Tianxi Cai

2021-11-08

Business Address

Department of Biostatistics, Harvard T. H. Chan School of Public Health
655 Huntington Ave, Boston MA 02115
Phone: (617)432-4923; Fax: (617)432-5619
E-mail: tcai@hsph.harvard.edu

Education

- **Sc.D. Biostatistics, Harvard University, (1996–1999)** Boston, MA
- **B.Sc. Mathematics, University of Science and Technology of China, (1991–1995)** P. R. China

Professional Experience

- 2018–Present **John Rock Professor of Population and Translational Data Science**, Harvard University
- 2016–Present **Professor**, Dept of Biomedical Informatics, Harvard Medical School
- 2012–Present **Professor**, Dept of Biostatistics, Harvard T. H. Chan School of Public Health
- 2007–2008 **Joy Foundation Fellow**, Radcliffe Institute for Advanced Studies
- 2006–2011 **Associate Professor**, Dept of Biostatistics, Harvard T. H. Chan School of Public Health
- 2002–2006 **Assistant Professor**, Dept of Biostatistics, Harvard T. H. Chan School of Public Health
- 2000–2002 **Assistant Professor**, Dept of Biostatistics, University of Washington
- 1999–2000 **Research Fellow**, Dept of Biostatistics, Harvard T. H. Chan School of Public Health
- 06–08/1997 **Summer Assistant Research Statistician**, Schering Plough Research Institute

Honors and Awards

- Myrto Lefkopoulou Distinguished Lecturer (2014)
- Fellow of American Statistical Association (2011)
- Joy Foundation Fellow, Radcliffe Institute for Advanced Studies (2007–2008)
- Distinguished Visiting Junior Scholar, Peter Wall Institute of Advanced Study (2001)
- Student Travel Award, ENAR International Biometrics Society (1999)
- Authur T. Lyman and Henry S. Grew Scholarship, Harvard University (1998–1999)
- Robert Reed Award for Academic Excellence in Biostatistics, Harvard University (1997)
- Pfizer Scholarship (1997–1998)
- Genetics Institute Scholarship (1996–1997)

Professional Activities

National and International Committee

- COPSS Presidents' Award Committee (2020–2023)
- George W. Snedecor Award Selection Committee (2012–2016)
- Recombinant DNA Advisory Committee (RAC), NIH (2011–2015)
- Program Committee Chair, International Chinese Statistical Association (2012–2013)
- Education Advisory Committee, International Biometric Society ENAR (2010)
- Program Committee Member and Biometrics Section Chair, International Chinese Statistical Association (2011-2012)
- Executive Committee co-Chair, 2012 ICSA Applied Statistical Symposium

- ASA Biometrics Section Program Chair 2011 Joint Statistical Meeting
- Organizing Committee, IMS-China International Conference on Stat. & Prob. (2009)
- IMS Program Chair, International Biometrics Society ENAR (2008–2009)

Other Professional Positions

- Scientific Advisor, Hanalytics, Singapore
- Academic Advisory Board member, Center for Data Science, Zhejiang University, China

Editorial Board

Guest Editor

- Lifetime Data Analysis (Special Issue on Risk Prediction and Evaluation 2013, Joint with Drs. M. Gail and R. Pfeiffer)
- Statistics and Its Interface (Special Issue 2010, Joint with Drs. J. Zhu and H. Zhou)

Associate Editor

- Journal of the American Statistical Association Theory and Methods (2011–)
- Life Time Data Analysis (2010–)
- Statistics in Biosciences (2009–)
- Journal of the Royal Statistical Society Series B (Statistical Methodology) (2014–)
- Journal of the American Statistical Association Applications and Case Studies (2010–2012)
- Biometrics (2008–2012)

Organizer and Chair of Invited Conference Sessions and Symposiums

Organizer/Co-organizer

- Symposium on “Paths to Precision Medicine: the role of statistics” (Harvard, 2014)
- Conference on “Sequencing and Complex Traits: beyond 1000 Genomes” (Harvard, 2012)
- NSF FRG Workshop on “Recent Advances on High Dimensional Data Analysis” (Boston, 2011)
- Radcliffe Institute for Advanced Studies Exploratory Seminar on “High Dimensional Data Analysis” (2008)
- Harvard Schering Plough Workshop (2010, 2007)

Invited Session Organizer

- ENAR (2015)
- ICSA international symposium (2013)
- New England Statistical Symposium (2011)
- IMS-China International conference (2011)
- First Joint Biostatistics Symposium in Beijing, China (2010)
- International Conference on Statistics and Society (2010)
- Joint Statistical Meeting (2010)
- Organizer and Chair, IMS Medallion Lecture at ENAR (2009)
- ICSA Applied Statistics Symposium (2008)
- Joint Statistical Meeting (2007)
- ICSA Applied Statistics Symposium (2005)

Review Panel for National and International Grants

- National Institute of Health SBIR- Health Informatics Special Emphasis panel (2017, 2018)
- National Institute of Health Big Data to Knowledge (BD2K) Mentored Career Development Award review panel (2016)
- National Science Foundation Statistics Career Panel (2016)
- Joint DMS/NIGMS National Science Foundation/National Institutes of Health review panel (2013)
- Susan G. Komen Investigator Initiated Research - Prevention/Early Detection Panel (2010)
- Doctoral Program Plus for Austrian National Science Foundation (2009)

- Susan G. Komen Grant Foundation - Prevention/Risk Reduction Panel (2007-2009)
- Clinical and Integrative Cardiovascular Sciences Study Section, NIH (2007-2009)
- Small Grants for Cancer Epidemiology (2008)
- Research Grants Council of HongKong (2006)
- Biostatistical Methods and Research Design Study Section, NIH (2004)

Referee

- Annals of Statistics, Biometrika, Biometrics, Biostatistics, Journal of the American Statistical Association, Journal of Computational Statistics and Data Analysis, Journal of the Royal Statistical Society, Medical Decision Making, Statistica Sinica, Statistics in Medicine, American Journal of Epidemiology, Circulation, Functional and Integrative Genomics

Publications

Methodology Research

1. **Cai T**, Cai TT, Guo Z. Optimal statistical inference for individualized treatment effects in high-dimensional models. J R Stat Soc Series B. 2021;83:669–719. https://doi.org/10.1111/rssb.12426
2. **Cai T**, Liu M, Xia Y. Individual Data Protected Integrative Regression Analysis of High-Dimensional Heterogeneous Data. Journal of the American Statistical Association. 2021; https://doi.org/10.1080/01621459.2021.1904958
3. Chan S, Wang X, Jazic I, Peskoe S, Zheng Y, **Cai T**. Developing and evaluating risk prediction models with panel current status data. Biometrics. 2021 06; 77(2):599-609. https://pubmed.ncbi.nlm.nih.gov/32562264
4. Cheng D, Ananthakrishnan AN, **Cai T**. Robust and efficient semi-supervised estimation of average treatment effects with application to electronic health records data. Biometrics. 2021 06; 77(2):413-423. https://pubmed.ncbi.nlm.nih.gov/32413171
5. Liu M, Xia Y, Cho K, **Cai T**. Integrative High Dimensional Multiple Testing with Heterogeneity under Data Sharing Constraints. Journal of Machine Learning Research. 2021; 22(126):1-26. https://jmlr.org/papers/v22/20-774.html
6. Wang X, **Cai T**, Tian L, Bourgeois F, Parast L. Quantifying the feasibility of shortening clinical trial duration using surrogate markers. Stat Med. 2021 Sep 02. https://pubmed.ncbi.nlm.nih.gov/34474500
7. Wang X, Zheng Y, Jensen MK, He Z, **Cai T**. Biomarker evaluation under imperfect nested case-control design. Stat Med. 2021 08 15; 40(18):4035-4052. https://pubmed.ncbi.nlm.nih.gov/33915597
8. Wang Y, Hong C, Palmer N, Di Q, Schwartz J, Kohane I, **Cai T**. A fast divide-and-conquer sparse Cox regression. Biostatistics. 2021 04 10; 22(2):381-401. https://pubmed.ncbi.nlm.nih.gov/31545341
9. Zhang HG, Hejblum BP, Weber GM, Palmer NP, Churchill SE, Szolovits P, Murphy SN, Liao KP, Kohane IS, **Cai T**. ATLAS: an automated association test using probabilistically linked health records with application to genetic studies. J Am Med Inform Assoc. 2021 Oct 05. https://pubmed.ncbi.nlm.nih.gov/34608931
10. Zheng Y, Lu T, **Cai T**. Efficient and robust estimation of t-year risk prediction models leveraging time varying intermediate outcomes. Statistica Sinica. 2021; 31:1531-1547. https://doi.org/10.5705/ss.202019.0066
11. Ahuja Y, Zhou D, He Z, Sun J, Castro VM, Gainer V, Murphy SN, Hong C, **Cai T**. sureLDA: A multidisease automated phenotyping method for the electronic health record. J Am Med Inform Assoc. 2020 08 01; 27(8):1235-1243. https://pubmed.ncbi.nlm.nih.gov/32548637
12. Chan SF, Hejblum BP, Chakrabortty A, **Cai T**. Semi-supervised estimation of covariance with application to phenome-wide association studies with electronic medical records data. Stat Methods Med Res. 2020 02; 29(2):455-465. https://pubmed.ncbi.nlm.nih.gov/30943854
13. Cheng D, Chakrabortty A, Ananthakrishnan AN, **Cai T**. Estimating average treatment effects with a double-index propensity score. Biometrics. 2020 09; 76(3):767-777. https://pubmed.ncbi.nlm.nih.gov/31797368
14. Hong C, Wang Y, **Cai T**. A divide-and-conquer method for sparse risk prediction and evaluation. Biostatistics. 2020 Sep 10. https://pubmed.ncbi.nlm.nih.gov/32909599
15. Parast L, **Cai T**, Tian L. Evaluating multiple surrogate markers with censored data. Biometrics. 2020 Sep 13. https://pubmed.ncbi.nlm.nih.gov/32920821
16. Parast L, Tian L, **Cai T**. Assessing the value of a censored surrogate outcome. Lifetime Data Anal. 2020 04; 26(2):245-265. https://pubmed.ncbi.nlm.nih.gov/30980316
17. Shi X, Li X, **Cai T**. Spherical Regression Under Mismatch Corruption With Application to Automated Knowledge Translation. Journal of the American Statistical Association. 2020; https://doi.org/10.1080/

