
 100. Whitsel EA, Boyko EJ, Rautaharju PM, Raghunathan TE, **Lin DY**, Pearce RM, Weinmann SA, Siscovick DS. Electrocardiographic QT interval prolongation and risk of primary cardiac arrest in diabetic patients. *Diabetes Care* 2005; 28: 2045–2047.
 101. **Lin DY**. On rapid simulation of p-values in association studies. *American Journal of Human Genetics* 2005; 77: 513–514. PMID: PMC1226216.
 102. Diao G, **Lin DY**. Semiparametric methods for mapping quantitative trait loci with censored data. *Biometrics* 2005; 61: 789–798.
 103. Pan Z, **Lin DY**. Goodness-of-fit methods for generalized linear mixed models. *Biometrics* 2005; 61: 1000–1009.
 104. **Lin DY**, Zeng D, Millikan R. Maximum likelihood estimation of haplotype effects and haplotype-environment interactions in association studies. *Genetic Epidemiology* 2005; 29: 299–312.
 105. Jin Z, **Lin DY**, Ying Z. Rank regression analysis of multivariate failure time data based on marginal linear models. *Scandinavian Journal of Statistics* 2006; 33: 1–23.
 106. **Lin DY**. Evaluating statistical significance in two-stage genomewide association studies. *American Journal of Human Genetics* 2006; 78: 505–509. PMID: PMC1380293.
 107. Diao G, **Lin DY**. Improving the power of association tests for quantitative traits in family studies. *Genetic Epidemiology* 2006; 30: 301–313.
 108. Jin Z, **Lin DY**, Ying Z. On least-squares regression with censored data. *Biometrika* 2006; 93: 147–161.
 109. **Lin DY**, Zeng D. Likelihood-based inference on haplotype effects in genetic association studies (with discussion). *Journal of the American Statistical Association* 2006; 101: 89–118.
 110. Zeng D, **Lin DY**. Efficient estimation of semiparametric transformation models for counting processes.

- Biometrika* 2006; 93: 627–640.
111. Zeng D, **Lin DY**, Avery CL, North KE, Bray MS. Efficient semiparametric estimation of haplotype-disease associations in case-cohort and nested case-control studies. *Biostatistics* 2006; 7: 486–502.
 112. Diao G, **Lin DY**. Semiparametric variance-component models for linkage and association analyses of censored trait data. *Genetic Epidemiology* 2006; 30: 570–581.
 113. Lange LA, Carlson CS, Hindorf LA, Lange EM, Walston J, Durda JP, Cushman M, Bis JC, Zeng D, **Lin DY**, Kuller LH, Nickerson DA, Psaty BM, Tracy RP, Reiner AP. Association of polymorphisms in the CRP gene with circulating c-reactive protein levels and cardiovascular events. *Journal of the American Medical Association* 2006; 296: 2703–2711.
 114. **Lin DY**, Huang BE. The use of inferred haplotypes in downstream analyses. *American Journal of Human Genetics* 2007; 80: 577–579. PMID: PMC1821109.
 115. Huang BE, **Lin DY**. Efficient association mapping of quantitative trait loci with selective genotyping. *American Journal of Human Genetics* 2007; 80: 567–576. PMID: PMC1821103.
 116. Zeng D, **Lin DY**. Semiparametric transformation models with random effects for recurrent events. *Journal of the American Statistical Association* 2007; 102: 167–180.
 117. Zeng D, **Lin DY**. Maximum likelihood estimation in semiparametric regression models with censored data (with discussion). *Journal of the Royal Statistical Society, Series B* 2007; 69: 507–564.
 118. **Lin DY**. On the Breslow estimator. *Lifetime Data Analysis* 2007; 13: 471–480.
 119. Huang BE, Amos CI, **Lin DY**. Detecting haplotype effects in genomewide association studies. *Genetic Epidemiology* 2007; 31: 803–812.
 120. Zeng D, **Lin DY**. Efficient estimation for the accelerated failure time model. *Journal of the American Statistical Association* 2007; 102: 1387–1396.
 121. Elston RC, **Lin DY**, Zheng G. Multi-stage sampling for genetic studies. *Annual Review of Genomics and Human Genetics* 2007; 8: 327–342.
 122. Zeng D, **Lin DY**, Lin X. Semiparametric transformation models with random effects for clustered failure time data. *Statistica Sinica* 2008; 18: 355–377. PMID: PMC2756664.
 123. **Lin DY**, Hu Y, Huang BE. Simple and efficient analysis of disease association with missing genotype data. *American Journal of Human Genetics* 2008; 82: 444–452. PMID: PMC2427170.
 124. Johnson BA, **Lin DY**, Zeng Z. Penalized estimating functions and variable selection in semiparametric regression models. *Journal of the American Statistical Association* 2008; 103: 672–680. PMID: PMC2850080.
 125. Zeng D, **Lin DY**. Efficient resampling methods for non-smooth estimating functions. *Biostatistics* 2008; 9: 355–363. PMID: PMC2673016.
 126. Sullivan PF, **Lin DY**, Tzeng J-Y, van den Oord E, Perkins D, Stroup TS, Wanger M, Lee S, Wright FA, Zou F, Liu W, Downing AM, Lieberman J, Close SL. Genomewide association for schizophrenia in the CATIE study. *Molecular Psychiatry* 2008; 13: 570–584. PMID: PMC3910086.
 127. Zeng D, **Lin DY**. Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. *Biometrics* 2009; 65: 746–752. PMID: PMC3030128.
 128. Sullivan PF, de Geus EJC, Willemsen G, James MR, Smit JH, Zandbelt T, Arolt V, Baune BT, Blackwood D, Cichon S, Coventry WL, Farmer A, Fava M, Gordon SD, He Q, Heath AC, Heutink P, Holsboer F, Hoogendijk WJ, Hottenga JJ, Hu Y, Kohli M, **Lin DY**, Lucae S, MacIntyre DJ, Maier W, McGhee KA, McGuffin P, Montgomery GW, Muir WJ, Nolen WA, Nöthen MM, Perlis RH, Pirlo K, Posthuma D, Rietschel M, Rizzu P, Schosser A, Smit AB, Smoller JW, Tzeng J-Y, van Dyck R, Verhage M, Zitman FG, Martin NG, Wray NR, Booma DI, Penninx B. Genomewide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Molecular Psychiatry* 2009; 14: 359–375. PMID: PMC2717726.
 129. **Lin DY**, Zeng D. Proper analysis of secondary phenotype data in case-control association studies. *Genetic Epidemiology* 2009; 33: 256–265. PMID: PMC2684820.

130. **Lin DY**, Sullivan PF. Meta-analysis of genome-wide association studies with overlapping subjects. *American Journal of Human Genetics* 2009; 85: 862–872. PMID: PMC2790578.
131. Joubert BR, Diao G, **Lin DY**, North KE, Franceschini N. Longitudinal age-dependent effect on systolic blood pressure. *BMC Proceedings* 2009; 3: S87. PMID: PMC2795990.
132. Yan Y, Hu YJ, North KE, Franceschini N, **Lin DY**. Evaluation of population impact of candidate polymorphisms for coronary heart disease in the Framingham study offspring cohort. *BMC Proceedings* 2009; 3: S118. PMID: PMC2795890.
133. Zeng D, **Lin DY**. A general asymptotic theory for maximum likelihood estimation in semiparametric regression models with censored data. *Statistica Sinica* 2010; 20: 871–910. PMID: PMC2888521.
134. **Lin DY**, Zeng D. Meta-analysis of genome-wide association studies: no efficiency gain in using individual participant data. *Genetic Epidemiology* 2010; 34: 60–66. PMID: PMC3878085.
135. Diao G, **Lin DY**. Variance-components methods for linkage and association analysis of ordinal traits in general pedigrees. *Genetic Epidemiology* 2010; 34: 232–237. PMID: PMC3003595.
136. **Lin DY**, Zeng D. On the relative efficiency of using summary statistics versus individual level data in meta-analysis. *Biometrika* 2010; 97: 321–332. PMID: PMC3412575.
137. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature Genetics* 2010; 42: 441–447 (chair of the Statistical Working Group and a senior author). PMID: PMC2914600.
138. Chen L, **Lin DY**, Zeng D. Attributable fraction functions for censored event times. *Biometrika* 2010; 97: 713–726. PMID: PMC3744602.
139. Hu YJ, **Lin DY**, Zeng D. A general framework for studying genetic effects and gene-environment interactions with missing data. *Biostatistics* 2010; 11: 583–598. PMID: PMC3294269.
140. Hu YJ, **Lin DY**. Analysis of untyped SNPs: Maximum likelihood and imputation methods. *Genetic Epidemiology* 2010; 34: 803–815. PMID: PMC3030127.
141. Liu Y, Blackwood DH, Caesar S, de Geus EJC, Farmer A, Ferreira MAR, Ferrier IN, Fraser C, Gordon-Smith K, Green EK, Grozeva D, Gurling HM, Hamshere ML, Heutink P, Holmans PA, Hoogendijk WJ, Hottenga JJ, Jones L, Jones IR, Kirov G, **Lin DY**, McGuffin P, Moskvina V, Nolen WA, Perlis RH, Posthuma D, Scolnick EM, Smit AB, Smit JH, Smoller JW, St. Clair D, van Dyck R, Verhage M, Wellcome Trust Case-Control Consortium, Willemsen G, Young AH, Zandbelt T, Boomsma DI, Craddock N, O'Donovan MC, Owen MJ, Penninx BWJH, Purcell S, Sklar P, Sullivan P. Meta-analysis of genome-wide association data of bipolar disorder and major depressive disorder. *Molecular Psychiatry* 2011; 16: 2–4. PMID: PMC3883627.
142. He Q, **Lin DY**. A variable selection method for genome-wide association studies. *Bioinformatics* 2011; 27: 1–8. PMID: PMC3025714.
143. Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, **Lin DY**, Duan J, Ophoff RA, Andreassen OA, et al. Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics* 2011; 43: 969–976 (contributions to statistical analyses and manuscript preparation). PMID: PMC3303194.
144. Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, Craddock N, Edenberg HJ, Nurnberger JI, Rietschel M, Blackwood D, Corvin A, Flickinger M, Guan WH, Mattingsdal M, McQuillin A, Kwan P, Wienker TF, Daly M, Dudbridge F, Holmans PA, **Lin DY**, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature Genetics* 2011; 43: 977–983 (contributions to statistical analyses). PMID: PMC3637176.
145. Avery CL, He QC, North KE, Ambite JL, Boerwinkle E, Fornage M, Hindorff LA, Kooperberg C, Meigs JB, Pankow JS, Pendergrass SA, Psaty BM, Ritchie MD, Rotter JI, Taylor KD, Wilkens LR, Heiss G, **Lin DY**. A phenomics-based strategy identifies loci on APOC1, BRAP, and PLCG1 associated with metabolic syndrome phenotype domains. *PLOS Genetics* 2011; 7: e1002322. PMID: PMC3192835.
146. Nyante SJ, Gammon MD, Kaufman JS, Bensen JT, **Lin DY**, Barnholtz-Sloan JS, Hu YJ, He QC, Luo

