

Tianxi Cai

2021-03-25

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

- 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. A Sonabend, AM Pellegrini, S Chan, HE Brown, JN Rosenquist, PJ Vuijk, AE Doyle, RH Perlis, **T Cai**. [Integrating questionnaire measures for transdiagnostic psychiatric phenotyping using word2vec](#). *PLOS ONE* 2020;15(4):e0230663.
2. C Hong, Y Wang, **T Cai**. [A divide-and-conquer method for sparse risk prediction and evaluation](#). *Biostatistics* 2020.
3. L Parast, **T Cai**, L Tian. [Evaluating multiple surrogate markers with censored data](#). *Biometrics* 2020.
4. S Chan, X Wang, I Jazi^{c}, S Peskoe, Y Zheng, **T Cai**. [Developing and evaluating risk prediction models with panel current status data](#). *Biometrics* 2020.
5. Y Ahuja, D Zhou, Z He, J Sun, VM Castro, V Gainer, SN Murphy, C Hong, **T Cai**. [sureLDA: A multidisease automated phenotyping method for the electronic health record](#). *Journal of the American Medical Informatics Association* 2020;27(8):1235–1243.
6. BP Hejblum, GM Weber, KP Liao, NP Palmer, S Churchill, NA Shadick, P Szolovits, SN Murphy, IS Kohane, **T Cai**. [Probabilistic record linkage of de-identified research datasets with discrepancies using diagnosis codes](#). *Scientific Data* 2019;6(1).
7. C Hong, KP Liao, **T Cai**. [Semi-supervised validation of multiple surrogate outcomes with application to electronic medical records phenotyping](#). *Biometrics* 2019;75(1):78–89.
8. D Cheng, A Chakraborty, AN Ananthakrishnan, **T Cai**. [Estimating average treatment effects with a double-index propensity score](#). *Biometrics* 2019.
9. J Gronsbell, J Minnier, S Yu, K Liao, **T Cai**. [Automated feature selection of predictors in electronic medical records data](#). *Biometrics* 2019;75(1):268–277.
10. KP Liao, J Sun, TA Cai, N Link, C Hong, J Huang, JE Huffman, J Gronsbell, Y Zhang, Y Ho, V Castro, V Gainer, SN Murphy, CJ O'Donnell, JM Gaziano, K Cho, P Szolovits, IS Kohane, S Yu, **T Cai**. [High-throughput multimodal automated phenotyping \(MAP\) with application to PheWAS](#). *Journal of the American Medical Informatics Association* 2019;26(11):1255–1262.
11. L Parast, L Tian, **T Cai**. [Assessing the value of a censored surrogate outcome](#). *Lifetime Data Analysis* 2019;26(2):245–265.
12. L Parast, **T Cai**, L Tian. [Using a surrogate marker for early testing of a treatment effect](#). *Biometrics* 2019;75(4):1253–1263.
13. L Zhang, Y Zhang, T Cai, Y Ahuja, Z He, Y Ho, A Beam, K Cho, R Carroll, J Denny, I Kohane, K Liao, **T Cai**. [Automated grouping of medical codes via multiview banded spectral clustering](#). *Journal of Biomedical Informatics* 2019;100:103322.
14. SF Chan, BP Hejblum, A Chakraborty, **T Cai**. [Semi-supervised estimation of covariance with application to phenome-wide association studies with electronic medical records data](#). *Statistical Methods in Medical Research* 2019;29(2):455–465.
15. **T Cai**, TT Cai, K Liao, W Liu. [Large-Scale Simultaneous Testing of Cross-Covariance Matrices with Applications to PheWAS](#). *Statistica Sinica* 2019;29:983–1005.
16. W Ning, S Chan, A Beam, M Yu, A Geva, K Liao, M Mullen, KD Mandl, I Kohane, **T Cai**, S Yu. [Feature extraction for phenotyping from semantic and knowledge resources](#). *Journal of Biomedical Informatics* 2019;91:103122.
17. X Wang, L Parast, L Tian, **T Cai**. [Model-free approach to quantifying the proportion of treatment effect explained by a surrogate marker](#). *Biometrika* 2019;107(1):107–122.
18. Y Wang, C Hong, N Palmer, Q Di, J Schwartz, I Kohane, **T Cai**. [A fast divide-and-conquer sparse Cox](#)