01621459.2020.1752219

18. Wang X, Parast L, Tian LU, **Cai T**. Model-free approach to quantifying the proportion of treatment effect explained by a surrogate marker. Biometrika. 2020 Mar; 107(1):107-122. https://pubmed.ncbi.nlm.nih.gov/32587413
19. **Cai T**, Cai TT, Liao K, Liu W. Large-Scale Simultaneous Testing of Cross-Covariance Matrices with Applications to PheWAS. Stat Sin. 2019 Apr; 29(2):983-1005. https://pubmed.ncbi.nlm.nih.gov/31889766
20. Gamerman V, **Cai T**, Elsäßer A. Pragmatic randomized clinical trials: best practices and statistical guidance. Health Serv Outcomes Res Method. 2019; 19:23–35. https://doi.org/10.1007/s10742-018-0192-5 https://doi.org/10.1007/s10742-018-0192-5
21. Goodman MO, Chibnik L, **Cai T**. Variance components genetic association test for zero-inflated count outcomes. Genet Epidemiol. 2019 02; 43(1):82-101. https://pubmed.ncbi.nlm.nih.gov/30353568
22. Gronsbell J, Minnier J, Yu S, Liao K, **Cai T**. Automated feature selection of predictors in electronic medical records data. Biometrics. 2019 03; 75(1):268-277. https://pubmed.ncbi.nlm.nih.gov/30353541
23. Hejblum BP, Weber GM, Liao KP, Palmer NP, Churchill S, Shadick NA, Szolovits P, Murphy SN, Kohane IS, **Cai T**. Probabilistic record linkage of de-identified research datasets with discrepancies using diagnosis codes. Sci Data. 2019 01 08; 6:180298. https://pubmed.ncbi.nlm.nih.gov/30620344
24. Hong C, Liao KP, **Cai T**. Semi-supervised validation of multiple surrogate outcomes with application to electronic medical records phenotyping. Biometrics. 2019 03; 75(1):78-89. https://pubmed.ncbi.nlm.nih.gov/30267536
25. Liao KP, Sun J, Cai TA, Link N, Hong C, Huang J, Huffman JE, Gronsbell J, Zhang Y, Ho YL, Castro V, Gainer V, Murphy SN, O'Donnell CJ, Gaziano JM, Cho K, Szolovits P, Kohane IS, Yu S, **Cai T**. High-throughput multimodal automated phenotyping (MAP) with application to PheWAS. J Am Med Inform Assoc. 2019 11 01; 26(11):1255-1262. https://pubmed.ncbi.nlm.nih.gov/31613361
26. Maziarz M, **Cai T**, Qi L, Lok AS, Zheng Y. Evaluating longitudinal markers under two-phase study designs. Biostatistics. 2019 07 01; 20(3):485-498. https://pubmed.ncbi.nlm.nih.gov/29912281
27. Ning W, Chan S, Beam A, Yu M, Geva A, Liao K, Mullen M, Mandl KD, Kohane I, **Cai T**, Yu S. Feature extraction for phenotyping from semantic and knowledge resources. J Biomed Inform. 2019 03; 91:103122. https://pubmed.ncbi.nlm.nih.gov/30738949
28. Parast L, **Cai T**, Tian L. Using a surrogate marker for early testing of a treatment effect. Biometrics. 2019 12; 75(4):1253-1263. https://pubmed.ncbi.nlm.nih.gov/31009073
29. Zhang L, Zhang Y, Cai T, Ahuja Y, He Z, Ho YL, Beam A, Cho K, Carroll R, Denny J, Kohane I, Liao K, **Cai T**. Automated grouping of medical codes via multiview banded spectral clustering. J Biomed Inform. 2019 12; 100:103322. https://pubmed.ncbi.nlm.nih.gov/31672532
30. Zhang Y, Cai T, Yu S, Cho K, Hong C, Sun J, Huang J, Ho YL, Ananthakrishnan AN, Xia Z, Shaw SY, Gainer V, Castro V, Link N, Honerlaw J, Huang S, Gagnon D, Karlson EW, Plenge RM, Szolovits P, Savova G, Churchill S, O'Donnell C, Murphy SN, Gaziano JM, Kohane I, **Cai T**, Liao KP. High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). Nat Protoc. 2019 12; 14(12):3426-3444. https://pubmed.ncbi.nlm.nih.gov/31748751
31. Agniel D, Xie W, Essex M, **Cai T**. Functional principal variance component testing for a genetic association study of HIV progression. Ann Appl Stat. 2018 Sep; 12(3):1871-1893. https://pubmed.ncbi.nlm.nih.gov/32256927
32. Liu D, **Cai T**, Lok A, Zheng Y. Nonparametric Maximum Likelihood Estimators of Time-Dependent Accuracy Measures for Survival Outcome Under Two-Stage Sampling Designs. J Am Stat Assoc. 2018; 113(522):882-892. https://pubmed.ncbi.nlm.nih.gov/30555194
33. Sinnott JA, **Cai T**. Pathway aggregation for survival prediction via multiple kernel learning. Stat Med. 2018 07 20; 37(16):2501-2515. https://pubmed.ncbi.nlm.nih.gov/29664143
34. Xia Y, **Cai T**, Cai TT. Multiple Testing of Submatrices of a Precision Matrix with Applications to Identification of Between Pathway Interactions. J Am Stat Assoc. 2018; 113(521):328-339. https://pubmed.ncbi.nlm.nih.gov/29881130
35. Xia Y, **Cai T**, Cai TT. Two-Sample Tests for High-Dimensional Linear Regression with an Application to Detecting Interactions. Stat Sin. 2018 Jan; 28:63-92. https://pubmed.ncbi.nlm.nih.gov/29386856
36. Yu S, Ma Y, Gronsbell J, Cai T, Ananthakrishnan AN, Gainer VS, Churchill SE, Szolovits P, Murphy SN, Kohane IS, Liao KP, **Cai T**. Enabling phenotypic big data with PheNorm. J Am Med Inform Assoc. 2018 01 01; 25(1):54-60. https://pubmed.ncbi.nlm.nih.gov/29126253
37. Agniel D, **Cai T**. Analysis of multiple diverse phenotypes via semiparametric canonical correlation analysis. Biometrics. 2017 12; 73(4):1254-1265. https://pubmed.ncbi.nlm.nih.gov/28407213

38. Chen S, Tian L, **Cai T**, Yu M. A general statistical framework for subgroup identification and comparative treatment scoring. Biometrics. 2017 12; 73(4):1199-1209. https://pubmed.ncbi.nlm.nih.gov/28211943
39. Dai W, Yang M, Wang C, **Cai T**. Sequence robust association test for familial data. Biometrics. 2017 09; 73(3):876-884. https://pubmed.ncbi.nlm.nih.gov/28273695
40. Koch M, Furtado JD, Jiang GZ, Gray BE, **Cai T**, Sacks F, Tjønneland A, Overvad K, Jensen MK. Associations of anthropometry and lifestyle factors with HDL subspecies according to apolipoprotein C-III. J Lipid Res. 2017 06; 58(6):1196-1203. https://pubmed.ncbi.nlm.nih.gov/28365588
41. Maziarz M, Heagerty P, **Cai T**, Zheng Y. On longitudinal prediction with time-to-event outcome: Comparison of modeling options. Biometrics. 2017 03; 73(1):83-93. https://pubmed.ncbi.nlm.nih.gov/27438160
42. Parast L, **Cai T**, Tian L. Evaluating surrogate marker information using censored data. Stat Med. 2017 05 20; 36(11):1767-1782. https://pubmed.ncbi.nlm.nih.gov/28088843
43. Wang X, Zhang Z, Morris N, **Cai T**, Lee S, Wang C, Yu TW, Walsh CA, Lin X. Rare variant association test in family-based sequencing studies. Brief Bioinform. 2017 Nov 01; 18(6):954-961. https://pubmed.ncbi.nlm.nih.gov/27677958
44. Yu S, Chakrabortty A, Liao KP, Cai T, Ananthakrishnan AN, Gainer VS, Churchill SE, Szolovits P, Murphy SN, Kohane IS, **Cai T**. Surrogate-assisted feature extraction for high-throughput phenotyping. J Am Med Inform Assoc. 2017 Apr 01; 24(e1):e143-e149. https://pubmed.ncbi.nlm.nih.gov/27632993
45. Zheng Y, **Cai T**. Augmented estimation for t-year survival with censored regression models. Biometrics. 2017 12; 73(4):1169-1178. https://pubmed.ncbi.nlm.nih.gov/28294286
46. Zheng Y, Brown M, Lok A, **Cai T**. Improving efficiency in biomarker incremental value evaluation under two-phase designs. Ann Appl Stat. 2017 Jun; 11(2):638-654. https://pubmed.ncbi.nlm.nih.gov/28943991
47. Zhou QM, Dai W, Zheng Y, **Cai T**. Robust Dynamic Risk Prediction with Longitudinal Studies. Stat Theory Relat Fields. 2017; 1(2):159-170. https://pubmed.ncbi.nlm.nih.gov/29335682
48. **Cai T**, Cai TT, Zhang A. Structured Matrix Completion with Applications to Genomic Data Integration. J Am Stat Assoc. 2016; 111(514):621-633. https://pubmed.ncbi.nlm.nih.gov/28042188
49. Agniel D, Liao KP, **Cai T**. Estimation and testing for multiple regulation of multivariate mixed outcomes. Biometrics. 2016 12; 72(4):1194-1205. https://pubmed.ncbi.nlm.nih.gov/26910481
50. He Q, **Cai T**, Liu Y, Zhao N, Harmon QE, Almli LM, Binder EB, Engel SM, Ressler KJ, Conneely KN, Lin X, Wu MC. Prioritizing individual genetic variants after kernel machine testing using variable selection. Genet Epidemiol. 2016 Dec; 40(8):722-731. https://pubmed.ncbi.nlm.nih.gov/27488097
51. Huang YT, **Cai T**, Kim E. Integrative genomic testing of cancer survival using semiparametric linear transformation models. Stat Med. 2016 07 20; 35(16):2831-44. https://pubmed.ncbi.nlm.nih.gov/26887583
52. Huang YT, **Cai T**. Mediation analysis for survival data using semiparametric probit models. Biometrics. 2016 06; 72(2):563-74. https://pubmed.ncbi.nlm.nih.gov/26618735
53. Li J, Zhao L, Tian L, **Cai T**, Claggett B, Callegaro A, Dizier B, Spiessens B, Ulloa-Montoya F, Wei LJ. A predictive enrichment procedure to identify potential responders to a new therapy for randomized, comparative controlled clinical studies. Biometrics. 2016 09; 72(3):877-87. https://pubmed.ncbi.nlm.nih.gov/26689167
54. Neykov M, Liu JS, **Cai T**. L1-Regularized Least Squares for Support Recovery of High Dimensional Single Index Models with Gaussian Designs. J Mach Learn Res. 2016 May; 17(1):2976-3012. https://pubmed.ncbi.nlm.nih.gov/28503101
55. Payne R, Neykov M, Jensen MK, **Cai T**. Kernel machine testing for risk prediction with stratified case cohort studies. Biometrics. 2016 06; 72(2):372-81. https://pubmed.ncbi.nlm.nih.gov/26692376
56. Payne R, Yang M, Zheng Y, Jensen MK, **Cai T**. Robust risk prediction with biomarkers under two-phase stratified cohort design. Biometrics. 2016 12; 72(4):1037-1045. https://pubmed.ncbi.nlm.nih.gov/27037494
57. Shen Y, **Cai T**. Identifying predictive markers for personalized treatment selection. Biometrics. 2016 12; 72(4):1017-1025. https://pubmed.ncbi.nlm.nih.gov/26999054
58. Sinnott JA, **Cai T**. Inference for survival prediction under the regularized Cox model. Biostatistics. 2016 10; 17(4):692-707. https://pubmed.ncbi.nlm.nih.gov/27107008
59. Yong FH, Tian L, Yu S, **Cai T**, Wei LJ. Optimal stratification in outcome prediction using baseline information. Biometrika. 2016 12; 103(4):817-828. https://pubmed.ncbi.nlm.nih.gov/29422691
60. Minnier J, Yuan M, Liu JS, **Cai T**. Risk Classification with an Adaptive Naive Bayes Kernel Machine Model. J Am Stat Assoc. 2015 Apr 22; 110(509):393-404. https://pubmed.ncbi.nlm.nih.gov/26236061
61. Shen Y, **Cai T**, Chen Y, Yang Y, Chen J. Retrospective likelihood-based methods for analyzing case-cohort genetic association studies. Biometrics. 2015 Dec; 71(4):960-8. https://pubmed.ncbi.nlm.nih.gov/26177343
62. Shen Y, Liao KP, **Cai T**. Sparse kernel machine regression for ordinal outcomes. Biometrics. 2015 Mar; 71(1):63-70. https://pubmed.ncbi.nlm.nih.gov/25196727