- JC, Millikan RC. Common genetic variation in adiponectin, leptin, and leptin receptor and association with breast cancer subtypes. *Breast Cancer Research and Treatment* 2011; 129: 593–606. PMID: PMC3355661.
147. **Lin DY**, Zeng D. Correcting for population stratification in genome-wide association studies. *Journal of the American Statistical Association* 2011; 106: 997–1008. PMID: PMC3314247.
148. **Lin DY**, Tang ZZ. A general framework for detecting disease associations with rare variants in sequencing studies. *American Journal of Human Genetics* 2011; 89: 354–367. PMID: PMC3169821.
149. Franceschini N, Carty C, Bužková P, Reiner A, Garrett T, Lin Y, Vöckler JS, Hindorff LA, Cole SA, Boerwinkle E, **Lin DY**, Bookman E, Best LG, Bella JN, Eaton C, Greenland P, Jenny N, North KE, Taverna D, Young AM, Deelman E, Kooperberg C, Psaty B, Heiss G. Association of genetic variants and incident coronary heart disease in multi-ethnic cohorts: The PAGE study. *Circulation: Cardiovascular Genetics* 2011; 4: 661–672. PMID: PMC3293207.
150. Chen L, **Lin DY**, Zeng D. Checking semiparametric transformation models with censored data. *Biostatistics* 2012; 13: 18–31. PMID: PMC3276276.
151. Chen L, **Lin DY**, Zeng D. Predictive accuracy of covariates for event times. *Biometrika* 2012; 99: 615–630. PMID: PMC3635702.
152. Avery CL, Sethupathy P, Buyske S, He Q, **Lin DY**, Arking DE, Carty CL, Duggan D, Fesinmeyer MD, Hindorff LA, Jeff JM, Klein L, Patton KK, Peters U, Shohet R, Sotoodehnia N, Kooperberg C, Haiman CA, Mohlke KL, Whitsel EA, North KE. Fine-mapping and initial characterization of QT interval loci in African Americans. *PLOS Genetics* 2012; 8: e1002870. PMID: PMC3415454.
153. Berndt SI, Gustafsson S, Mägi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-Chonka DC, Day FR, Esko T, Fall T, Ferreira T, Gentilini D, Jackson AU, Luan J, Randall JC, Vedantam S, Willer CJ, Winkler TW, Wood AR, Workalemahu T, Hu YJ, Lee SH, Liang L, **Lin DY**, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nature Genetics* 2013; 45: 501–512. PMID: PMC3973018.
154. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium (Sullivan PF, Daly MJ, Ripke S, Lewis CM, **Lin DY**, et al). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry* 2013; 18: 497–511. PMID: PMC3837431.
155. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics* 2013; 45: 984–994. PMID: PMC3800159.
156. **Lin DY**. An overview of Ross Prentice’s contributions to statistical science. *Statistics in Biosciences* 2013; 5: 224–231. PMID: PMC3821774.
157. Tang ZZ, **Lin DY**. MASS: meta-analysis of score statistics for sequencing studies. *Bioinformatics* 2013; 29: 1803–1805. PMID: PMC3702254.
158. **Lin DY**, Zeng D, Tang, ZZ. Quantitative trait analysis in sequencing studies under trait-dependent sampling. *Proceedings of the National Academy of Sciences* 2013; 110: 12247–12252. PMID: PMC3725118.
159. Hu YJ, Berndt SI, Gustafsson S, Ganna A, Genetic Investigation of ANthropometric Traits (GIANT) Consortium, Hirschhorn J, North KE, Ingelsson E, **Lin DY**. Meta-analysis of gene-level associations for rare variants based on single-variant statistics. *American Journal of Human Genetics* 2013; 93: 236–248. PMID: PMC3738834.
160. He Q, Avery CL, **Lin DY**. A general framework for association tests with multivariate traits in large-scale genomics studies. *Genetic Epidemiology* 2013; 37: 759–767. PMID: PMC3926135.
161. **Lin DY**. Survival analysis with incomplete genetic data. *Lifetime Data Analysis* 2014; 20: 16–22. PMID: PMC3806886.
162. Lange LA, Hu Y, Zhang H, Xue C, Schmidt EM, Tang ZZ, Bizon C, Lange EM, Smith JD, Turner EH, Jun G, Kang HM, Peloso G, Auer P, Li KP, Flannick H, Zhang J, Fuchsberger C, Gaulton K, Lindgren

- C, Locke A, Manning A, Sim X, Rivas MA, Holmen OL, Gottesman O, Lu Y, Ruderfer D, Stahl EA, Duan Q, Li Y, Durda P, Jiao S, Isaacs A, Hofman A, Bis JC, Correa A, Griswold ME, Jakobsdottir J, Smith AV, Schreiner PJ, Feitosa MF, Zhang Q, Huffman JE, Crosby J, Wassel CL, Do R, Franceschini N, Martin LW, Robinson JG, Assimes TL, Crosslin DR, Rosenthal EA, Tsai M, Rieder MJ, Farlow DN, Folsom AR, Lumley T, Fox ER, Carlson CS, Peters U, Jackson RD, van Duijn CM, Uitterlinden AG, Levy D, Rotter JI, Taylor Jr HA, Gudnason V, Siscovick DS, Fornage M, Borecki IB, Hayward C, Rudan I, Chen YE, Bottinger EP, Loos RJJ, Sætrom P, Hveem K, Boehnke M, Groop L, McCarthy M, Meitinger T, Ballantyne CM, Gabriel SB, O'Donnell CJ, Post WS, North KE, Reiner AP, Boerwinkle E, Psaty MB, Altshuler D, Kathiresan S, **Lin DY**, Jarvik G, Cupples LA, Kooperberg C, Wilson JG, Nickerson DA, Abecasis GR, Rich SS, Tracy RP, Willer CJ, on behalf of the NHLBI Exome Sequencing Project. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. *American Journal of Human Genetics* 2014; 94: 233–245. PMID: PMC3928660.
163. Zeng D, **Lin DY**. Efficient estimation of semiparametric transformation models for two-phase cohort studies. *Journal of the American Statistical Association* 2014; 109: 371–383. PMID: PMC3960088.
164. Tang ZZ, **Lin DY**. Meta-analysis of sequencing studies with heterogeneous genetic associations. *Genetic Epidemiology* 2014; 38: 389–401. PMID: PMC4157393.
165. Crosby J, Peloso GM, Auer PL, Crosslin DR, Stitzel NO, Lange LA, Lu Y, Tang ZZ, Zhang H, Hindy G, Masca N, Stirrups K, Kanoni S, Do R, Jun G, Hu Y, Kang HM, Xue C, Goel A, Farrall M, Duga S, Merlini PA, Asselta R, Girelli D, Martinelli N, Yin W, Reilly D, Speliotes E, Fox C, Hveem K, Hølem O, Nikpay M, NHLBI Exome Sequencing Project, Farlow DN, Assimes TL, Franceschini N, Robinson J, North KE, Martin LW, DePristo M, Gupta N, Escher SA, Jansson JH, Van Zuydam N, Palmer CNA, Wareham N, Koch W, Meitinger T, Peters A, Lieb W, Erbel R, König IR, Kruppa J, Degenhardt F, Gottesman O, Bottinger EP, O'Donnell CJ, Psaty BM, Ballantyne CM, Abecasis G, Ordovas JM, Melander O, Watkins H, Orho-Melander M, Ardissino D, Loos RJJ, McPherson R, Willer CJ, Erdmann J, Hall AS, Samani NJ, Deloukas P, Schunkert H, Wilson JG, Kooperberg C, Rich SS, Tracy RP, **Lin DY**, Altshuler D, Gabriel S, Nickerson DA, Jarvik GP, Cupples LA, Reiner AP, Boerwinkle E, Kathiresan S. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *New England Journal of Medicine* 2014; 371: 22–31. PMID: PMC4180269.
166. Stitzel NO, Won HH, Morrison AC, Peloso GM, Do R, Lange LA, Fontanillas P, Gupta N, Duga S, Goel A, Farrall M, Saleheen D, Ferrario P, König I, Asselta R, Merlini P, Marziliano N, Notarangelo MF, Schick U, Auer P, Assimes TL, Reilly M, Wilensky R, Rader DJ, Hovingh GK, Meitinger T, Kessler T, Kastrati A, Laugwitz KL, Siscovick D, Rotter JI, Hazen S, Tracy R, Cresci S, Spertus J, Jackson R, Schwartz SM, Natarajan P, Crosby J, Muzny D, Ballantyne C, Rich SS, O'Donnell CJ, Abecasis G, Sunyaev S, Nickerson DA, Buring JE, Ridker PM, Chasman DI, Austin E, Ye Z, Kullo IJ, Weeke PE, Shaffer CM, Bastarache LA, Denny JC, Roden DM, Palmer C, Deloukas P, **Lin DY**, Tang ZZ, Erdmann J, Schunkert H, Danesh J, Marrugat J, Elosua R, Ardissino D, McPherson R, Watkins H, Reiner AP, Wilson JG, Altshuler D, Gibbs RA, Lander ES, Boerwinkle E, Gabriel S, Kathiresan S, on behalf of The Myocardial Infarction Genetics Consortium. Inactivating mutations in NPC1L1 and protection from coronary heart disease. *New England Journal of Medicine* 2014; 371: 2072–2082. PMID: PMC4335708.
167. **Lin DY**, Tao R, Kalsbeek WD, Zeng D, Gonzalez II F, Fernandez-Rhodes L, Graff M, Koch GG, North KE, Heiss G. Genetic association analysis under complex survey sampling: the Hispanic Community Health Study/Study of Latinos. *American Journal of Human Genetics* 2014; 95: 675–688. PMID: PMC4259979.
168. Hu YJ, **Lin DY**, Sun W, Zeng D. A likelihood-based framework for association analysis of allele-specific copy numbers. *Journal of the American Statistical Association* 2014; 109: 1533–1545. PMID: PMC4315366.
169. Franceschini N, Hu Y, Reiner AP, Buyske S, Nalls M, Yanek LR, Li Y, Hindorf LA, Cole SA, Howard