- regression. *Biostatistics* 2019.
19. Y Zhang*, T Cai*, S Yu*, K Cho, C Hong, J Sun, J Huang, Y Ho, AN Ananthakrishnan, Z Xia, SY Shaw, V Gainer, V Castro, N Link, J Honerlaw, S Huang, D Gagnon, EW Karlson, RM Plenge, P Szolovits, G Savova, S Churchill, C O'Donnell, SN Murphy, JM Gaziano, I Kohane, **T Cai***, KP Liao*. [High-throughput phenotyping with electronic medical record data using a common semi-supervised approach \(PheCAP\)](#). *Nature Protocols* 2019;14(12):3426–3444. (*: contributed equally to the work)
 20. A Chakraborty, **T Cai***. [Efficient and adaptive linear regression in semi-supervised settings](#). *The Annals of Statistics* 2018;46(4):1541–1572. (*: a PhD student thesis paper)
 21. JA Sinnott, **T Cai***. [Pathway aggregation for survival prediction via multiple kernel learning](#). *Statistics in Medicine* 2018;37(16):2501–2515. (*: a PhD student thesis paper)
 22. M Maziarz, **T Cai**, L Qi, AS Lok, Y Zheng. [Evaluating longitudinal markers under two-phase study designs](#). *Biostatistics* 2018;20(3):485–498.
 23. MO Goodman, L Chibnik, **T Cai**. [Variance components genetic association test for zero-inflated count outcomes](#). *Genetic Epidemiology* 2018;43(1):82–101.
 24. Y Xia, **T Cai**, TT Cai. [Two-Sample Tests for High-Dimensional Linear Regression With an Application to Detecting Interactions](#). *Statistica Sinica* 2018;28:63–92.
 25. D Agniel, **T Cai***. [Analysis of Multiple Diverse Phenotypes via Semiparametric Canonical Correlation Analysis](#). *Biometrics* 2017;73(4):1254–1265. (*: a PhD student thesis paper)
 26. D Liu, **T Cai**, A Lok, Y Zheng. [Nonparametric Maximum Likelihood Estimators of Time-Dependent Accuracy Measures for Survival Outcome Under Two-Stage Sampling Designs](#). *Journal of the American Statistical Association* 2017;113(522):882–892.
 27. JL Gronsbell, **T Cai***. [Semi-supervised approaches to efficient evaluation of model prediction performance](#). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2017;80(3):579–594. (*: a PhD student thesis paper)
 28. L Parast, **T Cai**, L Tian. [Evaluating surrogate marker information using censored data](#). *Statistics in Medicine* 2017;36(11):1767–1782.
 29. QM Zhou, W Dai, Y Zheng, **T Cai**. [Robust dynamic risk prediction with longitudinal studies](#). *Statistical Theory and Related Fields* 2017;1(2):159–170.
 30. S Yu, Y Ma, J Gronsbell, T Cai, AN Ananthakrishnan, VS Gainer, SE Churchill, P Szolovits, SN Murphy, IS Kohane, KP Liao, **T Cai**. [Enabling phenotypic big data with PheNorm](#). *Journal of the American Medical Informatics Association* 2017;25(1):54–60.
 31. W Dai, M Yang, C Wang, **T Cai***. [Sequence robust association test for familial data](#). *Biometrics* 2017;73(3):876–884. (*: a PhD student thesis paper)
 32. Y Xia, **T Cai**, TT Cai. [Multiple Testing of Submatrices of a Precision Matrix With Applications to Identification of Between Pathway Interactions](#). *Journal of the American Statistical Association* 2017;113(521):328–339.
 33. Y Zheng, M Brown, A Lok, **T Cai**. [Improving efficiency in biomarker incremental value evaluation under two-phase designs](#). *The Annals of Applied Statistics* 2017;11(2):638–654.
 34. Y Zheng, **T Cai***. [Augmented estimation for t-year survival with censored regression models](#). *Biometrics* 2017;73(4):1169–1178. (*: a PhD student thesis paper)
 35. D Agniel, KP Liao, **T Cai***. [Estimation and testing for multiple regulation of multivariate mixed outcomes](#). *Biometrics* 2016;72(4):1194–1205. (*: a PhD student thesis paper)
 36. FH Yong, L Tian, S Yu, **T Cai**, LJ Wei. [Optimal stratification in outcome prediction using baseline information](#). *Biometrika* 2016;103(4):817–828.
 37. M Maziarz, P Heagerty, **T Cai**, Y Zheng. [On longitudinal prediction with time-to-event outcome: Comparison of modeling options](#). *Biometrics* 2016;73(1):83–93.
 38. M Neykov, JS Liu, **T Cai***. [L1-Regularized Least Squares for Support Recovery of High Dimensional Single Index Models with Gaussian Designs](#). *Journal of Machine Learning Research* 2016;17(87):1–37. (*: a PhD student thesis paper)
 39. M Neykov, JS Liu, **T Cai***. [On the Characterization of a Class of Fisher-Consistent Loss Functions and its Application to Boosting](#). *Journal of Machine Learning Research* 2016;17(70):1–32. (*: a PhD student thesis paper)
 40. Q He, **T Cai**, Y Liu, N Zhao, QE Harmon, LM Almli, EB Binder, SM Engel, KJ Ressler, KN Conneely, X Lin, MC Wu. [Prioritizing individual genetic variants after kernel machine testing using variable selection](#). *Genetic Epidemiology* 2016;40(8):722–731.
 41. R Payne, M Yang, Y Zheng, MK Jensen, **T Cai***. [Robust risk prediction with biomarkers under two-phase stratified cohort design](#). *Biometrics* 2016;72(4):1037–1045. (*: a PhD student thesis paper)