63. Xia Y, **Cai T**, Cai TT. Testing Differential Networks with Applications to Detecting Gene-by-Gene Interactions. Biometrika. 2015 Jun; 102(2):247-266. https://pubmed.ncbi.nlm.nih.gov/28502988
64. Yu S, Liao KP, Shaw SY, Gainer VS, Churchill SE, Szolovits P, Murphy SN, Kohane IS, **Cai T**. Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. J Am Med Inform Assoc. 2015 Sep; 22(5):993-1000. https://pubmed.ncbi.nlm.nih.gov/25929596
65. Zhou QM, Zheng Y, Chibnik LB, Karlson EW, **Cai T**. Assessing incremental value of biomarkers with multiphase nested case-control studies. Biometrics. 2015 Dec; 71(4):1139-49. https://pubmed.ncbi.nlm.nih.gov/26195245
66. Matsouaka RA, Li J, **Cai T**. Evaluating marker-guided treatment selection strategies. Biometrics. 2014 Sep; 70(3):489-499. https://pubmed.ncbi.nlm.nih.gov/24779731
67. Parast L, Tian L, **Cai T**. Landmark Estimation of Survival and Treatment Effect in a Randomized Clinical Trial. J Am Stat Assoc. 2014 Jan 01; 109(505):384-394. https://pubmed.ncbi.nlm.nih.gov/24659838
68. Sinnott JA, Dai W, Liao KP, Shaw SY, Ananthakrishnan AN, Gainer VS, Karlson EW, Churchill S, Szolovits P, Murphy S, Kohane I, Plenge R, **Cai T**. Improving the power of genetic association tests with imperfect phenotype derived from electronic medical records. Hum Genet. 2014 Nov; 133(11):1369-82. https://pubmed.ncbi.nlm.nih.gov/25062868
69. **Cai T**, Betensky R. Reply to Sabanés Bové and Held's "comment on Cai and Betensky (2003), on the poisson approximation for hazard regression". Biometrics. 2013 Sep; 69(3):796. https://pubmed.ncbi.nlm.nih.gov/24073867
70. **Cai T**, Zheng Y. Resampling Procedures for Making Inference under Nested Case-control Studies. J Am Stat Assoc. 2013 Jan 01; 108(504):1532-1544. https://pubmed.ncbi.nlm.nih.gov/24436503
71. Parast L, **Cai T**. Landmark risk prediction of residual life for breast cancer survival. Stat Med. 2013 Sep 10; 32(20):3459-71. https://pubmed.ncbi.nlm.nih.gov/23494768
72. Sinnott JA, **Cai T**. Omnibus risk assessment via accelerated failure time kernel machine modeling. Biometrics. 2013 Dec; 69(4):861-73. https://pubmed.ncbi.nlm.nih.gov/24328713
73. Uno H, Tian L, **Cai T**, Kohane IS, Wei LJ. A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data. Stat Med. 2013 Jun 30; 32(14):2430-42. https://pubmed.ncbi.nlm.nih.gov/23037800
74. Zhao L, Tian L, **Cai T**, Claggett B, Wei LJ. Effectively selecting a target population for a future comparative study. J Am Stat Assoc. 2013 Jan 01; 108(502):527-539. https://pubmed.ncbi.nlm.nih.gov/24058223
75. Zheng Y, **Cai T**, Pepe MS. Adopting nested case-control quota sampling designs for the evaluation of risk markers. Lifetime Data Anal. 2013 Oct; 19(4):568-88. https://pubmed.ncbi.nlm.nih.gov/23807695
76. Zheng Y, Parast L, **Cai T**, Brown M. Evaluating incremental values from new predictors with net reclassification improvement in survival analysis. Lifetime Data Anal. 2013 Jul; 19(3):350-70. https://pubmed.ncbi.nlm.nih.gov/23254468
77. Zhou QM, Zheng Y, **Cai T**. Assessment of biomarkers for risk prediction with nested case-control studies. Clin Trials. 2013 Oct; 10(5):677-9. https://pubmed.ncbi.nlm.nih.gov/24013405
78. Zhou QM, Zheng Y, **Cai T**. Subgroup specific incremental value of new markers for risk prediction. Lifetime Data Anal. 2013 Apr; 19(2):142-69. https://pubmed.ncbi.nlm.nih.gov/23263882
79. **Cai T**, Lin X, Carroll RJ. Identifying genetic marker sets associated with phenotypes via an efficient adaptive score test. Biostatistics. 2012 Sep; 13(4):776-90. https://pubmed.ncbi.nlm.nih.gov/22734045
80. **Cai T**, Zheng Y. Evaluating prognostic accuracy of biomarkers in nested case-control studies. Biostatistics. 2012 Jan; 13(1):89-100. https://pubmed.ncbi.nlm.nih.gov/21856652
81. León LF, **Cai T**. Model Checking Techniques for Assessing Functional Form Specifications in Censored Linear Regression Models. Stat Sin. 2012 Apr 01; 22(2):509-530. https://pubmed.ncbi.nlm.nih.gov/23825917
82. Liu D, **Cai T**, Zheng Y. Evaluating the predictive value of biomarkers with stratified case-cohort design. Biometrics. 2012 Dec; 68(4):1219-27. https://pubmed.ncbi.nlm.nih.gov/23173848
83. Parast L, Cheng SC, **Cai T**. Landmark Prediction of Long Term Survival Incorporating Short Term Event Time Information. J Am Stat Assoc. 2012; 107(500):1492-1501. https://pubmed.ncbi.nlm.nih.gov/23293405
84. Tian L, **Cai T**, Zhao L, Wei LJ. On the covariate-adjusted estimation for an overall treatment difference with data from a randomized comparative clinical trial. Biostatistics. 2012 Apr; 13(2):256-73. https://pubmed.ncbi.nlm.nih.gov/22294672
85. Zheng Y, **Cai T**, Jin Y, Feng Z. Evaluating prognostic accuracy of biomarkers under competing risk. Biometrics. 2012 Jun; 68(2):388-96. https://pubmed.ncbi.nlm.nih.gov/22150576
86. **Cai T**, Gerds TA, Zheng Y, Chen J. Robust prediction of t-year survival with data from multiple studies. Biometrics. 2011 Jun; 67(2):436-44. https://pubmed.ncbi.nlm.nih.gov/20670303

87. **Cai T**, Tian L, Lloyd-Jones DM. Comparing costs associated with risk stratification rules for t-year survival. Biostatistics. 2011 Oct; 12(4):597-609. https://pubmed.ncbi.nlm.nih.gov/21415016
88. **Cai T**, Tian L, Wong PH, Wei LJ. Analysis of randomized comparative clinical trial data for personalized treatment selections. Biostatistics. 2011 Apr; 12(2):270-82. https://pubmed.ncbi.nlm.nih.gov/20876663
89. **Cai T**, Tonini G, Lin X. Kernel machine approach to testing the significance of multiple genetic markers for risk prediction. Biometrics. 2011 Sep; 67(3):975-86. https://pubmed.ncbi.nlm.nih.gov/21281275
90. **Cai T**, Zheng Y. Non-parametric Evaluation of Biomarker Accuracy under Nested Case-control Studies. J Am Stat Assoc. 2011 Jan 01; 106(494):569-580. https://pubmed.ncbi.nlm.nih.gov/22844169
91. Lin X, **Cai T**, Wu MC, Zhou Q, Liu G, Christiani DC, Lin X. Kernel machine SNP-set analysis for censored survival outcomes in genome-wide association studies. Genet Epidemiol. 2011 Nov; 35(7):620-31. https://pubmed.ncbi.nlm.nih.gov/21818772
92. Minnier J, Tian L, **Cai T**. A Perturbation Method for Inference on Regularized Regression Estimates. J Am Stat Assoc. 2011 Jan 01; 106(496):1371-1382. https://pubmed.ncbi.nlm.nih.gov/22844171
93. Parast L, Cheng SC, **Cai T**. Incorporating short-term outcome information to predict long-term survival with discrete markers. Biom J. 2011 Mar; 53(2):294-307. https://pubmed.ncbi.nlm.nih.gov/21337601
94. Tian L, Wang R, **Cai T**, Wei LJ. The highest confidence density region and its usage for joint inferences about constrained parameters. Biometrics. 2011 Jun; 67(2):604-10. https://pubmed.ncbi.nlm.nih.gov/20825392
95. Uno H, **Cai T**, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med. 2011 May 10; 30(10):1105-17. https://pubmed.ncbi.nlm.nih.gov/21484848
96. Uno H, **Cai T**, Tian L, Wei LJ. Graphical procedures for evaluating overall and subject-specific incremental values from new predictors with censored event time data. Biometrics. 2011 Dec; 67(4):1389-96. https://pubmed.ncbi.nlm.nih.gov/21504421
97. Wu MC, Lee S, **Cai T**, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011 Jul 15; 89(1):82-93. https://pubmed.ncbi.nlm.nih.gov/21737059
98. **Cai T**, Parast L, Ryan L. Meta-analysis for rare events. Stat Med. 2010 Sep 10; 29(20):2078-89. https://pubmed.ncbi.nlm.nih.gov/20623822
99. **Cai T**, Tian L, Uno H, Solomon SD, Wei LJ. Calibrating parametric subject-specific risk estimation. Biometrika. 2010 Jun; 97(2):389-404. https://pubmed.ncbi.nlm.nih.gov/23049123
100. McDaniel S, Minnier J, Betensky RA, Mohapatra G, Shen Y, Gusella JF, Louis DN, **Cai T**. Assessing Population Level Genetic Instability via Moving Average. Stat Biosci. 2010 Dec; 2(2):120-136. https://pubmed.ncbi.nlm.nih.gov/22866169
101. Wang R, Tian L, **Cai T**, Wei LJ. Nonparametric inference procedure for percentiles of the random effects distribution in meta-analysis. Ann Appl Stat. 2010; 4(1):520-532. https://pubmed.ncbi.nlm.nih.gov/25678939
102. Zheng Y, **Cai T**, Stanford JL, Feng Z. Semiparametric models of time-dependent predictive values of prognostic biomarkers. Biometrics. 2010 Mar; 66(1):50-60. https://pubmed.ncbi.nlm.nih.gov/19397579
103. **Cai T**, Huang J, Tian L. Regularized estimation for the accelerated failure time model. Biometrics. 2009 Jun; 65(2):394-404. https://pubmed.ncbi.nlm.nih.gov/18573133
104. Tian L, **Cai T**, Pfeffer MA, Piankov N, Cremieux PY, Wei LJ. Exact and efficient inference procedure for meta-analysis and its application to the analysis of independent 2 x 2 tables with all available data but without artificial continuity correction. Biostatistics. 2009 Apr; 10(2):275-81. https://pubmed.ncbi.nlm.nih.gov/18922759
105. **Cai T**, Cheng S. Robust combination of multiple diagnostic tests for classifying censored event times. Biostatistics. 2008 Apr; 9(2):216-33. https://pubmed.ncbi.nlm.nih.gov/18056687
106. Gerds TA, **Cai T**, Schumacher M. The performance of risk prediction models. Biom J. 2008 Aug; 50(4):457-79. https://pubmed.ncbi.nlm.nih.gov/18663757
107. Ware JH, **Cai T**. Comments on 'Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond' by M. J. Pencina et al., Statistics in Medicine (DOI: 10.1002/sim.2929). Stat Med. 2008 Jan 30; 27(2):185-7. https://pubmed.ncbi.nlm.nih.gov/17668917
108. Zheng Y, **Cai T**, Pepe MS, Levy WC. Time-dependent Predictive Values of Prognostic Biomarkers with Failure Time Outcome. J Am Stat Assoc. 2008; 103(481):362-368. https://pubmed.ncbi.nlm.nih.gov/19655041
109. **Cai T**, Zheng Y. Model checking for ROC regression analysis. Biometrics. 2007 Mar; 63(1):152-63. https://pubmed.ncbi.nlm.nih.gov/17447940
110. **Cai T**, Gilbert PB, Self SG. Joint inferences on vaccine efficacy against infection and disease with application to the first HIV vaccine efficacy trial. J Biopharm Stat. 2006; 16(4):517-38. https://pubmed.ncbi.nlm.nih.

gov/16892911

111. **Cai T**, Pepe MS, Zheng Y, Lumley T, Jenny NS. The sensitivity and specificity of markers for event times. Biostatistics. 2006 Apr; 7(2):182-97. https://pubmed.ncbi.nlm.nih.gov/16079162
112. Pepe MS, **Cai T**, Longton G. Combining predictors for classification using the area under the receiver operating characteristic curve. Biometrics. 2006 Mar; 62(1):221-9. https://pubmed.ncbi.nlm.nih.gov/16542249
113. Zheng Y, **Cai T**, Feng Z. Application of the time-dependent ROC curves for prognostic accuracy with multiple biomarkers. Biometrics. 2006 Mar; 62(1):279-87. https://pubmed.ncbi.nlm.nih.gov/16542256
114. **Cai T**, Cheng S. Semiparametric regression analysis for doubly censored data. Biometrika. 2004; 91(2): 277-290. https://doi.org/10.1093/biomet/91.2.277
115. **Cai T**, Moskowitz CS. Semi-parametric estimation of the binormal ROC curve for a continuous diagnostic test. Biostatistics. 2004 Oct; 5(4):573-86. https://pubmed.ncbi.nlm.nih.gov/15475420
116. **Cai T**. Semi-parametric ROC regression analysis with placement values. Biostatistics. 2004 Jan; 5(1):45-60. https://pubmed.ncbi.nlm.nih.gov/14744827
117. Pepe MS, **Cai T**. The analysis of placement values for evaluating discriminatory measures. Biometrics. 2004 Jun; 60(2):528-35. https://pubmed.ncbi.nlm.nih.gov/15180681
118. **Cai T**, Betensky RA. Hazard regression for interval-censored data with penalized spline. Biometrics. 2003 Sep; 59(3):570-9. https://pubmed.ncbi.nlm.nih.gov/14601758
119. **Cai T**, Cheng SC, Wei LJ. Semiparametric Mixed-Effects Models for Clustered Failure Time Data. Journal of the American Statistical Association. 2002; 97(458):514-522. https://doi.org/10.1198/016214502760047041
120. **Cai T**, Hyndman RJ, Wand MP. Mixed Model-Based Hazard Estimation. Journal of Computational and Graphical Statistics. 2002; 11(4):784-798. https://doi.org/10.1198/106186002862
121. **Cai T**, Pepe MS. Semiparametric Receiver Operating Characteristic Analysis to Evaluate Biomarkers for Disease. Journal of the American Statistical Association. 2002; 97(460):1099-1107. https://doi.org/10.1198/016214502388618915
122. **Cai T**, Wei LJ, Wilcox M. Semiparametric regression analysis for clustered failure time data. Biometrika. 2000; 87(4):867-878. https://doi.org/10.1093/biomet/87.4.867
123. **Cai T**, Wei LJ. Regression Analysis for Multivariate Failure Time Observations. Asymptotics in Statistics and Probability. 2000; 33-45. https://doi.org/10.1515/9783110942002-006
124. **Cai T**, Shen J. Boundedness Is Redundant in a Theorem of Daubechies. Applied and Computational Harmonic Analysis. 1999; 6(3):400-404. https://doi.org/10.1006/acha.1997.0257