- BV, Andrews JS, Carty CL, Sethupathy P, Martin LW, **Lin DY**, Johnson KC, Becker LC, North KE, Dehghan A, Bis JC, Liu Y, Greenland P, Manson J, Maeda N, Garcia M, Harris TB, Becker DM, O'Donnell C, Heiss G, Kooperberg C, Boerwinkle E. Prospective associations of coronary heart disease loci in African Americans using the MetaboChip: the PAGE study. *PLOS ONE* 2014; 9: e113203. PMID: PMC4277270.
170. Hu YJ, Li Y, Auer PL, **Lin DY**. Integrative analysis of sequencing and array genotype data for discovering disease associations with rare mutations. *Proceedings of the National Academy of Sciences* 2015; 112: 1019–1024. PMID: PMC4313847.
171. Nyante SJ, Gammon MD, Kaufman JS, Bensen JT, **Lin DY**, Barnholtz-Sloan JS, Hu Y, He Q, Luo J, Millikan RC. Genetic variation in estrogen and progesterone pathway genes and breast cancer risk: an exploration of tumor subtype-specific effects. *Cancer Causes and Control* 2015; 26: 121–131. PMID: PMC4291841.
172. Zeng D, **Lin DY**. On random-effects meta-analysis. *Biometrika* 2015; 102: 281–294. PMID: PMC4681410.
173. Tao R, Zeng D, Franceschini N, North KE, Boerwinkle E, **Lin DY**. Analysis of sequence data under multivariate trait-dependent sampling. *Journal of the American Statistical Association* 2015; 110: 560–572. PMID: PMC4565625.
174. Tang ZZ, **Lin DY**. Meta-analysis for discovering rare-variant associations: statistical methods and software programs. *American Journal of Human Genetics* 2015; 97: 35–53. PMID: PMC4571037.
175. The Network and Pathway Analysis Subgroup of the Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nature Neuroscience* 2015; 18: 199–209. PMID: PMC4378867.
176. Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. Genetic differences in the immediate transcriptome response to stress predict risk-related brain function and psychiatric disorders. *Neuron* 2015; 86: 1189–1202.
177. Psychiatric Genomics Consortium and Rheumatoid Arthritis Consortium International. New data and an old puzzle: the negative association between schizophrenia and rheumatoid arthritis. *International Journal of Epidemiology* 2015; 44: 1706–1721.
178. He Q, Zhang HH, Avery CL, **Lin DY**. Sparse meta-analysis with high-dimensional data. *Biostatistics* 2016; 17: 205–220. PMID: PMC4834947.
179. Mao L, **Lin DY**. Semiparametric regression for the weighted composite endpoint of recurrent and terminal events. *Biostatistics* 2016; 17: 390–403. PMID: PMC4804115.
180. Yu B, Pulit SL, Hwang SJ, Brody JA, Amin N, Auer PL, Bis JC, Boerwinkle E, Burke GL, Chakravarti A, Correa A, Dreisbach AW, Ehret GB, Franceschini N, Hofman A, **Lin DY**, Metcalf GA, Musani SK, Muzny D, Palmas W, Raffel L, Reiner A, Rice K, Rotter JI, Veeraghavan N, Wilson JG, Fox E, Guo X, North KE, Gibbs RA, van Duijn CM, Psaty BM, Levy D, Newton-Cheh C, Morrison AC, on behalf of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium and the National Heart, Lung, and Blood Institute GO Exome Sequencing Project (ESP). Rare exome sequence variants in CLCN6 reduce blood pressure levels and hypertension risk. *Circulation: Cardiovascular Genetics* 2016; 9: 64–70. PMID: PMC4771070.
181. Zeng D, Mao L, **Lin DY**. Maximum likelihood estimation for semiparametric transformation models with interval-censored data. *Biometrika* 2016; 103: 253–271. PMID: PMC4890294.
182. Kan M, Auer PL, Wang GT, Bucayas KL, Hooker S, Rodriguez A, Li B, Ellis J, Cupples LA, Chen YDI, Dupuis J, Fox C, Gross MD, Smith JD, Heard-Costa N, Meigs JB, Pankow JS, Rotter JI, Siscovick D, Wilson JG, Shendure J, Jackson R, Peters U, Whitsel E, Zhong H, **Lin DY**, Hsu L, Franceschini N, Carlson C, Abecasis G, Gabriel S, Bamshad MJ, Altshuler D, Nickerson DA, North KE, Lange LA, Reiner AP, NHLBI-Exome Sequencing Project, Leal SM. Rare variant associations with waist-to-hip ratio in European-American and African-American women from the NHLBI-Exome Sequencing Project.

- European Journal of Human Genetics* 2016; 24: 1181–1187. PMID: PMC4970686.
183. **Lin DY**, Gong J, Gallo P, Bunn PH, Couper D. Simultaneous inference on treatment effects in survival studies with factorial designs. *Biometrics* 2016; 72: 1078–1085. PMID: PMC5026867.
 184. **Lin DY**, Dai L, Cheng G, Sailer MO. On confidence intervals for the hazard ratio in randomized clinical trials. *Biometrics* 2016; 72: 1098–1102. PMID: PMC5085885.
 185. Avery CL, Holliday KM, Chakladar S, Engeda JC, Hardy ST, Reis JP, Schreiner PJ, Shay CM, Daviglus ML, Heiss G, **Lin DY**, Zeng D. Disparities in early transitions to obesity in contemporary multi-ethnic U.S. populations. *PLOS ONE* 2016; 11: e0158025. PMID: PMC4922630.
 186. Zubair N, Graff M, Luis Ambite J, Bush WS, Kichaev G, Lu Y, Manichaikul A, Sheu WH, Absher D, Assimes TL, Bielinski SJ, Bottinger EP, Buzkova P, Chuang LM, Chung RH, Cochran B, Dumitrescu L, Gottesman O, Haessler JW, Haiman C, Heiss G, Hsiung CA, Hung YJ, Hwu CM, Juang JJ, Le Marchand L, Lee IT, Lee WJ, Lin LA, **Lin DY**, Lin SY, Mackey RH, Martin LW, Pasaniuc B, Peters U, Predazzi I, Quertermous T, Reiner AP, Robinson J, Rotter JI, Ryckman KK, Schreiner PJ, Stahl E, Tao R, Tsai MY, Waite LL, Wang TD, Buyske S, Ida Chen YD, Cheng I, Crawford DC, Loos RJJ, Rich SS, Fornage M, North KE, Kooperberg C, Carty CL. Fine-mapping of lipid regions in global populations discovers ethnic-specific signals and refines previously identified lipid loci. *Human Molecular Genetics* 2016; 25: 5500–5512. PMID: PMC5721937.
 187. Avery CL, Wassel CL, Richard MA, Highland HM, Bien S, Zubair N, Soliman EZ, Fornage M, Bielinski SJ, Tao R, Seyerle AA, Shah SJ, Lloyd-Jones DM, Buyske S, Rotter JI, Post WS, Rich SS, Hindorff LA, Jeff JM, Shohet RV, Sotoodehnia N, **Lin DY**, Whitsel EA, Peters U, Haiman CA, Crawford DC, Kooperberg C, North KE. Fine mapping of QT interval regions in global populations refines previously identified QT interval loci and identifies signals unique to African and Hispanic descent populations. *Heart Rhythm* 2017; 14: 572–580. PMID: PMC5448160.
 188. Mao L, **Lin DY**. Efficient estimation of semiparametric transformation models for the cumulative incidence of competing risks. *Journal of the Royal Statistical Society, Series B* 2017; 79: 573–587. PMID: PMC5319638.
 189. Hardy ST, Holliday KM, Chakladar S, Engeda JC, Allen NB, Heiss G, Lloyd-Jones DM, Schreiner PJ, Shay CM, **Lin DY**, Zeng D, Avery CL. Heterogeneity in blood pressure transitions over the life course: Age-specific emergence of racial/ethnic and sex disparities in the United States. *Journal of the American Medical Association Cardiology* 2017; 2: 653–661. PMID: PMC5634332.
 190. Tang ZZ, Bunn P, Tao R, Liu Z, **Lin DY**. PreMeta: a tool to facilitate meta-analysis of rare variant associations. *BMC Genomics* 2017; 18: 160. PMID: PMC5310051.
 191. Holliday KM, **Lin DY**, Chakladar S, Castañeda SF, Daviglus ML, Evenson KR, Marquez DX, Qi Q, Shay CM, Sotres-Alvarez D, Vidot DC, Zeng D, Avery CL. Targeting physical activity interventions for adults: When should intervention occur? *Preventive Medicine* 2017; 97: 13–18. PMID: PMC5337155.
 192. Mao L, **Lin DY**, Zeng D. Semiparametric regression analysis of interval-censored competing risks data. *Biometrics* 2017; 73: 857–865. PMID: PMC5561531.
 193. Zeng D, Gao F, **Lin DY**. Maximum likelihood estimation for semiparametric regression models with multivariate interval-censored data. *Biometrika* 2017; 104: 505–525. PMID: PMC5787874.
 194. Tao R, Zeng D, **Lin DY**. Efficient semiparametric inference under two-phase sampling, with applications to genetic association studies. *Journal of the American Statistical Association* 2017; 112: 1468–1476. PMID: PMC5823539.
 195. Gao F, Zeng D, **Lin DY**. Semiparametric estimation of the accelerated failure time model with partly interval-censored data. *Biometrics* 2017; 73: 1161–1168. PMID: PMC5785785.
 196. Sun W, Bunn P, Jin C, Little P, Zhabotynsky V, Perou CM, Hayes DN, Chen M, **Lin DY**. The association between copy number aberration, DNA methylation and gene expression in tumor samples. *Nucleic Acids Research* 2018; 46: 3009–3018. PMID: PMC5887505.
 197. Gong J, Nishimura K, Fernandez-Rhodes L, Haessler J, Bien S, Graff M, Lim U, Lu Y, Gross M,