Collaborative Research

125. Ahuja Y, Kim N, Liang L, Cai T, Dahal K, Seyok T, Lin C, Finan S, Liao K, Savovoa G, Chitnis T, **Cai T**, Xia Z. Leveraging electronic health records data to predict multiple sclerosis disease activity. Ann Clin Transl Neurol. 2021 04; 8(4):800-810. https://pubmed.ncbi.nlm.nih.gov/33626237
126. Bourgeois FT, Gutiérrez-Sacristán A, Keller MS, Liu M, Hong C, Bonzel CL, Tan ALM, Aronow BJ, Boeker M, Booth J, Cruz Rojo J, Devkota B, García Barrio N, Gehlenborg N, Geva A, Hanauer DA, Hutch MR, Issitt RW, Klann JG, Luo Y, Mandl KD, Mao C, Moal B, Moshal KL, Murphy SN, Neuraz A, Ngiam KY, Omenn GS, Patel LP, Jiménez MP, Sebire NJ, Balazote PS, Serret-Larmande A, South AM, Spiridou A, Taylor DM, Tippmann P, Visweswaran S, Weber GM, Kohane IS, **Cai T**, Avillach P. International Analysis of Electronic Health Records of Children and Youth Hospitalized With COVID-19 Infection in 6 Countries. JAMA Netw Open. 2021 06 01; 4(6):e2112596. https://pubmed.ncbi.nlm.nih.gov/34115127
127. Cai T, Cai F, Dahal KP, Cremone G, Lam E, Golnik C, Seyok T, Hong C, **Cai T**, Liao KP. Improving the Efficiency of Clinical Trial Recruitment Using an Ensemble Machine Learning to Assist With Eligibility Screening. ACR Open Rheumatol. 2021 Sep; 3(9):593-600. https://pubmed.ncbi.nlm.nih.gov/34296815
128. Cheng D, Cushing KC, **Cai T**, Ananthakrishnan AN. Safety and Efficacy of Tumor Necrosis Factor Antagonists in Older Patients With Ulcerative Colitis: Patient-Level Pooled Analysis of Data From Randomized Trials. Clin Gastroenterol Hepatol. 2021 05; 19(5):939-946.e4. https://pubmed.ncbi.nlm.nih.gov/32371165
129. Chepelev LL, Wang X, Gold B, Bonzel CL, Rybicki F, Uyeda JW, Sheikh A, Anderson D, Lindaman J, Mogel G, Mitsouras D, Mahoney MC, **Cai T**, Rybicki FJ. Improved Appropriateness of Advanced Diagnostic Imaging After Implementation of Clinical Decision Support Mechanism. J Digit Imaging. 2021 04; 34(2):397-403. https://pubmed.ncbi.nlm.nih.gov/33634414
130. Huang S, Cai T, Weber BN, He Z, Dahal KP, Hong C, Hou J, Seyok T, Cagan A, DiCarli MF, Joseph J, Kim SC, Solomon DH, **Cai T**, Liao KP. The Association Between Inflammation, Incident Heart Failure, and Heart Failure Subtypes in Patients with Rheumatoid Arthritis. Arthritis Care Res (Hoboken). 2021 Oct 08. https://pubmed.ncbi.nlm.nih.gov/34623035

131. Kohane IS, Aronow BJ, Avillach P, Beaulieu-Jones BK, Bellazzi R, Bradford RL, Brat GA, Cannataro M, Cimino JJ, García-Barrio N, Gehlenborg N, Ghassemi M, Gutiérrez-Sacristán A, Hanauer DA, Holmes JH, Hong C, Klann JG, Loh NHW, Luo Y, Mandl KD, Daniar M, Moore JH, Murphy SN, Neuraz A, Ngiam KY, Omenn GS, Palmer N, Patel LP, Pedrera-Jiménez M, Sliz P, South AM, Tan ALM, Taylor DM, Taylor BW, Torti C, Vallejos AK, Wagholikar KB, Weber GM, **Cai T**. What Every Reader Should Know About Studies Using Electronic Health Record Data but May Be Afraid to Ask. J Med Internet Res. 2021 03 02; 23(3):e22219. https://pubmed.ncbi.nlm.nih.gov/33600347
132. Le TT, Gutiérrez-Sacristán A, Son J, Hong C, South AM, Beaulieu-Jones BK, Loh NHW, Luo Y, Morris M, Ngiam KY, Patel LP, Samayamuthu MJ, Schriver E, Tan AL, Moore J, **Cai T**, Omenn GS, Avillach P, Kohane IS, Visweswaran S, Mowery DL, Xia Z. Multinational Prevalence of Neurological Phenotypes in Patients Hospitalized with COVID-19. medRxiv. 2021 Jan 29. https://pubmed.ncbi.nlm.nih.gov/33655281
133. Le TT, Gutiérrez-Sacristán A, Son J, Hong C, South AM, Beaulieu-Jones BK, Loh NHW, Luo Y, Morris M, Ngiam KY, Patel LP, Samayamuthu MJ, Schriver E, Tan ALM, Moore J, **Cai T**, Omenn GS, Avillach P, Kohane IS, Visweswaran S, Mowery DL, Xia Z. Multinational characterization of neurological phenotypes in patients hospitalized with COVID-19. Sci Rep. 2021 10 12; 11(1):20238. https://pubmed.ncbi.nlm.nih.gov/34642371
134. Tedeschi SK, Cai T, He Z, Ahuja Y, Hong C, Yates KA, Dahal K, Xu C, Lyu H, Yoshida K, Solomon DH, **Cai T**, Liao KP. Classifying Pseudogout Using Machine Learning Approaches With Electronic Health Record Data. Arthritis Care Res (Hoboken). 2021 03; 73(3):442-448. https://pubmed.ncbi.nlm.nih.gov/31910317
135. Weber GM, Hong C, Palmer NP, Avillach P, Murphy SN, Gutiérrez-Sacristán A, Xia Z, Serret-Larmande A, Neuraz A, Omenn GS, Visweswaran S, Klann JG, South AM, Loh NHW, Cannataro M, Beaulieu-Jones BK, Bellazzi R, Agapito G, Alessiani M, Aronow BJ, Bell DS, Bellasi A, Benoit V, Beraghi M, Boeker M, Booth J, Bosari S, Bourgeois FT, Brown NW, Bucalo M, Chiovato L, Chiudinelli L, Dagliati A, Devkota B, DuVall SL, Follett RW, Ganslandt T, García Barrio N, Gradinger T, Griffier R, Hanauer DA, Holmes JH, Horki P, Huling KM, Issitt RW, Jouhet V, Keller MS, Kraska D, Liu M, Luo Y, Lynch KE, Malovini A, Mandl KD, Mao C, Maram A, Matheny ME, Maulhardt T, Mazzitelli M, Milano M, Moore JH, Morris JS, Morris M, Mowery DL, Naughton TP, Ngiam KY, Norman JB, Patel LP, Pedrera Jimenez M, Ramoni RB, Schriver ER, Scudeller L, Sebire NJ, Serrano Balazote P, Spiridou A, Tan AL, Tan BW, Tibollo V, Torti C, Trecarichi EM, Vitacca M, Zambelli A, Zucco C, Kohane IS, **Cai T**, Brat GA. International Comparisons of Harmonized Laboratory Value Trajectories to Predict Severe COVID-19: Leveraging the 4CE Collaborative Across 342 Hospitals and 6 Countries: A Retrospective Cohort Study. medRxiv. 2021 Feb 05. https://pubmed.ncbi.nlm.nih.gov/33564777
136. Weber GM, Zhang HG, L'Yi S, Bonzel CL, Hong C, Avillach P, Gutiérrez-Sacristán A, Palmer NP, Tan ALM, Wang X, Yuan W, Gehlenborg N, Alloni A, Amendola DF, Bellasi A, Bellazzi R, Beraghi M, Bucalo M, Chiovato L, Cho K, Dagliati A, Estiri H, Follett RW, García Barrio N, Hanauer DA, Henderson DW, Ho YL, Holmes JH, Hutch MR, Kavuluru R, Kirchoff K, Klann JG, Krishnamurthy AK, Le TT, Liu M, Loh NHW, Lozano-Zahonero S, Luo Y, Maidlow S, Makoudjou A, Malovini A, Martins MR, Moal B, Morris M, Mowery DL, Murphy SN, Neuraz A, Ngiam KY, Okoshi MP, Omenn GS, Patel LP, Pedrera Jiménez M, Prudente RA, Samayamuthu MJ, Sanz Vidorreta FJ, Schriver ER, Schubert P, Serrano Balazote P, Tan BW, Tanni SE, Tibollo V, Visweswaran S, Wagholikar KB, Xia Z, Zöller D, Kohane IS, **Cai T**, South AM, Brat GA. International Changes in COVID-19 Clinical Trajectories Across 315 Hospitals and 6 Countries: Retrospective Cohort Study. J Med Internet Res. 2021 10 11; 23(10):e31400. https://pubmed.ncbi.nlm.nih.gov/34533459
137. Yuan Q, Cai T, Hong C, Du M, Johnson BE, Lanuti M, **Cai T**, Christiani DC. Performance of a Machine Learning Algorithm Using Electronic Health Record Data to Identify and Estimate Survival in a Longitudinal Cohort of Patients With Lung Cancer. JAMA Netw Open. 2021 Jul 01; 4(7):e2114723. https://pubmed.ncbi.nlm.nih.gov/34232304
138. Yuan W, Beaulieu-Jones BK, Yu KH, Lipnick SL, Palmer N, Loscalzo J, **Cai T**, Kohane IS. Temporal bias in case-control design: preventing reliable predictions of the future. Nat Commun. 2021 02 17; 12(1):1107. https://pubmed.ncbi.nlm.nih.gov/33597541
139. Zhang J, Can A, Lai PMR, Mukundan S, Castro VM, Dligach D, Finan S, Gainer V, Shadick N, Savova G, Murphy SN, **Cai T**, Weiss ST, Du R. Geometric variations associated with posterior communicating artery aneurysms. J Neurointerv Surg. 2021 Nov; 13(11):1049-1052. https://pubmed.ncbi.nlm.nih.gov/33479035
140. Zhang J, Can A, Lai PMR, Mukundan S, Castro VM, Dligach D, Finan S, Gainer VS, Shadick NA, Savova G, Murphy SN, **Cai T**, Weiss ST, Du R. Morphological variables associated with ruptured basilar tip aneurysms. Sci Rep. 2021 01 28; 11(1):2526. https://pubmed.ncbi.nlm.nih.gov/33510194
141. Zhang J, Lai PMR, Can A, Mukundan S, Castro VM, Dligach D, Finan S, Gainer VS, Shadick NA, Savova G,

Murphy SN, **Cai T**, Weiss ST, Du R. Tobacco use and age are associated with different morphologic features of anterior communicating artery aneurysms. Sci Rep. 2021 Feb 26; 11(1):4791. https://pubmed.ncbi.nlm.nih.gov/33307259

142. Zhang J, Lai PMR, Can A, Mukundan S, Castro VM, Dligach D, Finan S, Gainer VS, Shadick NA, Savova G, Murphy SN, **Cai T**, Weiss ST, Du R. Tobacco use and age are associated with different morphologic features of anterior communicating artery aneurysms. Sci Rep. 2021 Feb 26; 11(1):4791. https://pubmed.ncbi.nlm.nih.gov/33637879
143. Ananthakrishnan AN, Cagan A, **Cai T**, Gainer VS, Savova G, Shaw SY, Churchill S, Burke KE, Karlson EW, Murphy SN, Kohane I, Liao KP, Xavier RJ. Use of Narrative Concepts in Electronic Health Records to Validate Associations Between Genetic Factors and Response to Treatment of Inflammatory Bowel Diseases. Clin Gastroenterol Hepatol. 2020 07; 18(8):1890-1892. https://pubmed.ncbi.nlm.nih.gov/31404664
144. Beam AL, Kompa B, Schmaltz A, Fried I, Weber G, Palmer N, Shi X, **Cai T**, Kohane IS. Clinical Concept Embeddings Learned from Massive Sources of Multimodal Medical Data. Pac Symp Biocomput. 2020; 25:295-306. https://pubmed.ncbi.nlm.nih.gov/31797605
145. Brat GA, Weber GM, Gehlenborg N, Avillach P, Palmer NP, Chiovato L, Cimino J, Waitman LR, Omenn GS, Malovini A, Moore JH, Beaulieu-Jones BK, Tibollo V, Murphy SN, Yi SL, Keller MS, Bellazzi R, Hanauer DA, Serret-Larmande A, Gutierrez-Sacristan A, Holmes JJ, Bell DS, Mandl KD, Follett RW, Klann JG, Murad DA, Scudeller L, Bucalo M, Kirchoff K, Craig J, Obeid J, Jouhet V, Griffier R, Cossin S, Moal B, Patel LP, Bellasi A, Prokosch HU, Kraska D, Sliz P, Tan ALM, Ngiam KY, Zambelli A, Mowery DL, Schiver E, Devkota B, Bradford RL, Daniar M, Daniel C, Benoit V, Bey R, Paris N, Serre P, Orlova N, Dubiel J, Hilka M, Jannot AS, Breant S, Leblanc J, Griffon N, Burgun A, Bernaux M, Sandrin A, Salamanca E, Cormont S, Ganslandt T, Gradinger T, Champ J, Boeker M, Martel P, Esteve L, Gramfort A, Grisel O, Leprovost D, Moreau T, Varoquaux G, Vie JJ, Wassermann D, Mensch A, Caucheteux C, Haverkamp C, Lemaitre G, Bosari S, Krantz ID, South A, **Cai T**, Kohane IS. International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium. NPJ Digit Med. 2020; 3:109. https://pubmed.ncbi.nlm.nih.gov/32864472
146. Huang S, Huang J, Cai T, Dahal KP, Cagan A, He Z, Stratton J, Gorelik I, Hong C, **Cai T**, Liao KP. Impact of ICD10 and secular changes on electronic medical record rheumatoid arthritis algorithms. Rheumatology (Oxford). 2020 12 01; 59(12):3759-3766. https://pubmed.ncbi.nlm.nih.gov/32413107
147. Koch M, DeKosky ST, Goodman M, Sun J, Furtado JD, Fitzpatrick AL, Mackey RH, **Cai T**, Lopez OL, Kuller LH, Mukamal KJ, Jensen MK. Association of Apolipoprotein E in Lipoprotein Subspecies With Risk of Dementia. JAMA Netw Open. 2020 07 01; 3(7):e209250. https://pubmed.ncbi.nlm.nih.gov/32648923
148. Koch M, DeKosky ST, Goodman M, Sun J, Furtado JD, Fitzpatrick AL, Mackey RH, **Cai T**, Lopez OL, Kuller LH, Mukamal KJ, Jensen MK. High density lipoprotein and its apolipoprotein-defined subspecies and risk of dementia. J Lipid Res. 2020 03; 61(3):445-454. https://pubmed.ncbi.nlm.nih.gov/31892526
149. Sacks FM, Liang L, Furtado JD, **Cai T**, Davidson WS, He Z, McClelland RL, Rimm EB, Jensen MK. Protein-Defined Subspecies of HDLs (High-Density Lipoproteins) and Differential Risk of Coronary Heart Disease in 4 Prospective Studies. Arterioscler Thromb Vasc Biol. 2020 11; 40(11):2714-2727. https://pubmed.ncbi.nlm.nih.gov/32907368
150. Sonabend W A, Pellegrini AM, Chan S, Brown HE, Rosenquist JN, Vuijk PJ, Doyle AE, Perlis RH, **Cai T**. Integrating questionnaire measures for transdiagnostic psychiatric phenotyping using word2vec. PLoS One. 2020; 15(4):e0230663. https://pubmed.ncbi.nlm.nih.gov/32243452
151. Wang Y, Zhu J, Aroner S, Overvad K, **Cai T**, Yang M, Tjønneland A, Handberg A, Jensen MK. Plasma CD36 and Incident Diabetes: A Case-Cohort Study in Danish Men and Women. Diabetes Metab J. 2020 02; 44(1):134-142. https://pubmed.ncbi.nlm.nih.gov/31701685
152. Zhang J, Can A, Lai PMR, Mukundan S, Castro VM, Dligach D, Finan S, Gainer VS, Shadick NA, Savova G, Murphy SN, **Cai T**, Weiss ST, Du R. Surrounding vascular geometry associated with basilar tip aneurysm formation. Sci Rep. 2020 10 21; 10(1):17928. https://pubmed.ncbi.nlm.nih.gov/33087795
153. Zhang J, Can A, Lai PMR, Mukundan S, Castro VM, Dligach D, Finan S, Yu S, Gainer VS, Shadick NA, Savova G, Murphy SN, **Cai T**, Weiss ST, Du R. Age and morphology of posterior communicating artery aneurysms. Sci Rep. 2020 07 14; 10(1):11545. https://pubmed.ncbi.nlm.nih.gov/32665589
154. Zhang T, Goodman M, Zhu F, Healy B, Carruthers R, Chitnis T, Weiner H, **Cai T**, De Jager P, Tremlett H, Xia Z. Phenome-wide examination of comorbidity burden and multiple sclerosis disease severity. Neurol Neuroimmunol Neuroinflamm. 2020 11; 7(6). https://pubmed.ncbi.nlm.nih.gov/32817202
155. Zhao SS, Hong C, Cai T, Xu C, Huang J, Ermann J, Goodson NJ, Solomon DH, **Cai T**, Liao KP. Incorporating natural language processing to improve classification of axial spondyloarthritis using electronic health records. Rheumatology (Oxford). 2020 05 01; 59(5):1059-1065. https://pubmed.ncbi.nlm.nih.gov/31535693

156. Cai T, Zhang L, Yang N, Kumamaru KK, Rybicki FJ, **Cai T**, Liao KP. EXTraction of EMR numerical data: an efficient and generalizable tool to EXTEND clinical research. BMC Med Inform Decis Mak. 2019 11 15; 19(1):226. https://pubmed.ncbi.nlm.nih.gov/31730484
157. Can A, Lai PMR, Castro VM, Yu S, Dligach D, Finan S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Decreased Total Iron Binding Capacity May Correlate with Ruptured Intracranial Aneurysms. Sci Rep. 2019 04 15; 9(1):6054. https://pubmed.ncbi.nlm.nih.gov/30988354
158. Chen CY, Lee P, Castro V, Minnier J, Murphy S, Gainer V, **Cai T**, Sklar P, Perlis R, Smoller J. M27 - genetic validation of bipolar disorder identified by automated phenotyping using electronic health records. European Neuropsychopharmacology. 2019; S968-S969. https://doi.org/10.1016/j.euroneuro.2017.08.334
159. Jorge A, Castro VM, Barnado A, Gainer V, Hong C, **Cai T**, Cai T, Carroll R, Denny JC, Crofford L, Costenbader KH, Liao KP, Karlson EW, Feldman CH. Identifying lupus patients in electronic health records: Development and validation of machine learning algorithms and application of rule-based algorithms. Semin Arthritis Rheum. 2019 08; 49(1):84-90. https://pubmed.ncbi.nlm.nih.gov/30665626
160. Zhang J, Can A, Mukundan S, Steigner M, Castro VM, Dligach D, Finan S, Yu S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Wang Z, Weiss ST, Du R. Morphological Variables Associated With Ruptured Middle Cerebral Artery Aneurysms. Neurosurgery. 2019 07 01; 85(1):75-83. https://pubmed.ncbi.nlm.nih.gov/29850834
161. Zhong QY, Gelaye B, Karlson EW, Avillach P, Smoller JW, **Cai T**, Williams MA. Associations of antepartum suicidal behaviour with adverse infant and obstetric outcomes. Paediatr Perinat Epidemiol. 2019 03; 33(2):137-144. https://pubmed.ncbi.nlm.nih.gov/30790331
162. Zhong QY, Mittal LP, Nathan MD, Brown KM, Knudson González D, Cai T, Finan S, Gelaye B, Avillach P, Smoller JW, Karlson EW, **Cai T**, Williams MA. Use of natural language processing in electronic medical records to identify pregnant women with suicidal behavior: towards a solution to the complex classification problem. Eur J Epidemiol. 2019 Feb; 34(2):153-162. https://pubmed.ncbi.nlm.nih.gov/30535584
163. **Cai T**, Zhang Y, Ho YL, Link N, Sun J, Huang J, Cai TA, Damrauer S, Ahuja Y, Honerlaw J, Huang J, Costa L, Schubert P, Hong C, Gagnon D, Sun YV, Gaziano JM, Wilson P, Cho K, Tsao P, O'Donnell CJ, Liao KP. Association of Interleukin 6 Receptor Variant With Cardiovascular Disease Effects of Interleukin 6 Receptor Blocking Therapy: A Phenome-Wide Association Study. JAMA Cardiol. 2018 09 01; 3(9):849-857. https://pubmed.ncbi.nlm.nih.gov/30090940
164. Aghayev A, Giannopoulos AA, Gronsbell J, George E, **Cai T**, Steigner ML, Mitsouras D, Rybicki FJ. Common First-Pass CT Angiography Findings Associated With Rapid Growth Rate in Abdominal Aorta Aneurysms Between 3 and 5 cm in Largest Diameter. AJR Am J Roentgenol. 2018 Feb; 210(2):431-437. https://pubmed.ncbi.nlm.nih.gov/29261347
165. Can A, Castro VM, Dligach D, Finan S, Yu S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Elevated International Normalized Ratio Is Associated With Ruptured Aneurysms. Stroke. 2018 09; 49(9):2046-2052. https://pubmed.ncbi.nlm.nih.gov/30354989
166. Can A, Castro VM, Dligach D, Finan S, Yu S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Lipid-Lowering Agents and High HDL (High-Density Lipoprotein) Are Inversely Associated With Intracranial Aneurysm Rupture. Stroke. 2018 05; 49(5):1148-1154. https://pubmed.ncbi.nlm.nih.gov/29622625
167. Can A, Castro VM, Ozdemir YH, Dagen S, Dligach D, Finan S, Yu S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Alcohol Consumption and Aneurysmal Subarachnoid Hemorrhage. Transl Stroke Res. 2018 02; 9(1):13-19. https://pubmed.ncbi.nlm.nih.gov/28752411
168. Can A, Castro VM, Ozdemir YH, Dagen S, Dligach D, Finan S, Yu S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Heroin Use Is Associated with Ruptured Saccular Aneurysms. Transl Stroke Res. 2018 08; 9(4):340-346. https://pubmed.ncbi.nlm.nih.gov/29103103
169. Can A, Castro VM, Yu S, Dligach D, Finan S, Gainer VS, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Antihyperglycemic Agents Are Inversely Associated With Intracranial Aneurysm Rupture. Stroke. 2018 01; 49(1):34-39. https://pubmed.ncbi.nlm.nih.gov/29203688
170. Can A, Rudy RF, Castro VM, Dligach D, Finan S, Yu S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Low Serum Calcium and Magnesium Levels and Rupture of Intracranial Aneurysms. Stroke. 2018 07; 49(7):1747-1750. https://pubmed.ncbi.nlm.nih.gov/29844027
171. Can A, Rudy RF, Castro VM, Yu S, Dligach D, Finan S, Gainer V, Shadick NA, Savova G, Murphy S, **Cai T**, Weiss ST, Du R. Association between aspirin dose and subarachnoid hemorrhage from saccular aneurysms: A case-control study. Neurology. 2018 09 18; 91(12):e1175-e1181. https://pubmed.ncbi.nlm.nih.gov/30135253
172. Chen CY, Lee PH, Castro VM, Minnier J, Charney AW, Stahl EA, Ruderfer DM, Murphy SN, Gainer V, **Cai T**, Jones I, Pato CN, Pato MT, Landén M, Sklar P, Perlis RH, Smoller JW. Genetic validation of bipolar

disorder identified by automated phenotyping using electronic health records. Transl Psychiatry. 2018 04 18; 8(1):86. https://pubmed.ncbi.nlm.nih.gov/29666432