- Fornage M, Yoneyama S, Isasi C, Buzkova P, Daviglus ML, **Lin DY**, Tao R, Goodloe R, Bush W, Farber-Eger E, Boston J, Dilks H, Ehret G, Gu C, Lewis C, Nguyen KD, Cooper R, Leppert M, Irvin M, Bottinger E, Wilkens L, Haiman C, Park L, Monroe K, Cheng I, Stram D, Carlson C, Jackson R, Kuller L, Houston D, Kooperberg C, Buyske S, Hindorff L, Crawford D, Loos R, Le Marchand L, Matisse T, North K, Peters U. Trans-ethnic analysis of MetaboChip data identifies two new loci associated with BMI. *International Journal of Obesity* 2018; 42: 384–390. PMID: PMC5876082.
198. Zeng D, Pan J, Hu K, Chi E, **Lin DY**. Improving the power to establish clinical similarity in a phase 3 efficacy trial by incorporating prior evidence of analytical and pharmacokinetic similarity. *Journal of Biopharmaceutical Statistics* 2018; 28: 320–332. PMID: PMC5909990.
199. Engeda, JC, Holliday KM, Hardy ST, Chakladar S, **Lin DY**, Talavera GA, Howard BV, Daviglus ML, Pirzada A, Schreiner PJ, Zeng D, Avery CL. Transitions from ideal to intermediate cholesterol levels may vary by cholesterol metric. *Scientific Reports* 2018; 8: 2782. PMID: PMC5807429.
200. Wong KY, Zeng D, **Lin DY**. Efficient estimation for semiparametric structural equation models with censored data. *Journal of the American Statistical Association* 2018; 113: 893–905. PMID: PMC6075718.
201. Hodonsky CJ, Schurmann C, Schick UM, Kocarnik J, Tao R, van Rooij FJA, Wassel C, Buyske S, Fornage M, Hindorff LA, Floyd JS, Ganesh SK, **Lin DY**, North KE, Reiner AP, Loos R, Kooperberg C, Avery CL. Generalization and fine mapping of red blood cell trait genetic associations to multi-ethnic populations: The PAGE Study. *American Journal of Hematology* 2018; 93: 1061–1073. PMID: PMC6300146.
202. Gao F, Zeng D, **Lin DY**. Semiparametric regression analysis of interval-censored data with informative dropout. *Biometrics* 2018; 74: 1213–1222. PMID: PMC6309250.
203. Wong KY, Fan C, Tanioka M, Parker JS, Nobel AB, Zeng D, **Lin DY**, Perou CM. I-BOOST: an integrative boosting approach for predicting survival time with multiple genomics platforms. *Genome Biology* 2019; 20: 52. PMID: PMC6404283.
204. **Lin DY**. A simple and accurate method to determine genomewide significance for association tests in sequencing studies. *Genetic Epidemiology* 2019; 43: 365–372. PMID: PMC6520182.
205. Lin BM, Nadkarni GN, Tao R, Graff M, Fornage M, Buyske S, Matisse TC, Highland H, Wilkens LR, Carlson CS, Park SL, Setiawan VW, Ambite JL, Heiss G, Boerwinkle E, **Lin DY**, Morris AP, Loos R, Kooperberg C, North K, Wassel C, Franceschini N. Genetics of chronic kidney disease stages across ancestries: The PAGE study. *Frontiers in Genetics* 2019; 10: 494. PMID: PMC6544117.
206. Wojcik GL,* Graff M,* Nishimura KK,* Tao R,* Haessler J,* Gignoux CR,* Highland HM,* Patel YM,* Sorokin EP, Avery CL, Belbin GM, Bien SA, Cheng I, Cullina S, Hodonsky CJ, Hu Y, Huckins LM, Jeff J, Justice AE, Kocarnik JM, Lim U, Lin BM, Lu Y, Nelson SC, Park SSL, Poisner H, Preuss MH, Richard MA, Schurmann C, Setiawan VW, Sockell A, Vahi K, Verbanck M, Vishnu A, Walker RW, Young KL, Zubair N, Acuna-Alonso V, Ambite JL, Barnes KC, Boerwinkle E, Bottinger E, Bustamante CD, Caberto C, Canizales-Quintero S, Conomos MP, Deelman E, Do R, Doheny K, Fernandez-Rhodes L, Fornage M, Hailu B, Heiss G, Henn B, Hindorff LA, Jackson RD, Laurie CA, Laurie CC, Li Y, **Lin DY**, Moreno-Estrada A, Nadkarni G, Norman PJ, Pooler LC, Reiner AP, Romm J, Sabati C, Sandoval K, Sheng X, Stahl EA, Stram DO, Thornton TA, Wassel CL, Wilkens LR, Winkler CA, Yoneyama S, Buyske S, Haiman CA, Kooperberg C, Le Marchand L, Loos R, Matisse TC, North KE, Peters U, Kenny EE, Carlson CS. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 2019; 570: 514–518. *Shared first authorship. (Dr. Tao is Dr. Lin's Ph.D. student.) PMID: PMC6785182.
207. Raulerson CK, Ko A, Kidd JC, Currin KW, Brotman SM, Cannon ME, Wu Y, Spracklen CN, Jackson AU, Stringham HM, Welch RP, Fuchsberger C, Locke AE, Narisu N, Lusi AJ, Civelek M, Furey TS, Kuusisto J, Collins FS, Boehnke M, Scott LJ, **Lin DY**, Love MI, Laakso M, Pajukanta P, Mohlke KL. Adipose tissue gene expression associations reveal hundreds of candidate genes for cardiometabolic

- traits. *American Journal of Human Genetics* 2019; 105: 773–787. PMID: PMC6817527.
208. Gao F, Zeng D, Couper D, **Lin DY**. Semiparametric regression analysis of multiple right- and interval-censored events. *Journal of the American Statistical Association* 2019; 114: 1232–1240. PMID: PMC6777710.
209. Deng Q, Bai X, Liu D, Roy D, Ying Z, **Lin DY**. Power and sample size for dose-finding studies with survival endpoints under model uncertainty. *Biometrics* 2019; 75: 308–314. PMID: PMC6411454.
210. Palace J, **Lin DY**, Zeng D, Majed M, Elson L, Hamid S, Messina S, Misu T, Sagen J, Whittam D, Yoshiki T, Leite MI, Weinshenker B, Cabre P, Jacob A, Nakashima I, Fujihara K, Pittock SJ. Outcome prediction models in AQP4-IgG positive neuromyelitis optica spectrum disorders. *Brain* 2019; 142: 1310–1323. PMID: PMC6487334.
211. Little PL, **Lin DY**, Sun W. Associating somatic mutations to clinical outcomes: a pan-cancer study of survival time. *Genome Medicine* 2019; 11: 37. PMID: PMC6540540.
212. Wong KY, Zeng D, **Lin DY**. Robust score tests with missing data in genomics studies. *Journal of the American Statistical Association* 2019; 114: 1778–1786. PMID: PMC6951249.
213. McCabe SD, **Lin DY**, Love ML. Consistency and overfitting of multi-omics methods on experimental data. *Briefings in Bioinformatics* 2020; 21: 1277–1284. PMID: PMC7373174.
214. Zeng D, Pan Z, **Lin DY**. Design and analysis of bridging studies with prior probabilities on the null and alternative hypotheses. *Biometrics* 2020; 76: 224–234. PMID: PMC7983026.
215. Diao G, **Lin DY**. Statistically efficient association analysis of quantitative traits with haplotypes and untyped SNPs in family studies. *BMC Genetics* 2020; 21: 1–11. PMID: PMC7487716.
216. **Lin DY**, Zeng D, Couper D. A general framework for integrative analysis of incomplete multi-omics data. *Genetic Epidemiology* 2020; 44: 646–664. PMID: PMC7951090.
217. Wang R, **Lin DY**, Jiang Y. SCOPE: a normalization and copy number estimation method for single-cell DNA sequencing. *Cell Systems* 2020; 10: 445–452.e6. PMID: PMC7250054.
218. Tao R, Zeng D, **Lin DY**. Optimal designs of two-phase studies. *Journal of the American Statistical Association* 2020; 115: 1946–1959. PMID: PMC7954143.
219. Hodonsky CJ, Baldassari AR, Bien SA, Raffield LM, Highland HM, Sitlani C, Wojcik GL, Tao R, Graff M, Tang W, Thyagarajan B, Buyske S, Fornage M, Hindorff LA, Li Y, **Lin DY**, Reiner AP, North KE, Loos RJJ, Kooperberg CL, Avery CL. Ancestry-specific associations identified in genome-wide combined-phenotype study of red blood cell traits emphasize benefits of diversity in genomics. *BMC Genomics* 2020; 21: 228. PMID: PMC7071748.
220. Lin BM, Grinde KE, Brody JA, Breeze CE, Raffield LM, Mychaleckyj JC, Thornton TA, Perry JA, Baier LJ, Fuentes LL, Guo X, Heavner BD, Hanson RL, Hung YJ, Qian H, Hsiung CA, Hwang SJ, Irvin MR, Jain D, Kelly TN, Kobes S, Lange L, Lash JP, Li Y, Liu X, Mi X, Musani SK, Papanicolaou GJ, Parsa A, Reiner AP, Salimi S, Sheu WHH, Shuldiner AR, Taylor KD, Smith AV, Smith JA, Tin A, Vaidya D, Wallace RB, Yamamoto K, Sakaue S, Matsuda K, Kamatani Y, Momozawa Y, Yanek LR, Young BA, Zhao W, Okada Y, Abecasis G, Psaty BM, Arnett DK, Boerwinkle E, Cai J, Chen IYD, Correa A, Cupples LA, He J, Kardia SLR, Kooperberg C, Mathias RA, Mitchell BD, Nickerson DA, Turner ST, Ramachandran VS, Rotter JI, Levy D, Kramer HJ, Köttgen A, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, TOPMed Kidney Working Group, Rich SS, **Lin DY**, Browning SR, Franceschini N. Whole genome sequence analyses of eGFR in 23,732 people representing multiple ancestries in the NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium. *EBIOMed* 2021; 63: 103157. PMID: PMC7804602.
221. **Lin DY**, Zeng D, Eron JJ. Evaluating the efficacy of therapies in COVID-19 patients. *Clinical Infectious Diseases* 2021; 72: 1093–1100. PMID: PMC7454372.
222. Hu Y, Bien SA, Nishimura KK, Haessler J, Hodonsky CJ, Baldassari AR, Highland HM, Wang Z, Preuss M, Sitlani CM, Wojcik GL, Tao R, Graff M, Huckins LM, Sun Q, Chen MH, Mousas A, Auer PL, Lettre G, Tang W, Qi L, Thyagarajan B, Buyske S, Fornage M, Hindorff LA, Li Y, **Lin DY**, Reiner AP, North