173. Furtado JD, Yamamoto R, Melchior JT, Andraski AB, Gamez-Guerrero M, Mulcahy P, He Z, **Cai T**, Davidson WS, Sacks FM. Distinct Proteomic Signatures in 16 HDL (High-Density Lipoprotein) Subspecies. Arterioscler Thromb Vasc Biol. 2018 12; 38(12):2827-2842. https://pubmed.ncbi.nlm.nih.gov/30571168
174. Hejblum BP, Cui J, Lahey LJ, Cagan A, Sparks JA, Sokolove J, **Cai T**, Liao KP. Association Between Anti-Citrullinated Fibrinogen Antibodies and Coronary Artery Disease in Rheumatoid Arthritis. Arthritis Care Res (Hoboken). 2018 07; 70(7):1113-1117. https://pubmed.ncbi.nlm.nih.gov/28992379
175. Imran TF, Posner D, Honerlaw J, Vassy JL, Song RJ, Ho YL, Kittner SJ, Liao KP, **Cai T**, O'Donnell CJ, Djousse L, Gagnon DR, Gaziano JM, Wilson PW, Cho K. A phenotyping algorithm to identify acute ischemic stroke accurately from a national biobank: the Million Veteran Program. Clin Epidemiol. 2018; 10:1509-1521. https://pubmed.ncbi.nlm.nih.gov/30425582
176. McCoy TH, Castro VM, Hart KL, Pellegrini AM, Yu S, **Cai T**, Perlis RH. Genome-wide Association Study of Dimensional Psychopathology Using Electronic Health Records. Biol Psychiatry. 2018 06 15; 83(12):1005-1011. https://pubmed.ncbi.nlm.nih.gov/29496196
177. McCoy TH, Yu S, Hart KL, Castro VM, Brown HE, Rosenquist JN, Doyle AE, Vuijk PJ, **Cai T**, Perlis RH. High Throughput Phenotyping for Dimensional Psychopathology in Electronic Health Records. Biol Psychiatry. 2018 06 15; 83(12):997-1004. https://pubmed.ncbi.nlm.nih.gov/29496195
178. Qian B, Rudy RF, **Cai T**, Du R. Cerebral Artery Diameter in Inbred Mice Varies as a Function of Strain. Front Neuroanat. 2018; 12:10. https://pubmed.ncbi.nlm.nih.gov/29515376
179. Zhong QY, Gelaye B, Smoller JW, Avillach P, **Cai T**, Williams MA. Adverse obstetric outcomes during delivery hospitalizations complicated by suicidal behavior among US pregnant women. PLoS One. 2018; 13(2):e0192943. https://pubmed.ncbi.nlm.nih.gov/29447245
180. Zhong QY, Karlson EW, Gelaye B, Finan S, Avillach P, Smoller JW, **Cai T**, Williams MA. Screening pregnant women for suicidal behavior in electronic medical records: diagnostic codes vs. clinical notes processed by natural language processing. BMC Med Inform Decis Mak. 2018 05 29; 18(1):30. https://pubmed.ncbi.nlm.nih.gov/29843698
181. Aroner SA, Yang M, Li J, Furtado JD, Sacks FM, Tjønneland A, Overvad K, **Cai T**, Jensen MK. Apolipoprotein C-III and High-Density Lipoprotein Subspecies Defined by Apolipoprotein C-III in Relation to Diabetes Risk. Am J Epidemiol. 2017 Sep 15; 186(6):736-744. https://pubmed.ncbi.nlm.nih.gov/28520887
182. Can A, Castro VM, Ozdemir YH, Dagen S, Yu S, Dligach D, Finan S, Gainer V, Shadick NA, Murphy S, **Cai T**, Savova G, Dammers R, Weiss ST, Du R. Association of intracranial aneurysm rupture with smoking duration, intensity, and cessation. Neurology. 2017 Sep 26; 89(13):1408-1415. https://pubmed.ncbi.nlm.nih.gov/28855408
183. Castro VM, Dligach D, Finan S, Yu S, Can A, Abd-El-Barr M, Gainer V, Shadick NA, Murphy S, **Cai T**, Savova G, Weiss ST, Du R. Large-scale identification of patients with cerebral aneurysms using natural language processing. Neurology. 2017 Jan 10; 88(2):164-168. https://pubmed.ncbi.nlm.nih.gov/27927935
184. Geva A, Gronsbell JL, **Cai T**, Cai T, Murphy SN, Lyons JC, Heinz MM, Natter MD, Patibandla N, Bickel J, Mullen MP, Mandl KD. A Computable Phenotype Improves Cohort Ascertainment in a Pediatric Pulmonary Hypertension Registry. J Pediatr. 2017 09; 188:224-231.e5. https://pubmed.ncbi.nlm.nih.gov/28625502
185. Liao KP, Sparks JA, Hejblum BP, Kuo IH, Cui J, Lahey LJ, Cagan A, Gainer VS, Liu W, Cai TT, Sokolove J, **Cai T**. Phenome-Wide Association Study of Autoantibodies to Citrullinated and Noncitrullinated Epitopes in Rheumatoid Arthritis. Arthritis Rheumatol. 2017 04; 69(4):742-749. https://pubmed.ncbi.nlm.nih.gov/27792870
186. Naxerova K, Reiter JG, Brachtel E, Lennerz JK, van de Wetering M, Rowan A, **Cai T**, Clevers H, Swanton C, Nowak MA, Elledge SJ, Jain RK. Origins of lymphatic and distant metastases in human colorectal cancer. Science. 2017 Jul 07; 357(6346):55-60. https://pubmed.ncbi.nlm.nih.gov/28684519
187. Ong MS, Mullen MP, Austin ED, Szolovits P, Natter MD, Geva A, **Cai T**, Kong SW, Mandl KD. Learning a Comorbidity-Driven Taxonomy of Pediatric Pulmonary Hypertension. Circ Res. 2017 Aug 04; 121(4):341-353. https://pubmed.ncbi.nlm.nih.gov/28611076
188. Stein MB, Ware EB, Mitchell C, Chen CY, Borja S, **Cai T**, Dempsey CL, Fullerton CS, Gelernter J, Heeringa SG, Jain S, Kessler RC, Naifeh JA, Nock MK, Ripke S, Sun X, Beckham JC, Kimbrel NA, Ursano RJ, Smoller JW. Genomewide association studies of suicide attempts in US soldiers. . 2017 Dec; 174(8):786-797. https://pubmed.ncbi.nlm.nih.gov/28902444
189. Xie W, Agniel D, Shevchenko A, Malov SV, Svitin A, Cherkasov N, Baum MK, Campa A, Gaseitsiwe S, Bussmann H, Makhema J, Marlink R, Novitsky V, Lee TH, **Cai T**, O'Brien SJ, Essex M. Genome-Wide

Analyses Reveal Gene Influence on HIV Disease Progression and HIV-1C Acquisition in Southern Africa. AIDS Res Hum Retroviruses. 2017 06; 33(6):597-609. https://pubmed.ncbi.nlm.nih.gov/28132517

190. Ananthakrishnan AN, Cagan A, **Cai T**, Gainer VS, Shaw SY, Churchill S, Karlson EW, Murphy SN, Liao KP, Kohane I. Statin Use Is Associated With Reduced Risk of Colorectal Cancer in Patients With Inflammatory Bowel Diseases. Clin Gastroenterol Hepatol. 2016 07; 14(7):973-9. https://pubmed.ncbi.nlm.nih.gov/26905907
191. Ananthakrishnan AN, Cagan A, **Cai T**, Gainer VS, Shaw SY, Savova G, Churchill S, Karlson EW, Kohane I, Liao KP, Murphy SN. Comparative Effectiveness of Infliximab and Adalimumab in Crohn's Disease and Ulcerative Colitis. Inflamm Bowel Dis. 2016 Apr; 22(4):880-5. https://pubmed.ncbi.nlm.nih.gov/26933751
192. Ananthakrishnan AN, Cagan A, **Cai T**, Gainer VS, Shaw SY, Savova G, Churchill S, Karlson EW, Murphy SN, Liao KP, Kohane I. Identification of Nonresponse to Treatment Using Narrative Data in an Electronic Health Record Inflammatory Bowel Disease Cohort. Inflamm Bowel Dis. 2016 Jan; 22(1):151-8. https://pubmed.ncbi.nlm.nih.gov/26332313
193. Kumamaru KK, Kumamaru H, Bateman BT, Gronsbell J, **Cai T**, Liu J, Higgins LD, Aoki S, Ohtomo K, Rybicki FJ, Patorno E. Limited Hospital Variation in the Use and Yield of CT for Pulmonary Embolism in Patients Undergoing Total Hip or Total Knee Replacement Surgery. Radiology. 2016 Dec; 281(3):826-834. https://pubmed.ncbi.nlm.nih.gov/27228331
194. Stein MB, Chen CY, Ursano RJ, **Cai T**, Gelernter J, Heeringa SG, Jain S, Jensen KP, Maihofer AX, Mitchell C, Nievergelt CM, Nock MK, Neale BM, Polimanti R, Ripke S, Sun X, Thomas ML, Wang Q, Ware EB, Borja S, Kessler RC, Smoller JW. Genome-wide Association Studies of Posttraumatic Stress Disorder in 2 Cohorts of US Army Soldiers. JAMA Psychiatry. 2016 07 01; 73(7):695-704. https://pubmed.ncbi.nlm.nih.gov/27167565
195. Zhong QY, Gelaye B, Miller M, Fricchione GL, **Cai T**, Johnson PA, Henderson DC, Williams MA. Suicidal behavior-related hospitalizations among pregnant women in the USA, 2006-2012. Arch Womens Ment Health. 2016 06; 19(3):463-72. https://pubmed.ncbi.nlm.nih.gov/26680447
196. Ananthakrishnan AN, Cagan A, **Cai T**, Gainer VS, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP, Xavier RJ. Common Genetic Variants Influence Circulating Vitamin D Levels in Inflammatory Bowel Diseases. Inflamm Bowel Dis. 2015 Nov; 21(11):2507-14. https://pubmed.ncbi.nlm.nih.gov/26241000
197. Ananthakrishnan AN, Cagan A, **Cai T**, Gainer VS, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP. Colonoscopy is associated with a reduced risk for colon cancer and mortality in patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2015 Feb; 13(2):322-329.e1. https://pubmed.ncbi.nlm.nih.gov/25041865
198. Ananthakrishnan AN, Cheng A, Cagan A, **Cai T**, Gainer VS, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP. Mode of childbirth and long-term outcomes in women with inflammatory bowel diseases. Dig Dis Sci. 2015 Feb; 60(2):471-7. https://pubmed.ncbi.nlm.nih.gov/25213079
199. Castro V, Shen Y, Yu S, Finan S, Pau CT, Gainer V, Keefe CC, Savova G, Murphy SN, **Cai T**, Welt CK. Identification of subjects with polycystic ovary syndrome using electronic health records. Reprod Biol Endocrinol. 2015 Oct 29; 13:116. https://pubmed.ncbi.nlm.nih.gov/26510685
200. Castro VM, Minnier J, Murphy SN, Kohane I, Churchill SE, Gainer V, **Cai T**, Hoffnagle AG, Dai Y, Block S, Weill SR, Nadal-Vicens M, Pollastri AR, Rosenquist JN, Goryachev S, Ongur D, Sklar P, Perlis RH, Smoller JW. Validation of electronic health record phenotyping of bipolar disorder cases and controls. Am J Psychiatry. 2015 Apr; 172(4):363-72. https://pubmed.ncbi.nlm.nih.gov/25827034
201. Kessler RC, Warner CH, Ivany C, Petukhova MV, Rose S, Bromet EJ, Brown M, **Cai T**, Colpe LJ, Cox KL, Fullerton CS, Gilman SE, Gruber MJ, Heeringa SG, Lewandowski-Romps L, Li J, Millikan-Bell AM, Naifeh JA, Nock MK, Rosellini AJ, Sampson NA, Schoenbaum M, Stein MB, Wessely S, Zaslavsky AM, Ursano RJ. Predicting suicides after psychiatric hospitalization in US Army soldiers: the Army Study To Assess Risk and rEsilience in Servicemembers (Army STARRS). JAMA Psychiatry. 2015 Jan; 72(1):49-57. https://pubmed.ncbi.nlm.nih.gov/25390793
202. Liao KP, **Cai T**, Savova GK, Murphy SN, Karlson EW, Ananthakrishnan AN, Gainer VS, Shaw SY, Xia Z, Szolovits P, Churchill S, Kohane I. Development of phenotype algorithms using electronic medical records and incorporating natural language processing. BMJ. 2015 Apr 24; 350:h1885. https://pubmed.ncbi.nlm.nih.gov/25911572
203. Liao KP, Ananthakrishnan AN, Kumar V, Xia Z, Cagan A, Gainer VS, Goryachev S, Chen P, Savova GK, Agniel D, Churchill S, Lee J, Murphy SN, Plenge RM, Szolovits P, Kohane I, Shaw SY, Karlson EW, **Cai T**. Methods to Develop an Electronic Medical Record Phenotype Algorithm to Compare the Risk of Coronary Artery Disease across 3 Chronic Disease Cohorts. PLoS One. 2015; 10(8):e0136651. https://pubmed.ncbi.nlm.nih.gov/26301417

204. Uno H, Wittes J, Fu H, Solomon SD, Claggett B, Tian L, **Cai T**, Pfeffer MA, Evans SR, Wei LJ. Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies. Ann Intern Med. 2015 Jul 21; 163(2):127-34. https://pubmed.ncbi.nlm.nih.gov/26054047
205. Ananthakrishnan AN, Cagan A, Gainer VS, Cheng SC, **Cai T**, Scoville E, Konijeti GG, Szolovits P, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP. Thromboprophylaxis is associated with reduced post-hospitalization venous thromboembolic events in patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2014 Nov; 12(11):1905-10. https://pubmed.ncbi.nlm.nih.gov/24632349
206. Ananthakrishnan AN, Cagan A, Gainer VS, Cheng SC, **Cai T**, Szolovits P, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP. Mortality and extraintestinal cancers in patients with primary sclerosing cholangitis and inflammatory bowel disease. J Crohns Colitis. 2014 Sep; 8(9):956-63. https://pubmed.ncbi.nlm.nih.gov/24559536
207. Ananthakrishnan AN, Cheng SC, **Cai T**, Cagan A, Gainer VS, Szolovits P, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP. Association between reduced plasma 25-hydroxy vitamin D and increased risk of cancer in patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2014 May; 12(5):821-7. https://pubmed.ncbi.nlm.nih.gov/24161349
208. Ananthakrishnan AN, Cheng SC, **Cai T**, Cagan A, Gainer VS, Szolovits P, Shaw SY, Churchill S, Karlson EW, Murphy SN, Kohane I, Liao KP. Serum inflammatory markers and risk of colorectal cancer in patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2014 Aug; 12(8):1342-8.e1. https://pubmed.ncbi.nlm.nih.gov/24407106
209. Lai PS, Hang JQ, Zhang FY, Lin X, Zheng BY, Dai HL, Su L, **Cai T**, Christiani DC. Gender differences in the effect of occupational endotoxin exposure on impaired lung function and death: the Shanghai Textile Worker Study. Occup Environ Med. 2014 Feb; 71(2):118-125. https://pubmed.ncbi.nlm.nih.gov/24297825
210. Liao KP, Diogo D, Cui J, **Cai T**, Okada Y, Gainer VS, Murphy SN, Gupta N, Mirel D, Ananthakrishnan AN, Szolovits P, Shaw SY, Raychaudhuri S, Churchill S, Kohane I, Karlson EW, Plenge RM. Association between low density lipoprotein and rheumatoid arthritis genetic factors with low density lipoprotein levels in rheumatoid arthritis and non-rheumatoid arthritis controls. Ann Rheum Dis. 2014 Jun; 73(6):1170-5. https://pubmed.ncbi.nlm.nih.gov/23716066
211. Ong MS, Kohane IS, **Cai T**, Gorman MP, Mandl KD. Population-level evidence for an autoimmune etiology of epilepsy. JAMA Neurol. 2014 May; 71(5):569-74. https://pubmed.ncbi.nlm.nih.gov/24687183
212. van Loo HM, **Cai T**, Gruber MJ, Li J, de Jonge P, Petukhova M, Rose S, Sampson NA, Schoevers RA, Wardenaar KJ, Wilcox MA, Al-Hamzawi AO, Andrade LH, Bromet EJ, Bunting B, Fayyad J, Florescu SE, Gureje O, Hu C, Huang Y, Levinson D, Medina-Mora ME, Nakane Y, Posada-Villa J, Scott KM, Xavier M, Zarkov Z, Kessler RC. Major depressive disorder subtypes to predict long-term course. Depress Anxiety. 2014 Sep; 31(9):765-77. https://pubmed.ncbi.nlm.nih.gov/24425049
213. Yu S, Kumamaru KK, George E, Dunne RM, Bedayat A, Neykov M, Hunsaker AR, Dill KE, **Cai T**, Rybicki FJ. Classification of CT pulmonary angiography reports by presence, chronicity, and location of pulmonary embolism with natural language processing. J Biomed Inform. 2014 Dec; 52:386-93. https://pubmed.ncbi.nlm.nih.gov/25117751
214. Altschuler GM, Hofmann O, Kalatskaya I, Payne R, Ho Sui SJ, Saxena U, Krivtsov AV, Armstrong SA, **Cai T**, Stein L, Hide WA. Pathprinting: An integrative approach to understand the functional basis of disease. Genome Med. 2013; 5(7):68. https://pubmed.ncbi.nlm.nih.gov/23890051
215. Ananthakrishnan AN, **Cai T**, Savova G, Cheng SC, Chen P, Perez RG, Gainer VS, Murphy SN, Szolovits P, Xia Z, Shaw S, Churchill S, Karlson EW, Kohane I, Plenge RM, Liao KP. Improving case definition of Crohn's disease and ulcerative colitis in electronic medical records using natural language processing: a novel informatics approach. Inflamm Bowel Dis. 2013 Jun; 19(7):1411-20. https://pubmed.ncbi.nlm.nih.gov/23567779
216. Ananthakrishnan AN, Cagan A, Gainer VS, **Cai T**, Cheng SC, Savova G, Chen P, Szolovits P, Xia Z, De Jager PL, Shaw SY, Churchill S, Karlson EW, Kohane I, Plenge RM, Murphy SN, Liao KP. Normalization of plasma 25-hydroxy vitamin D is associated with reduced risk of surgery in Crohn's disease. Inflamm Bowel Dis. 2013 Aug; 19(9):1921-7. https://pubmed.ncbi.nlm.nih.gov/23751398
217. Ananthakrishnan AN, Gainer VS, **Cai T**, Perez RG, Cheng SC, Savova G, Chen P, Szolovits P, Xia Z, De Jager PL, Shaw S, Churchill S, Karlson EW, Kohane I, Perlis RH, Plenge RM, Murphy SN, Liao KP. Similar risk of depression and anxiety following surgery or hospitalization for Crohn's disease and ulcerative colitis. Am J Gastroenterol. 2013 Apr; 108(4):594-601. https://pubmed.ncbi.nlm.nih.gov/23337479
218. Ananthakrishnan AN, Gainer VS, Perez RG, **Cai T**, Cheng SC, Savova G, Chen P, Szolovits P, Xia Z, De Jager PL, Shaw SY, Churchill S, Karlson EW, Kohane I, Perlis RH, Plenge RM, Murphy SN, Liao KP.

Psychiatric co-morbidity is associated with increased risk of surgery in Crohn's disease. Aliment Pharmacol Ther. 2013 Feb; 37(4):445-54. https://pubmed.ncbi.nlm.nih.gov/23289600

219. Liao KP, **Cai T**, Gainer VS, Cagan A, Murphy SN, Liu C, Churchill S, Shaw SY, Kohane I, Solomon DH, Plenge RM, Karlson EW. Lipid and lipoprotein levels and trend in rheumatoid arthritis compared to the general population. Arthritis Care Res (Hoboken). 2013 Dec; 65(12):2046-50. https://pubmed.ncbi.nlm.nih.gov/23925980
220. Liao KP, Kurreeman F, Li G, Duclos G, Murphy S, Guzman R, **Cai T**, Gupta N, Gainer V, Schur P, Cui J, Denny JC, Szolovits P, Churchill S, Kohane I, Karlson EW, Plenge RM. Associations of autoantibodies, autoimmune risk alleles, and clinical diagnoses from the electronic medical records in rheumatoid arthritis cases and non-rheumatoid arthritis controls. Arthritis Rheum. 2013 Mar; 65(3):571-81. https://pubmed.ncbi.nlm.nih.gov/23233247
221. Xia Z, Secor E, Chibnik LB, Bove RM, Cheng S, Chitnis T, Cagan A, Gainer VS, Chen PJ, Liao KP, Shaw SY, Ananthakrishnan AN, Szolovits P, Weiner HL, Karlson EW, Murphy SN, Savova GK, **Cai T**, Churchill SE, Plenge RM, Kohane IS, De Jager PL. Modeling disease severity in multiple sclerosis using electronic health records. PLoS One. 2013; 8(11):e78927. https://pubmed.ncbi.nlm.nih.gov/24244385
222. Ananthakrishnan AN, Guzman-Perez R, Gainer V, **Cai T**, Churchill S, Kohane I, Plenge RM, Murphy S. Predictors of severe outcomes associated with Clostridium difficile infection in patients with inflammatory bowel disease. Aliment Pharmacol Ther. 2012 Apr; 35(7):789-95. https://pubmed.ncbi.nlm.nih.gov/22360370
223. Carroll RJ, Thompson WK, Eyler AE, Mandelin AM, **Cai T**, Zink RM, Pacheco JA, Boomershine CS, Lasko TA, Xu H, Karlson EW, Perez RG, Gainer VS, Murphy SN, Ruderman EM, Pope RM, Plenge RM, Kho AN, Liao KP, Denny JC. Portability of an algorithm to identify rheumatoid arthritis in electronic health records. J Am Med Inform Assoc. 2012 Jun; 19(e1):e162-9. https://pubmed.ncbi.nlm.nih.gov/22374935
224. Lu MT, Demehri S, **Cai T**, Parast L, Hunsaker AR, Goldhaber SZ, Rybicki FJ. Axial and reformatted four-chamber right ventricle-to-left ventricle diameter ratios on pulmonary CT angiography as predictors of death after acute pulmonary embolism. AJR Am J Roentgenol. 2012 Jun; 198(6):1353-60. https://pubmed.ncbi.nlm.nih.gov/22623548
225. Perlis RH, Iosifescu DV, Castro VM, Murphy SN, Gainer VS, Minnier J, **Cai T**, Goryachev S, Zeng Q, Gallagher PJ, Fava M, Weilburg JB, Churchill SE, Kohane IS, Smoller JW. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. Psychol Med. 2012 Jan; 42(1):41-50. https://pubmed.ncbi.nlm.nih.gov/21682950
226. Kurreeman F, Liao K, Chibnik L, Hickey B, Stahl E, Gainer V, Li G, Bry L, Mahan S, Ardlie K, Thomson B, Szolovits P, Churchill S, Murphy SN, **Cai T**, Raychaudhuri S, Kohane I, Karlson E, Plenge RM. Genetic basis of autoantibody positive and negative rheumatoid arthritis risk in a multi-ethnic cohort derived from electronic health records. Am J Hum Genet. 2011 Jan 07; 88(1):57-69. https://pubmed.ncbi.nlm.nih.gov/21211616
227. Love TJ, **Cai T**, Karlson EW. Validation of psoriatic arthritis diagnoses in electronic medical records using natural language processing. Semin Arthritis Rheum. 2011 Apr; 40(5):413-20. https://pubmed.ncbi.nlm.nih.gov/20701955
228. Qi L, Parast L, **Cai T**, Powers C, Gervino EV, Hauser TH, Hu FB, Doria A. Genetic susceptibility to coronary heart disease in type 2 diabetes: 3 independent studies. J Am Coll Cardiol. 2011 Dec 13; 58(25):2675-82. https://pubmed.ncbi.nlm.nih.gov/22152955
229. Gao X, Lai CQ, Scott T, Shen J, **Cai T**, Ordovas JM, Tucker KL. Urinary 8-hydroxy-2-deoxyguanosine and cognitive function in Puerto Rican adults. Am J Epidemiol. 2010 Aug 01; 172(3):271-8. https://pubmed.ncbi.nlm.nih.gov/20621990
230. Hunsaker AR, Oliva IB, **Cai T**, Trotman-Dickenson B, Gill RR, Hatabu H, Rybicki FJ. Contrast opacification using a reduced volume of iodinated contrast material and low peak kilovoltage in pulmonary CT angiography: Objective and subjective evaluation. AJR Am J Roentgenol. 2010 Aug; 195(2):W118-24. https://pubmed.ncbi.nlm.nih.gov/20651170
231. Liao KP, **Cai T**, Gainer V, Goryachev S, Zeng-treitler Q, Raychaudhuri S, Szolovits P, Churchill S, Murphy S, Kohane I, Karlson EW, Plenge RM. Electronic medical records for discovery research in rheumatoid arthritis. Arthritis Care Res (Hoboken). 2010 Aug; 62(8):1120-7. https://pubmed.ncbi.nlm.nih.gov/20235204
232. Rybicki FJ, Mitsouras D, Owens CD, Whitmore A, Gerhard-Herman M, Wake N, **Cai T**, Zhou Q, Conte MS, Creager MA, Mulkern RV. Multi-contrast high spatial resolution black blood inner volume three-dimensional fast spin echo MR imaging in peripheral vein bypass grafts. Int J Cardiovasc Imaging. 2010 Aug; 26(6):683-91. https://pubmed.ncbi.nlm.nih.gov/20333469
233. Steigner ML, Mitsouras D, Whitmore AG, Otero HJ, Wang C, Buckley O, Levit NA, Hussain AZ, **Cai T**, Mather RT, Smedby O, DiCarli MF, Rybicki FJ. Iodinated contrast opacification gradients in normal coronary

arteries imaged with prospectively ECG-gated single heart beat 320-detector row computed tomography. Circ Cardiovasc Imaging. 2010 Mar; 3(2):179-86. https://pubmed.ncbi.nlm.nih.gov/20044512