- KE, Loos RJF, Raffield LM, Peters U, Avery CL, Kooperberg C. Multi-ethnic genome-wide association analyses of white blood cell and platelet traits in the Population Architecture using Genomics and Epidemiology (PAGE) study. *BMC Genomics* 2021; 22: 432. PMID: PMC8191001.
223. Jin C, Chen M, **Lin DY**, Sun W. Cell-type-aware analysis of RNA-seq data. *Nature Computational Science* 2021; 1: 253–261. PMID: PMC8697413.
224. **Lin DY**, Zeng D, Mehrotra DV, Corey L, Gilbert PB. Evaluating the efficacy of Coronavirus Disease 2019 vaccines. *Clinical Infectious Diseases* 2021; 73: 1540–1544. PMID: PMC7799296.
225. Zeng D, **Lin DY**. Maximum likelihood estimation for semiparametric regression models with panel count data. *Biometrika* 2021; 108: 947–963. PMID: PMC8691743.
226. **Lin DY**, Zeng D, Gilbert PB. Evaluating the long-term efficacy of Coronavirus Disease 2019 (COVID-19) vaccines. *Clinical Infectious Diseases* 2021; 73: 1927–1939. PMID: PMC7989522.
227. Raffield LM, Howard AG, Graff M, **Lin DY**, Cheng S, Demerath E, Ndumele C, Palta P, Rebholz C, Seidemann S, Yu B, Gordon-Larsen P, North KE, Avery C. Obesity duration, severity, and distribution trajectories and cardiovascular disease risk in the Atherosclerosis Risk in Communities Study. *Journal of the American Heart Association* 2021; 10: e019946. PMID: PMC9075238.
228. Quintanilha JCF, Racioppi A, Wang J, Ethridge A, Denning S, Peña C, Skol A, Crona D, **Lin DY**, Innocenti F. PIK3R5 genetic predictors of hypertension induced by VEGF-pathway inhibitors. *The Pharmacogenomics Journal* 2022; 22: 82–88. PMID: PMC8799524.
229. Quintanilha JCF, Wang J, Sibley AB, Xu W, Espin-Garcia O, Jiang C, Etheridge AS, Ratain MJ, Lenz HJ, Bertagnolli M, Kindler HL, Dickler MN, Venook A, Liu G, Owzar K, **Lin DY**, Innocenti F. Genome-wide association studies of survival in 1,520 cancer patients treated with bevacizumab-containing regimens. *The International Journal of Cancer* 2022; 150: 279–289. PMID: PMC8627468.
230. **Lin DY**, Gu Y, Zeng D, Janes HE, Gilbert PB. Evaluating vaccine efficacy against SARS-CoV-2 infection. *Clinical Infectious Diseases* 2022; 74: 544–552. PMID: PMC8057249.
231. Wong KY, Zeng D, **Lin DY**. Semiparametric latent-class models for multivariate longitudinal and survival data. *The Annals of Statistics* 2022; 50: 487–510. PMID: PMC9269993.
232. **Lin DY**, Baden LR, El Sahly HM, Essink B, Neuzil KM, Corey L, Miller J, for the COVE Study Group. Durability of protection against symptomatic COVID-19 among participants of the mRNA-1273 SARS-CoV-2 vaccine trial. *JAMA Network Open* 2022; 5: e2215984. PMID: PMC9178430.
233. **Lin DY**, Gu Y, Wheeler B, Young H, Holloway S, Sunny SK, Moore Z, Zeng D. Effectiveness of Covid-19 vaccines over a 9-month period in North Carolina. *New England Journal of Medicine* 2022; 386: 933–941. PMID: PMC8781317.
234. Highland HM, Wojcik GL, Graff M, Nishimura KK, Hodonsky CJ, Baldassari AR, Cheng I, Gignoux CR, Tao R, Li Y, Boerwinkle E, Fornage M, Haessler J, Hindorff LA, Hu Y, Justice AE, Lin BM, **Lin DY**, Stram DO, Haiman CA, Kooperberg C, Le Marchand L, Matise TC, Kenny EE, Carlson CS, Stahl EA, Avery CL, North KE, Ambite JL, Buyske S, Loos RJ, Peters U, Young KL, Bien SA, Huckins LM. Predicted gene expression in ancestrally diverse populations leads to discovery of novel susceptibility loci for lifestyle and cardiometabolic traits. *The American Journal of Human Genetics* 2022; 109: 669–679. PMID: PMC9069067.
235. Wang R, **Lin DY**, Jiang Y. EPIC: inferring relevant cell types for complex traits by integrating genome-wide association studies and single-cell RNA sequencing. *PLOS Genetics* 2022; 18: e1010251. PMID: PMC9242467.
236. Wang J, Zeng D, **Lin DY**. Semiparametric single-index models for optimal treatment regimens with censored outcomes. *Lifetime Data Analysis* 2022; 28: 744–763. PMID: PMC10349573.
237. **Lin DY**, Gu Y, Xu Y, Zeng D, Wheeler B, Young H, Sunny SK, Moore Z. Effects of vaccination and previous infection on omicron infections in children. *New England Journal of Medicine* 2022; 387: 1141–1143. PMID: PMC9511630.
238. **Lin DY**, Gu Y, Xu Y, Wheeler B, Young H, Sunny SK, Moore Z, Zeng D. Association of primary and

- booster vaccination and prior infection with SARS-CoV-2 infection and severe COVID-19 outcomes. *JAMA* 2022; 328: 1415–1426. PMID: PMC9513711.
239. **Lin DY**, Zeng D, Gu Y, Krause P, Fleming TR. Reliably assessing duration of protection for COVID-19 vaccines. *The Journal of Infectious Diseases* 2022; 226: 1863–1866. PMID: PMC9383791.
240. Young KL, Fisher V, Deng X, Brody JA, Graff M, Lim E, Lin BM, Xu H, Amin N, An P, Aslibekyan S, Fohner AE, Hidalgo B, Lenzini P, Kraaij R, Medina-Gomez C, Prokić I, Rivadeneira F, Sitlani C, Tao R, van Rooij J, Zhang D, Broome JG, Buth EJ, Heavner BD, Jain D, Smith AV, Barnes K, Boorgula MP, Chavan S, Darbar D, De Andrade M, Guo X, Haessler J, Irvin MR, Kalyani RR, Kardia SLR, Kooperberg C, Kim W, Mathias RA, McDonald ML, Mitchell BD, Peyser PA, Regan EA, Redline S, Reiner AP, Rich SS, Rotter JI, Smith JA, Weiss S, Wiggins KL, Yanek LR, Arnett D, Heard-Costa NL, Leal S, **Lin DY**, McKnight B, Province M, van Duijn CM, North KE, Cupples LA, Liu CT. Whole-exome sequence analysis of anthropometric traits illustrates challenges in identifying effects of rare genetic variants. *HGG Advances* 2023; 4: 100163. PMID: PMC9772568.
241. Wong KY, Zeng D, **Lin DY**. Penalized regression for multiple types of many features with missing data. *Statistica Sinica* 2023; 33: 633–662. PMID: PMC10187615.
242. Kidd J, Raulerson CK, Mohlke KL, **Lin DY**. Meditation analysis of multiple mediators with incomplete omics data. *Genetic Epidemiology* 2023; 47: 61–77. PMID: PMC10423053.
243. Wang X, Zein J, Ji X, **Lin DY**. Impact of vaccination, prior infection and therapy on Omicron Infection and mortality. *The Journal of Infectious Diseases* 2023; 227: 970–976. PMID: PMC10319976.
244. **Lin DY**, Xu Y, Gu Y, Zeng D, Wheeler B, Young H, Sunny SK, Moore Z. Effectiveness of bivalent boosters against severe omicron infection. *New England Journal of Medicine* 2023; 388: 764–766. PMID: PMC9933929.
245. **Lin DY**, Xu Y, Gu Y, Zeng D, Sunny SK, Moore Z. Durability of bivalent boosters against omicron subvariants. *New England Journal of Medicine* 2023; 388: 1818–1820. PMID: PMC10120009.
246. Little P, Liu S, Zhabotynsky V, Li Y, **Lin DY**, Sun W. A computational method for Cell type-specific Expression Quantitative Trait Loci mapping using bulk RNA-seq data. *Nature Communications* 2023; 14: 3030. PMID: PMC10212972.
247. Yang H, **Lin DY**, Li Q. An efficient greedy search algorithm for high-dimensional linear discriminant analysis. *Statistica Sinica* 2023; 33: 1343–1364. PMID: PMC10348717.
248. Xu Y, Zeng D, **Lin DY**. Marginal proportional hazards models for multivariate interval-censored data. *Biometrika* 2023; 110: 815–830. PMID: PMC10434824.
249. **Lin DY**, Fleming TR. VV116 versus nirmatrelvir-ritonavir for oral treatment of Covid-19. *New England Journal of Medicine* 2023; 388: 2395–2397.
250. Wang X, Abi Fadel F, Huang S, Sacha GL, Bartley P, Duggal A, **Lin DY**. Effectiveness of bebtelovimab against hospitalization and death from the Omicron BA.4/BA.5 and BQ.1/BQ.1.1 subvariants. *Chest* 2023; 164: 872–874. PMC Journal – In Process.
251. **Lin DY**, Abi Fadel F, Huang S, Milinovich AT, Sacha GL, Bartley P, Duggal A, Wang X. Nirmatrelvir or molnupiravir use and severe outcomes from Omicron infections. *JAMA Network Open* 2023; 6: e2335077. PMID: PMC10514733.
252. **Lin DY**, Xu Y, Zeng D, Sunny SK. A cost-benefit analysis of bivalent Covid-19 vaccines. *Journal of Biotechnology and Biomedicine* 2023; 6: 551–553.
253. **Lin DY**, Xu Y, Gu Y, Zeng D, Wheeler B, Young H, Moore Z, Sunny SK. Effects of COVID-19 vaccination and previous SARS-CoV-2 infection on Omicron infection and severe outcomes in children under 12 years of age in the USA: an observational cohort study. *The Lancet Infectious Diseases* 2023; 23: 1257–1265. PMID: PMC10275621.
254. Li Y, Wong KY, Howard AG, Gordon-Larsen P, Highland HM, Graff M, North KE, Downie CG, Avery CL, Yu B, Young KL, Buchanan VL, Kaplan R, Hou L, Joyce BT, Qi Q, Sofer T, Moon JY, **Lin DY**. Mendelian randomization with incomplete measurements on the exposure in the Hispanic Community

Health Study/Study of Latinos. *Human Genetics and Genomics Advances* 2023; 5: 100245. PMID: PMC10628889.

255. Paritala S, Xu Y, Du Y, Donahue M, Maloney P, **Lin DY**. Effectiveness of bivalent boosters over nine and half months. *Journal of Biotechnology and Biomedicine* 2023; 6: 585–589.

IN-PRESS ARTICLES

256. Xu Y, Zeng D, **Lin DY**. Proportional rates models for multivariate panel count data. *Biometrics*.
257. Gu Y, Zeng D, Heiss G, **Lin DY**. Maximum likelihood estimation for semiparametric regression models with interval-censored multistate data. *Biometrika*.

SUBMITTED MANUSCRIPTS

258. Zeng D, Kosorok MR, **Lin DY**. Nonparametric maximum likelihood estimation in linear regression with unknown transformation.
259. Kidd J, **Lin DY**. Improving the power to detect indirect effects in mediation analysis.
260. Wang J, Zeng D, **Lin DY**. Fitting the Cox proportional hazards model to big data.
261. **Lin DY**, Wang J, Anstrom KJ, LaVange LM, Wen J, Bozzette SA, Powderly WG. Efficacy of infliximab, abatacept, and cenicriviroc for the treatment of adults hospitalized with COVID-19.
262. **Lin DY**, Xu Y, Gu Y, Sunny SK., Moore Z, Zeng D. Impact of booster vaccination interval on SARS-CoV-2 infection, hospitalization, and death.
263. **Lin DY**, Wang J, Gu Y, Zeng D. Evaluating treatment efficacy in hospitalized Covid-19 patients, with applications to Adaptive COVID-19 Treatment Trials.
264. Gu Y, Zeng D, **Lin DY**. Semiparametric regression models for interval-censored multi-state data with an absorbing state.
265. Sizelove R, Zeng D, **Lin DY**. Semiparametric linear regression with an interval-censored covariate in the Atherosclerosis Risk in Communities Study.
266. Shrestha P, Graff M, Gu Y, Wang Y, Avery CL, Ginnis J, Simancas-Pallares MA, Ferreira Zandoná AG, Ahn HS, Nguyen K, **Lin DY**, Preisser JS, Slade GD, Marazita ML, North KE, Divaris K. Multi-ancestry genome-wide association study of early childhood caries.
267. Koenigsberg SH, Makarem N, Highland HM, Downie CG, Alotaibi M, Cheng S, Howard AG, **Lin DY**, North KE, Patel SR, Thyagarajan B, Avery CL. Insomnia is associated with fatty acid and oxylipin inflammatory biomarkers in the Hispanic Community Health Study/Study of Latinos.
268. **Lin DY**, Du Y, Xu Y, Paritala S, Donahue M, Maloney P. Updated XBB.1.5 vaccines and SARS-CoV-2 infections from Omicron subvariants.
269. **Lin DY**, Huang S, Milinovich A, Duggal A, Wang X. Effectiveness of XBB.1.5 vaccines and antiviral drugs against severe outcomes of omicron infection.
270. Li Y, Wong KY, Howard AG, Gordon-Larsen G, Highland HM, Graff M, North KE, Downie CG, Avery CL, Yu B, Young KL, Buchanan VL, Kaplan R, Hou L, Joyce BT, Qi Q, Sofer T, Moon JY, **Lin DY**. Multivariable Mendelian randomization with incomplete measurements on the exposure variables in the Hispanic Community Health Study/Study of Latinos.