234. Wu C, Sodickson A, **Cai T**, Levit NA, Steigner ML, Rybicki FJ, Ledbetter SM. Comparison of respiratory motion artifact from craniocaudal versus caudocranial scanning with 64-MDCT pulmonary angiography. AJR Am J Roentgenol. 2010 Jul; 195(1):155-9. https://pubmed.ncbi.nlm.nih.gov/20566810
235. Chien K, **Cai T**, Hsu H, Su T, Chang W, Chen M, Lee Y, Hu FB. A prediction model for type 2 diabetes risk among Chinese people. Diabetologia. 2009 Mar; 52(3):443-50. https://pubmed.ncbi.nlm.nih.gov/19057891
236. Cornelis MC, Qi L, Zhang C, Kraft P, Manson J, **Cai T**, Hunter DJ, Hu FB. Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of European ancestry. Ann Intern Med. 2009 Apr 21; 150(8):541-50. https://pubmed.ncbi.nlm.nih.gov/19380854
237. Lu MT, **Cai T**, Ersoy H, Whitmore AG, Levit NA, Goldhaber SZ, Rybicki FJ. Comparison of ECG-gated versus non-gated CT ventricular measurements in thirty patients with acute pulmonary embolism. Int J Cardiovasc Imaging. 2009 Jan; 25(1):101-7. https://pubmed.ncbi.nlm.nih.gov/18626787
238. Steigner ML, Otero HJ, **Cai T**, Mitsouras D, Nallamshetty L, Whitmore AG, Ersoy H, Levit NA, Di Carli MF, Rybicki FJ. Narrowing the phase window width in prospectively ECG-gated single heart beat 320-detector row coronary CT angiography. Int J Cardiovasc Imaging. 2009 Jan; 25(1):85-90. https://pubmed.ncbi.nlm.nih.gov/18663599
239. Wu CJ, **Cai T**, Rikova K, Merberg D, Kasif S, Steffen M. A predictive phosphorylation signature of lung cancer. PLoS One. 2009 Nov 25; 4(11):e7994. https://pubmed.ncbi.nlm.nih.gov/19946374
240. Lu MT, **Cai T**, Ersoy H, Whitmore AG, Quiroz R, Goldhaber SZ, Rybicki FJ. Interval increase in right-left ventricular diameter ratios at CT as a predictor of 30-day mortality after acute pulmonary embolism: initial experience. Radiology. 2008 Jan; 246(1):281-7. https://pubmed.ncbi.nlm.nih.gov/18096540
241. Rybicki FJ, Otero HJ, Steigner ML, Vorobiof G, Nallamshetty L, Mitsouras D, Ersoy H, Mather RT, Judy PF, **Cai T**, Coyner K, Schultz K, Whitmore AG, Di Carli MF. Initial evaluation of coronary images from 320-detector row computed tomography. Int J Cardiovasc Imaging. 2008 Jun; 24(5):535-46. https://pubmed.ncbi.nlm.nih.gov/18368512
242. Ersoy H, Goldhaber SZ, **Cai T**, Luu T, Rosebrook J, Mulkern R, Rybicki F. Time-resolved MR angiography: a primary screening examination of patients with suspected pulmonary embolism and contraindications to administration of iodinated contrast material. AJR Am J Roentgenol. 2007 May; 188(5):1246-54. https://pubmed.ncbi.nlm.nih.gov/17449767
243. MacKenzie JD, Nazario-Larrieu J, **Cai T**, Ledbetter MS, Duran-Mendicuti MA, Judy PF, Rybicki FJ. Reduced-dose CT: effect on reader evaluation in detection of pulmonary embolism. AJR Am J Roentgenol. 2007 Dec; 189(6):1371-9. https://pubmed.ncbi.nlm.nih.gov/18029873
244. Chuang SK, **Cai T**. Predicting clustered dental implant survival using frailty methods. J Dent Res. 2006 Dec; 85(12):1147-51. https://pubmed.ncbi.nlm.nih.gov/17122171
245. Chuang SK, **Cai T**, Douglass CW, Wei LJ, Dodson TB. Frailty approach for the analysis of clustered failure time observations in dental research. J Dent Res. 2005 Jan; 84(1):54-8. https://pubmed.ncbi.nlm.nih.gov/15615876
246. Rogus JJ, **Cai T**, Wei LJ. Issues in genomic screening: critical values, sample sizes, and the ability to detect linkage. Genet Epidemiol. 1999; 17 Suppl 1:S697-701. https://pubmed.ncbi.nlm.nih.gov/10597516

Recent Invited Talks

- Marvin Zelen lecture, Eastern Mediterranean Region of the International Biometric Society (April, 2021)
- Harvard Catalyst Biostatistics Symposium on Precision Medicine (April, 2021)
- ACM Conference on Health, Inference, and Learning, Keynote presentation (April, 2021)
- International Vilnius Conference on Probability Theory & Mathematical Statistics, Vilnius, Lithuania (Jul, 2018)
- Conference of the International Society for Nonparametric Statistics, Salerno, Italy (Jun, 2018)
- 10th DIA China Annual Meeting, Beijing, China (May 2018)
- University Hospital of Lausanne, Switzerland (Apr 2018)
- Department of Biomedical Data Science, Stanford University (Mar 2018)
- Department of Biostatistics, McGill University, Canada (Mar 2018)
- Department of Biostatistics, Yale University (Mar 2018)
- Workshop on “Perspectives and Analysis Methods for Personalized Medicine”, Institute for Mathematical Science, Singapore (Jul 2017)

- Workshop on “High-Dimensional Statistical Modeling and Analysis in the Age of Big Data”, Peking University, Peking China (May 2017)
- Department of Biostatistics, Vanderbilt University (Apr 2017)
- Michigan Institute for Data Science, University of Michigan (Mar 2017)
- Keynote Speaker, International Conference on Robust Statistics, Geneva Switzerland (Jul 2016)
- 3rd conference of the International Society for Non-Parametric Statistics, Avignon France (Jun 2016)
- Memorial Sloan Kettering Cancer Center (Oct 2015)
- Joint Statistical Meeting (Aug 2015)
- 60th World Statistics Congress, Rio de Janeiro, Brazil (Jul, 2015)
- IMS-Microsoft Research Workshop: Foundations of Data Science (Jun, 2015)
- Big Data Science symposium, University of Ghent, Belgium (May 2015)
- Plenary Speaker, First International Conference on Big Data & Applied Statistics, Beijing China (Nov 2014)
- Distinguished Lecture, MD Anderson Cancer Center (Nov 2014)
- University of Michigan (Oct 2014)
- International Society of Clinical Biostatistics, Vienna Austria (Aug 2014)
- IMS, Sydney Australia (Jul 2014)
- Banff workshop on Statistical genetics (Jun 2014)
- University of Wisconsin Madison (Apr 2014)
- Columbia University (Feb 2014)
- NIH/NICHD (Jan 2014)
- ICSA International Symposium (Dec 2013)
- University of Huston (Dec 2013)
- University of Washington (Oct 2013)
- University of Georgia (Oct 2013)
- FDA, Center for Devices and Radiological Health (Sept 2013)
- IMS China, Chengdu, China (Jul 2013)
- INSERM Workshop on the Evaluation of predictive models, France (May 2013)
- Eastern Mediterranean Region IBS conference (Apr 2013)
- Joint Statistical Meeting (Aug 2013)
- ENAR Spring Meeting (Mar 2013)
- Rand Corporation (Jan 2013)
- Department of Industrial and Systems Engineering, USC (Jan 2013)
- Johnson and Johnson Anual Statistical Conference (Nov 2012)
- Division of Biostatistics, Albert Eistein College of Medicine (Nov 2012)
- FDA/Industry Workshop 2011, Washington DC (Sep 2012)
- Joint Statistical Meeting (Aug 2012)
- Workshop on statistical inference in complex/high-dimensional problems, Vienna, Austria (Jul 2012)
- ISBIS, Bankok Thailand (Jun 2012)
- ICSA Applied Symposium, Boston MA (Jun 2012)
- Conference on “Emerging Statistical Issues in Biomarker Validation for Clinical Trials University of Pennsylvania (Apr 2012)
- Department of Biostatistics, Yale University (Apr 2012)
- Department of Biostatistics, Johns Hopkins University (Apr 2012)
- National Cancer Institute, NIH (Mar 2012)
- Department of Biostatistics, UNC (Mar 2012)
- DIA/FDA Co-sponsored Conference on Tailored Therapeutics (Oct 2011)
- Risk Assessment Conference, University of Maryland, College Park (Oct 2011)
- IMS Asia Pacific Rim Meeting, Tokyo, Japan (Jul 2011)
- Canada CRM workshop on Analysis of survival and event history data (May 2011)
- Dept of Statistics, University of Missouri (Apr 2011)
- 2011 International Indian Statistical Association Conference (Apr 2011)
- ASA Connecticut Chapter annual mini-conference, (Mar 2011)
- Dept of Statistics, Brown University (Mar 2011)
- Dept of Statistics, University of Connecticut (Feb 2011)
- Dept of Mathematics and Statistics, Boston University (Dec 2010)
- Ann. Conf. of the International Society for Clinical Biostatistics, Montpellier, France (Aug 2010)

- Joint Statistical Meeting (Aug 2010)
- International Conference on Statistical Analysis of Complex Data, Kunming, China (Jul 2010)
- International Conference on Statistics and Society, Beijing, China (Jul 2010)
- First Joint Biostatistics Symposium, Beijing, China (Jul 2010)
- Dept of Statistics, University of Vienna, Austria (Jun 2010)
- DFG German Research Foundation Symposium "Statistical Modeling and Data Analysis in Clinical Epidemiology, Freiburg, Germany (Jun 2010)
- ENAR Spring Meeting (Mar 2010)
- Conference on "Resampling Methods and High Dimensional Data, Texas A& M (Mar 2010)
- Oberwolfach Workshop on "Statistical Issues in Prediction: what can be learned for individualized predictive medicine", Mathematisches Forschungsinstitut Oberwolfach,
- Dept of Statistics, University of Urbana Champaign (Sep 2009)
- Joint Statistical Meeting (Aug 2009)
- International Conference on "Nonparametric Modeling & Its Application Chengdu, China (Jul 2009)
- IMS Asia Pacific Rim Meeting, Korea (Jun 2009)
- Workshop on "Innovation and Inventiveness in Statistics Methodologies", Yale University (May 2009)
- Dept of Biostatistics, University of Michigan (Nov 2008)
- Dept of Biostatistics, University of Pittsburgh (Oct 2008)
- Joint Statistical Meeting (Aug 2008)
- IMS-China International Conference on Statistics and Probability (Jun, 2008)
- International Chinese Statistical Association Conference (Jun 2008)
- Breast Cancer Surveillance Consortium Statistical Meeting (Apr 2008)
- Fred Hutchinson Cancer Research Center (Apr 2008)
- Computational Seminar Series, Harvard Medical School (Mar 2008)
- ENAR International Biometrics Society (Mar 2008)
- University of Copenhagen, Copenhagen Denmark (Jan 2008)
- University of Pennsylvania (Nov 2007)
- Radcliffe Institute for Advanced Studies (Sep 2007)
- International Chinese Statistical Association International Conference (Jun 2007)
- Harvard Schering Plough Workshop (Jun 2007)
- Dept of Biostatistics, University of Rochester (May 2007)
- Dana Farber Biostatistics Computational Biology Dinner Workshop (Mar 2007)
- Dept of Statistics, University of Minneapolis (Nov 2006)
- Joint Statistical Meeting (Aug 2006)
- National Cancer Institute (Jun 2006)
- Fred Hutchinson Cancer Research Center (Sep 2005)
- Joint Statistical Meeting (Aug 2005)
- Division of Biostatistics, Yale University (Jan 2005)
- Division of Biostatistics, Boston University (Nov 2004)
- Dept of Biostatistics, Columbia University (Oct 2004)
- Division of Biostatistics, University of California at San Francisco (Sep 2004)
- Joint Statistical Meeting (Aug 2004)
- Dept of Statistics, University of Munich, Germany (Jun 2004)
- National Institute of Child Health and Human Development (May 2004)
- Dept of Biostatistics, Brown University (Mar 2004)