BOOKS

271. **Lin DY**, Fleming TR. *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*. Springer-Verlag, 1997.
272. **Lin DY**, Heagerty P. *Proceedings of the Second Seattle Symposium in Biostatistics: Analysis of Correlated Data*. Springer-Verlag, 2004.

OTHER PUBLICATIONS (PROCEEDINGS, DISCUSSIONS, BOOK CHAPTERS)

273. Brown MB, **Lin DY**, Normolle DP, Herzog AR. Models for and the effect of nonresponse in a survey of the noninstitutionalized elderly. In *The 1988 Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 1988, pp. 615–620.
274. **Lin DY**, Wei LJ. Discussion of C Jennison and BW Turnbull, “Interim analyses: the repeated confidence interval approach.” *Journal of the Royal Statistical Society, Series B* 1989; 51: 347–348.
275. **Lin DY**, Wei LJ. Discussion of JH Ware, “Investigating therapies of potentially great benefit: ECMO.” *Statistical Science* 1989; 4: 324–325.
276. **Lin DY**, Wei LJ. Discussion of A Agresti, “A survey of exact inference for contingency tables.” *Statistical Science* 1992; 7: 166–167.
277. **Lin DY**. Multivariate failure time data. In *Recent Advances in the Designs and Analysis of Clinical Trials*. P. Thall, ed. 1995; pp. 73–93. Kluwer Academic Publishers.
278. **Lin DY**. Marginal models for multivariate survival data. In *Encyclopedia of Biostatistics*. P. Armitage, T. Colton, eds. 1998; pp. 2407–2412. John Wiley & Sons.
279. **Lin DY**, Ying Z. Additive hazards regression models for survival data. In *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*. D.Y. Lin, T.R. Fleming, eds. 1997; pp. 185–198. Springer-Verlag.
280. **Lin DY**, Kronmal RA, Psaty BM. Reply to the letter to the editor. *Biometrics* 1999; 55: 990.
281. **Lin DY**. Survival analysis. In *Advanced Medical Statistics*. Y. Lu, J. Q. Fan, eds. 2003; pp. 815–835. World Scientific Publishing Co.
282. **Lin DY**. A note on permutation tests in multistage association scans—Reply to Dudbridge. *American Journal of Human Genetics* 2006; 78: 1096.
283. Zeng D, **Lin DY**. Discussion of P. Diggle, D. Farewell and R. Henderson, “Analysis of longitudinal data with drop-out: objectives, assumptions and a proposal.” *Journal of the Royal Statistical Society, Series C* 2007; 56: 544–545.
284. **Lin DY**, Huang BE. The use of inferred haplotypes in downstream analysis—Reply to P. Kraft and D.O. Stram. *American Journal of Human Genetics* 2007; 81: 865–866.
285. Gijbels I, **Lin DY**, Ying Z. Non- and semi-parametric analysis of failure time data with missing failure indicators. In *Complex Datasets and Inverse Problems*. R. Liu, W. Strawderman, C.H. Zhang, eds. 2007; pp. 203–223, IMS Lecture Notes—Monograph Series, Ohio.
286. Marchenko YV, Carroll RJ, **Lin DY**, Amos CI. Semiparametric analysis of case-control genetic data in the presence of environmental factors. *STATA Journal* 2008; 8: 305–333.
287. **Lin DY**, Hu Y. Reply to Marchini and Howie. *American Journal of Human Genetics* 2008; 83: 539–540. PMID: PMC2561930.
288. **Lin DY**. Genetic association analysis. In *Encyclopedia of Clinical Trials*. R. D’Agostino, L. Sullivan, J. Massaro, eds. 2008. John Wiley & Sons.
289. **Lin DY**. Transformation models. In *Handbook of Survival Analysis*. J. P. Klein, H. C. van Houwelingen, J. G. Ibrahim, T. H. Scheike, eds. 2014. Chapman and Hall.
290. **Lin DY**. Discussion of the paper by R. L. Prentice and Y. Huang: Optimal designs and efficient inference for biomarker studies. *Statistical Theory and Related Fields* 2018; 2: 21–22.
291. Zeng D, **Lin DY**. Maximum likelihood estimation for case-cohort and nested case-control studies. In *Handbook of Statistical Methods for Case-Control Studies*. O. Borgan, N. Breslow, N. Chatterjee, M. H. Gail, A. Scott, C. J. Wild, eds. 2018; pp. 391–404. Chapman and Hall/CRC.
292. Diao G, Zeng D, **Lin DY**. Analysis of secondary phenotype data under case-control designs. In *Handbook of Statistical Methods for Case-Control Studies*. O. Borgan, N. Breslow, N. Chatterjee, M. H. Gail, A. Scott, C. J. Wild, eds. 2018; pp. 515–528. Chapman and Hall/CRC.

293. **Lin DY**. Review of the book “*The Statistical Analysis of Multivariate Failure Time Data: A Marginal Modeling Approach*” by R. L. Prentice and S. Zhao. *Biometrics* 2019; 75: 1416–1417.
294. **Lin DY**. Response to the Letter to the Editor by Ladhani et al on “More on omicron infections in children.” *New England Journal of Medicine* 2022; 387: 1911–1912.
295. **Lin DY**, Zeng D. Maximum likelihood estimation of semiparametric regression models with interval-censored data. In *Emerging Topics in Modeling Interval-Censored Survival Data*. J. Sun, D. G. Chen, eds. 2002; pp. 107–122. Springer.
296. **Lin DY**. Response to the Letter to the Editor by Arbel et al on “Bivalent boosters against severe omicron infection”. *New England Journal of Medicine* 2023; 388: 1726–1727.

INVITED PRESENTATIONS IN CONFERENCES

1. “Sequential log rank tests adjusting for covariates with the accelerated life model,” The International Biometric Society ENAR Spring Meeting, Houston, TX, 1991.
2. “Evaluating the role of CD4 change as a surrogate endpoint in AIDS clinical trials,” The International Biometric Society WNAR Meeting, Santa Barbara, CA, 1991.
3. “Analysis of the accelerated failure time model” (short course), The 46th Annual Conference on Applied Statistics, Atlantic City, NJ, 1990.
4. “Regression analysis of multivariate failure time data” (short course), The 49th Annual Conference on Applied Statistics, Atlantic City, NJ, 1993.
5. “Cox regression analysis with incomplete covariate measurements,” Joint Statistical Meetings, San Francisco, CA, 1993.
6. “Semiparametric analysis of general hazard-based survival models,” The International Biometric Society ENAR Spring Meeting, Cleveland, OH, 1994.
7. “Checking the marginal Cox model for correlated failure time data,” International Research Conference on Lifetime Data Models in Reliability and Survival Analysis, Boston, MA, 1994.
8. “Monitoring medical studies with survival endpoints,” Third World Congress of the Bernoulli Society and 57th Annual Meeting of the Institute of Mathematical Statistics, Chapel Hill, NC, 1994.
9. “Analysis of multivariate failure time data: Marginal modeling approaches,” XVIIth International Biometric Conference, Hamilton, Canada, 1994.
10. “On modeling the marginal distributions of multivariate failure times,” The International Biometric Society ENAR Spring Meeting, Birmingham, AL, 1995.
11. “On the analysis of competing risks data,” Harvard School of Public Health/Schering-Plough Workshop on Biostatistical Perspectives in Clinical and Pharmaceutical Research. Boston, MA, 1995.
12. “Analysis of multiple events data” (short course), ICSA 1997 Applied Statistics Symposium, Rutgers University, NJ, 1997.
13. “Estimating the proportion of treatment effect explained by a surrogate marker,” Joint Statistical Meetings, Anaheim, CA, 1997.
14. “Checking the adequacy of semiparametric regression models for current status data,” The International Biometric Society ENAR Spring Meeting, Pittsburgh, PA, 1998.
15. “An update on semiparametric additive hazards regression,” and “Nonparametric inferences for the gap time distributions of serial events,” Workshop on Event History Analysis. Montreal, Canada, 1998.
16. “Additive hazards regression for censored data,” IVth Taipei International Statistical Symposium, Taipei, Taiwan, 1998.
17. “Regression analysis of censored medical costs,” The International Biometric Society ENAR Spring Meeting, Chicago, IL, 2000.
18. “Semiparametric failure-time regression with mismeasured or missing covariates,” The International

- Biometric Society ENAR Spring Meeting, Charlotte, NC 2001.
19. "Regression analysis of incomplete medical cost data," 22nd Annual Conference of the International Society for Clinical Biostatisticians, Stockholm, Sweden, 2001.
 20. "Censoring and cost-effectiveness analysis," The International Biometric Society ENAR Spring Meeting, Arlington, VA, 2002.
 21. "Analysis of multivariate failure time data," INSERM Workshop, La Londe-Les-Maures, France, 2004.
 22. "Checking the Cox proportional hazards model with cumulative sums of martingale residuals," INSERM workshop, La Londe-Les-Maures, France, 2004.
 23. "Maximum likelihood estimation in transformation models for counting processes," Joint Statistical Meetings, Toronto, Canada, 2004.
 24. "Analysis of multivariate failure time data" (short course), FDA/Industry Workshop, Washington, D.C., 2004.
 25. "Efficient semiparametric estimation of haplotype-disease association in cohort studies," Joint Statistical Meetings, Minneapolis, MN, 2005.
 26. "Maximum likelihood estimation in semiparametric models with censored data," Workshop on Statistical Analysis of Complex Event History Data, Oslo, Norway, 2005.
 27. "Statistical analysis of haplotype-disease association," Joint Statistical Meetings, Seattle, WA, 2006.
 28. "Semiparametric transformation models with random effects for highly stratified survival data," Joint Statistical Meetings, Seattle, WA, 2006.
 29. "Analysis of multivariate failure time data" (short course), Joint Statistical Meetings, Seattle, WA, 2006.
 30. "Maximum likelihood estimation in semiparametric regression models with censored data," 12th Army Conference on Applied Statistics, RTP, NC, 2006.
 31. "Maximum likelihood estimation in semiparametric regression models with censored data," Ordinary Meeting, Royal Statistical Society, London, Great Britain, 2007.
 32. "Likelihood-based inference on haplotype-disease association," The International Biometric Society ENAR Spring Meeting, Atlanta, 2007.
 33. "Statistical analysis of haplotype-disease associations using HAPSTAT" (tutorial), The International Biometric Society ENAR Spring Meeting, Atlanta, 2007.
 34. "Analysis of multivariate failure time data" (short course), ICSA 2007 Applied Statistics Symposium, Raleigh, NC.
 35. "Analysis of censored data" (short course), Joint Statistical Meetings, Salt Lake City, 2007.
 36. "HAPSTAT: Statistical analysis of haplotype-disease association," Interface 2008, Durham, North Carolina, 2008.
 37. "Analysis of multivariate failure time data" (short course), Joint Statistical Meetings, Denver, Colorado, 2008.
 38. "Estimating genetic effects and gene-environment interactions with missing data," The International Biometric Society ENAR Spring Meeting, San Antonio, 2009.
 39. "Survival analysis in clinical trials," ASA LearnSTAT course, Teaneck, NJ, 2009.
 40. "Meta-analysis: Summary statistics vs. original data," Biostatistics 50/60 Anniversary Conference, University of Michigan, 2009.
 41. "Statistical issues in genomewide association studies," Biostatistics 60th Anniversary Conference, UNC, 2009.
 42. "Maximum likelihood estimation in semiparametric regression models with censored data," International Biostatistics Conference, Hefei, China, 2009.
 43. "Estimating genetic effects and gene-environment interactions with missing data," Joint Statistical Meetings, Vancouver, Canada, 2010.
 44. "Predictive accuracy of covariates for event times," The International Biometric Society ENAR Spring

- Meeting, Miami, FL, 2011.
45. "Survival analysis with incomplete genetic data," Workshop on Analysis of Survival and Event History Data, Montreal, Canada, 2011.
 46. "Statistical analysis of recurrent event data," ICSA 2011 Applied Statistics Symposium, NYC, 2011.
 47. "Statistical issues in the analysis of disease association with next-generation sequence data", Joint Statistical Meetings, Miami Beach, FL, 2011.
 48. "Ross Prentice's contributions to statistical science," Prentice Symposium, Seattle, WA, 2011.
 49. "A general framework for detecting disease associations with rare variants in sequencing studies," Annual Meeting of International Genetic Epidemiology Society, Heidelberg, Germany, 2011 (platform presentation).
 50. "Quantitative trait analysis in the NHLBI Exome Sequencing Project," Annual Meeting of the American Society of Human Genetics, Montreal, Canada, 2011 (platform presentation).
 51. "A general framework for detecting disease associations with rare variants in sequencing studies," RTP Statistical Genetics Conference, 2011.
 52. "Quantitative trait analysis under trait-dependent sampling," The International Biometric Society ENAR Spring Meeting, Washington, DC, 2012.
 53. "Variable selection in meta-analysis of high-dimensional genetic data," Joint Statistical Meetings, San Diego, CA, 2012.
 54. "Designs and analysis of sequencing studies with trait-dependent sampling," The International Biometric Society ENAR Spring Meeting, Orlando, FL, 2013.
 55. "Survival analysis in clinical trials" (short course), The International Biometric Society ENAR Spring Meeting, Orlando, FL, 2013.
 56. "Designs and analysis of sequencing studies with trait-dependent sampling," The Second Taihu International Statistics Forum, Wuxi, China, 2013.
 57. "Nonparametric maximum likelihood estimation of semiparametric regression models with censored data," IMS-SWUFE International Conference on Statistics and Probability, Chengdu, China, 2013.
 58. "Proportional means/rates models for the composite endpoint of recurrent and terminal events," Joint Statistical Meetings, Montreal, Canada, 2013.
 59. "Statistical methods in genetic association studies" (short course), Joint Statistical Meetings, Montreal, Canada, 2013.
 60. "Survival analysis in clinical trials" (short course), Food and Drug Administration, 2013.
 61. "Analysis of recurrent event data in the presence of a terminal event," Paths to Precision Medicine, Harvard University, Boston, 2014.
 62. "Analysis of recurrent event data," FDA/Advanced Statistics Workshop, Washington, DC, 2014.
 63. "The magic of score statistics," The International Biometric Society ENAR Spring Meeting, Baltimore, MD, 2014.
 64. "Efficient estimation of two-phase cohort studies," Joint Statistical Meetings, Boston, 2014.
 65. "Meta-analysis: Is there an efficiency gain by using original data?" International Biostatistics Conference, Changchun, China, 2014.
 66. "Some recent theoretical developments in meta-analysis," The 10th International Conference on Frontiers of Statistics, Beijing, China, 2015.
 67. "Biomarker discovery and two-stage studies," The Fourth International Symposium on Biopharmaceutical Statistics, Beijing, China, 2015.
 68. "Simultaneous inference on treatment effects in survival studies with factorial designs," ICSA China Statistics Conference, Shanghai, China, 2015.
 69. "Genetic association analysis under complex survey sampling," Joint Statistical Meetings, Seattle, 2015.
 70. "Recent theoretical developments in meta-analysis," The International Biometric Society ENAR Spring Meeting, Austin, TX, 2016.

71. "Maximum likelihood estimation for semiparametric regression models with interval-censored data," Joint Statistical Meetings, Chicago, IL, 2016.
72. "Maximum likelihood estimation for semiparametric regression models with interval-censored data," Duke-Industry Statistics Symposium, Durham, NC, 2016.
73. "Integrative association analysis of multiple genomics platforms," Joint Statistical Meetings, Baltimore, 2017.
74. "Designs and analysis of sequencing studies," Joint Statistical Meetings, Vancouver, Canada, 2018.
75. "Integrative analysis of incomplete multi-omics data," The International Biometric Society ENAR Spring Meeting, Philadelphia, PA, 2019.
76. "Semiparametric single-index models for optimal treatment regimens with censored outcomes," Duke-Industry Statistics Symposium, Durham, NC, 2019.
77. "Semiparametric regression analysis of interval-censored data," The Second Conference on Lifetime Data Science, Pittsburgh, PA, 2019.
78. "Evaluating the efficacy of therapies and vaccines for COVID-19," American Association for the Advancement of Science Annual Meeting, 2021.
79. "Evaluating the long-term efficacy of COVID-19 vaccines," The International Biometric Society ENAR Spring Meeting, 2021.
80. "Semiparametric regression analysis of interval-censored data," The 77th Annual Deming Conference on Applied Statistics, 2021.
81. "Semiparametric regression analysis of interval-censored data," The International Workshop on Statistical Theory and Related Fields, 2021.
82. "Evaluating the effectiveness of COVID-19 vaccines over time," The 35th New England Statistics Symposium, 2022.
83. "Improving the power to establish clinical similarity in a phase 3 efficacy trial by incorporating prior evidence of analytical and pharmacokinetic similarity," FDA Workshop: Increasing the Efficiency of Biosimilar Development Programs, 2022.
84. "Effects of vaccination and prior infection on SARS-CoV-2 infection and severe outcomes over time," Lifetime Data Science Conference, Raleigh, NC, 2023.
85. "Novel statistical analysis of COVID studies," Raising the Impact of Data Science Research: From Theory to Translation Workshop, Boston, MA, 2023.

SEMINARS (SINCE 2002)

1. "Assessment and selection of regression models," Harvard University, 2002.
2. "Assessment and selection of regression models," Columbia University, 2002.
3. "Assessment and selection of regression models," Memorial Sloan Kettering Cancer Center, 2002.
4. "Selection and assessment of regression models," University of Pittsburgh, 2003.
5. "Selection and assessment of regression models," University of Rochester, 2003.
6. "Likelihood-based inference on haplotype effects in genetic association studies," North Carolina State University, 2004.
7. "Maximum likelihood estimation of haplotype effects and haplotype-environment interactions in genetic association studies," George Washington University, 2005.
8. "A general theory for the maximum likelihood estimation in semiparametric models with censored data," University of Oslo, 2005.
9. "Maximum likelihood estimation of haplotype effects and haplotype-environment interactions in genetic association studies," John Hopkins University, 2005.
10. "Likelihood-based inference on haplotype effects in genetic association studies," Washington University

- in St. Louis, 2005.
11. "Maximum likelihood estimation of haplotype effects and haplotype-environment interactions in genetic association studies," National Institute of Environmental Sciences, 2006.
 12. "Likelihood-based inference on haplotype effects in genetic association studies," Duke University, 2006.
 13. "Likelihood-based inference on haplotype effects in genetic association studies," University of Washington, 2006.
 14. "Analysis of haplotype-disease association," Albert Einstein College of Medicine, 2007.
 15. "A general framework for estimating genetic effects and gene-environment interactions with missing data," University of Michigan, 2008.
 16. "A general framework for estimating genetic effects and gene-environment interactions with missing data," Columbia University, 2009.
 17. "Accounting for unknown population substructure in the statistical analysis of genome-wide association studies," Fred Hutchinson Cancer Research Center, 2010.
 18. "Accounting for unknown population substructure in the statistical analysis of genome-wide association studies," Stanford University, 2010.
 19. "Meta-analysis with summary statistics vs. individual level data, with applications to genome-wide association studies," University of Pennsylvania, 2010.
 20. "Meta-analysis with summary statistics vs. individual level data, with applications to genome-wide association studies," Harvard University, 2011.
 21. "Statistical methods for detecting disease associations with rare variants in sequencing studies," Case Western Reserve University, 2011.
 22. "Statistical methods for detecting disease associations with rare variants in sequencing studies," Yale University, 2011.
 23. "Quantitative trait analysis in sequencing studies under trait-dependent sampling," Emory University, 2012.
 24. "Quantitative trait analysis in sequencing studies under trait-dependent sampling," Duke University, 2013.
 25. "Public health, genetics, and statistics," East China Normal University, 2013.
 26. "On random-effects meta-analysis," Harvard University, 2014.
 27. "Meta-analysis: Is there an efficiency gain by using original data?" Peking University, 2014.
 28. "Nonparametric maximum likelihood estimation of semiparametric regression models with censored data," Tsinghua University, 2014.
 29. "Nonparametric maximum likelihood estimation of semiparametric regression models with censored data," Chinese Academy of Sciences, 2014.
 30. "Nonparametric maximum likelihood estimation of semiparametric regression models with censored data," East China Normal University, 2014.
 31. "Meta-analysis: Is there an efficiency gain by using original data?" University of Pennsylvania, 2015.
 32. "Maximum likelihood estimation for semiparametric regression models with interval-censored data," Fred Hutchinson Cancer Research Center, 2016.
 33. "Biostatistics: Research and training," Peking University, 2018.
 34. "Biostatistics: Research and training," Tsinghua University, 2018.
 35. "Design and analysis of two-phase studies, with applications to genetic association studies," Washington University in St. Louis, 2018.
 36. "Evaluating the long-term efficacy of COVID-19 vaccines," Fred Hutchinson Cancer Research Center, 2021.
 37. "Evaluating vaccine efficacy against SARS-CoV-2 infection," Fred Hutchinson Cancer Research Center, 2021.
 38. "Evaluating the durability of protection for COVID-19 vaccines," European Federation of Statisticians

in the Pharmaceutical Industry, 2021.

39. “Waning vaccine-induced immunity in children: Evidence from North Carolina,” The Center for Disease Control and Prevention, 2021.
40. “Strategies to confirm effectiveness over time,” World Health Organization R&D Blueprint Consultation, 2022.
41. “Durability of protection for Covid-19 vaccines,” The American Statistical Association Biopharmaceutical Section, 2022.
42. “Durability of protection for Covid-19 vaccines,” University of Tennessee, 2022.
43. “Semiparametric regression analysis of interval-censored data,” Hong Kong Polytech University, 2022.
44. “Evaluating the effects of vaccination and prior infection on SARS-CoV-2 infection and severe outcomes over time,” National Institutes of Health, 2023.
45. “Research and training in biostatistics,” Peking University, 2023.
46. “Research and training in biostatistics,” Tsinghua University, 2023.
47. “Research and training in biostatistics,” University of Science and Technology of China, 2023.

PH.D. DISSERTATION ADVISEES

Past

Name	Year Completed	Institution of Employment
Charles Spiekerman	1995	University of Washington
Michal Kulich	1997	Charles University
Debashis Ghosh	2000	University of Colorado
Patrick Arbogast	2000	Vanderbilt University
Chengcheng Hu	2001	University of Arizona
Zhiying Pan	2005	Amgen
Guoqing Diao	2006	George Mason University
Emma Huang	2007	Janssen R&D
Li Chen	2009	University of Kentucky
Yijuan Hu	2011	Emory University
Q. Chad He	2012	Fred Hutchinson Cancer Research Center
Zhengzheng Tang	2014	University of Wisconsin
Lu Mao	2016	University of Wisconsin
Ran Tao	2016	Vanderbilt University
Alex Wong	2017	The Hong Kong Polytechnic University
Fei Gao	2017	Fred Hutchinson Cancer Research Center
Paul Little	2019	Fred Hutchinson Cancer Research Center
Jin Wang	2019	Genentech
Sean McCabe	2020	Harvard University
Chong Jin	2020	University of Pennsylvania
John Kidd	2021	Utah Valley University
Rujin Wang	2021	Regeneron Pharmaceuticals
Hannan Yang	2022	Novartis
Yu Gu	2023	Hong Kong University
Richard Sizelove	2023	Eli Lilly

Current

Jianqiao Wang, Yangjianchen Xu, Qiyao Qin, Xinyu Zhang, Yilun Li

PH.D. DISSERTATION COMMITTEES***Past***

Marcia Ciol, Michael Kosorok, Li Hsu, Yu Shen, Ben Wang, Yihui Zhang, William Bradley McNeney, Monika Peckova, Ying Zhang, Antonio Eduardo Gomes, James Lovato, Daniela Gobinelli, Bin Nan, Zoe Moodie, Shuguang Song, Douglas E. Schaebel, Lan Kong, Guosheng Yin, Xiaoli Liu, Hana Lee, Andrew Sterrett, Sangwook Kang, Sarah Nyante, Se-Hee Kim, Alison Wise

Current

Vasyl Zhabotynsky, Heejoon Jo, Bridget M Lin, Alvin George Thomas

CLASSROOM TEACHING***University of Washington***

BIOST 514	Biostatistics I, 1991, 1992, 1993
BIOST 576	Survival Analysis Methods, 1994, 1995, 1996, 1998, 1999, 2000
BIOST 578	Special Topics in Advanced Biostatistics: Monitoring Clinical Trials, 1991
BIOST 578	Special Topics in Advanced Biostatistics: Survival Analysis, 1998

University of North Carolina

BIOS 280	Theory and Methods of Survival Analysis, 2002, 2004, 2005
BIOS 780	Theory and Methods of Survival Analysis, 2007, 2009, 2010, 2011, 2013, 2015, 2017, 2019, 2021, 2023
BIOS 740	Statistical Methods in Genetic Association Studies, 2012
BIOS 782	Statistical Methods in Genetic Association Studies, 2014, 2016, 2018, 2020, 2022

SERVICE

- Biostatistics Search Committee (Chair), 2001–2002
- Statistical Genetics Search Committee, 2001–2003
- Graduate Studies Committee, 2001–present
- Seminar Committee, 2001–present
- Biostatistics Information Technology Committee, 2005–present (Chair, 2021–present)
- Promotion and Tenure Committee, School of Public Health, 2002–2006
- Faculty Council, University of North Carolina, 2003–2008
- Research Council, School of Public Health, 2006–2016
- Search Committee (Cancer Center Biostatistics), 2006–2007
- Search Committee (Statistical Genetics), 2006–2007
- Biostatistics Self-Study Committee, 2007–2008

- Biostatistics Retreat Committee, 2007–2008
- Junior Faculty Mentoring Committees: Fei Zou (Chair), Hongtu Zhu (Chair), Amy Herring, Anastasia Ivanova, Michael Wu, Ethan Lange (Genetics), Yun Li (Genetics), Sun Wei (Chair), Di Wu (Periodontology), Yuchao Jiang (Chair), Michael Love (Chair), Tanya Garcia (Chair), Didong Li, Xihao Li (Chair)
- Search Committee (Clinical Trials), 2008–2009
- Search Committee (Statistical Genetics), 2008–2009
- Search Committee (Chair, Infectious Diseases), 2010–2011
- Search Committee (Chair, Statistical Genetics), 2013–2014
- Search Committee (Chair), 2014–2015
- Search Committee (Statistical Genetics), 2016–2017
- Search Committee for the Chair of the Genetics Department, 2016–2017
- Search Committee for the Assistant Chair of Biostatistics, 2017
- Gillings Study Section Specific Aims Review, 2020–2021
- Biostatistics Search Committee (Chair), 2021–2022

RESEARCH GRANTS

- **R01 HL149683 (4/21/20–3/31/25)**
(formerly R01 GM047845, 08/01/92–04/20/20; MERIT Award: 08/01/05–07/31/15)
National Heart, Lung and Blood Institute
Semiparametric Analysis of Big Censored Data
Role: Principal Investigator
The broad, long-term objectives of this project are to develop semiparametric regression methods for the analysis of censored data, which are commonly encountered in biomedical research on chronic diseases. This renewal application is focused on the computational challenges in the analysis of big data involving tens of millions of individuals with tens of millions of variables.
- **R01 HG009974 (03/01/18–01/31/25)**
(formerly R01 CA082659, 04/01/00–01/31/18)
National Human Genome Research Institute
Statistical Methods in Trans-Omics Chronic Disease Research
Role: Principal Investigator
The proposed research intends to develop novel and high-impact statistical methods for integrative analysis of trans-omics data from ongoing precision medicine studies of chronic diseases. The goal is to facilitate the creation of a new era of medicine in which each patient receives individualized care that matches their genetic code.
- **R01 AI029168 (08/01/89–04/30/29)**
(MERIT Award: 05/01/99–04/30/09; MERIT Award: 05/01/24–04/30/29)
National Institute of Allergy and Infectious Diseases
Statistical Issues in AIDS Research
Role: Principal Investigator for Subcontract
This project addresses important issues in the statistical analysis of data from studies for treatment or prevention of HIV infection, continuing current work on methods for clinical trials of AIDS therapeutics and extending the work scope to the exploration of design and analysis issues in field trials of vaccines for prevention of HIV infection.
- **R01 HL143885 (4/1/19–3/31/24)**

National Heart, Lung and Blood Institute

Leveraging Multi-Omics Approaches to Examine Metabolic Challenges of Obesity in Relation to Cardiovascular Diseases

Role: Co-Investigator

We will develop and employ cutting-edge metabolomics and statistical methods to characterize known and unknown metabolite signals.

- **R01 HG010297 (6/15/19–3/31/24)**

Rutgers University

PAGE III: Population Architecture using Genomics and Epidemiology

Role: Biostatistician

In the next phase of the PAGE study, we propose to develop new resources to further bridge genomic and translational research and close the knowledge gap for genomic translation into ancestrally diverse populations. The outcome of this project will bring large-scale trans-ethnic applications to high level conclusions on genomic architecture.

- **R01 HL147853 (1/1/20–12/31/24)**

National Heart, Lung and Blood Institute

Inflammatory Mediators of Cardiometabolic Risk in Latinos

Role: Co-Investigator

Cardiometabolic risk factors and type 2 diabetes (T2D) impart a substantial and growing morbidity and mortality burden that disproportionately affects racial/ethnic minorities, including Hispanic/Latinos (H/L). This work will shed insight on the mechanisms underlying cardiometabolic disease in H/L, identify potential sources of health disparities in a genetically admixed cohort, and provide an essential foundation for future studies of inflammatory-modulating therapies aimed at reducing the burden of cardiometabolic disease in the population at large.

- **R01 DK122503 (9/22/20–7/31/25)**

National Institute of Diabetes and Digestive and Kidney Diseases

Integrative Approaches to Identifying Function and Clinical Significance of Adiposity Susceptibility Genes

Role: Co-Investigator

We will identify genes underlying GWAS signals, perform clinical characterization, and conduct in vitro functional studies to characterize the molecular underpinnings and biological mechanisms of obesity-risk loci.

- **P01 CA142538 (4/1/15–3/31/21)**

National Cancer Institute

Statistical Methods for Cancer Clinical Trials–Project 3: Statistical/Computational Methods for Pharmacogenomics and Individualized Therapy

Role: Project Leader

This research intends to develop novel and high-impact statistical and computational tools for discovering genetic variants associated with interindividual differences in the efficacy and toxicity of cancer medications and for optimizing drug therapy on the basis of each patient's genetic constitution.

- **P01 CA142538 (4/1/15–3/31/21)**

National Cancer Institute

Statistical Methods for Cancer Clinical Trials–Core A: Administrative Core

Role: Core contributor

The Administrative Core will provide essential administrative and scientific leadership, coordinating

the investigators across three institutions into an effective team and facilitating integration of project research, effective dissemination of research results, and management of resources and logistics.

- **P01 CA142538 (4/1/15–3/31/21)**

National Cancer Institute

Statistical Methods for Cancer Clinical Trials–Core B: Computational Resources and Dissemination Core

Role: Core Director

To facilitate the acceptance and usage of new methods by the cancer research community, it is essential that the methods be evaluated rigorously and implemented in professional, robust, user-friendly software. Documentation and demonstrations of new software must be readily available to researchers and accessible to users with different backgrounds and experience levels. The Computational Resources and Dissemination Core (Core B) will be responsible for the critical functions required to achieve these objectives.