42. S Yu, A Chakraborty, KP Liao, T Cai, AN Ananthkrishnan, VS Gainer, SE Churchill, P Szolovits, SN Murphy, IS Kohane, **T Cai**. [Surrogate-assisted feature extraction for high-throughput phenotyping](#). *Journal of the American Medical Informatics Association* 2016;ocw135.
43. **T Cai**, TT Cai, A Zhang. [Structured Matrix Completion with Applications to Genomic Data Integration](#). *Journal of the American Statistical Association* 2016;111(514):621–633.
44. X Wang, Z Zhang, N Morris, **T Cai**, S Lee, C Wang, TW Yu, CA Walsh, X Lin. [Rare variant association test in family-based sequencing studies](#). *Briefings in Bioinformatics* 2016;bbw083.
45. Y Huang, **T Cai**, E Kim. [Integrative genomic testing of cancer survival using semiparametric linear transformation models](#). *Statistics in Medicine* 2016;35(16):2831–2844.
46. Y Shen, **T Cai***. [Identifying predictive markers for personalized treatment selection](#). *Biometrics* 2016;72(4):1017–1025. (*: a PhD student thesis paper, winner of the John Van Ryzin Award)
47. J Minnier, M Yuan, JS Liu, **T Cai***. [Risk Classification With an Adaptive Naive Bayes Kernel Machine Model](#). *Journal of the American Statistical Association* 2015;110(509):393–404. (*: a PhD student thesis paper)
48. QM Zhou, Y Zheng, LB Chibnik, EW Karlson, **T Cai**. [Assessing incremental value of biomarkers with multi-phase nested case-control studies](#). *Biometrics* 2015;71(4):1139–1149.
49. S Yu, KP Liao, SY Shaw, VS Gainer, SE Churchill, P Szolovits, SN Murphy, IS Kohane, **T Cai**. [Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources](#). *Journal of the American Medical Informatics Association* 2015;22(5):993–1000.
50. Y Shen, **T Cai**, Y Chen, Y Yang, J Chen. [Retrospective likelihood-based methods for analyzing case-cohort genetic association studies](#). *Biometrics* 2015;71(4):960–968.
51. Y Xia, **T Cai**, TT Cai. [Testing differential networks with applications to the detection of gene-gene interactions](#). *Biometrika* 2015;102(2):247–266.
52. JA Sinnott, W Dai, KP Liao, SY Shaw, AN Ananthkrishnan, VS Gainer, EW Karlson, S Churchill, P Szolovits, S Murphy, I Kohane, R Plenge, **T Cai**. [Improving the power of genetic association tests with imperfect phenotype derived from electronic medical records](#). *Human Genetics* 2014;133(11):1369–1382.
53. L Parast, L Tian, **T Cai***. [Landmark Estimation of Survival and Treatment Effect in a Randomized Clinical Trial](#). *Journal of the American Statistical Association* 2014;109(505):384–394. (*: a PhD student thesis paper)
54. RA Matsouaka, J Li, **T Cai***. [Evaluating marker-guided treatment selection strategies](#). *Biometrics* 2014;70(3):489–499. (*: a PhD student thesis paper)
55. Y Shen, KP Liao, **T Cai***. [Sparse kernel machine regression for ordinal outcomes](#). *Biometrics* 2014;71(1):63–70. (*: a PhD student thesis paper)
56. JA Sinnott, **T Cai***. [Omnibus Risk Assessment via Accelerated Failure Time Kernel Machine Modeling](#). *Biometrics* 2013;69(4):861–873. (*: a PhD student thesis paper)
57. L Parast, **T Cai***. [Landmark risk prediction of residual life for breast cancer survival](#). *Statistics in Medicine* 2013;32(20):3459–3471. (*: a PhD student thesis paper)
58. L Zhao, L Tian, **T Cai**, B Claggett, LJ Wei. [Effectively Selecting a Target Population for a Future Comparative Study](#). *Journal of the American Statistical Association* 2013;108(502):527–539.
59. QM Zhou, Y Zheng, **T Cai**. [Assessment of biomarkers for risk prediction with nested case-control studies](#). *Clinical Trials: Journal of the Society for Clinical Trials* 2013;10(5):677–679.
60. **T Cai**, L Tian, D Lloyd-Jones, LJ Wei. [Evaluating subject-level incremental values of new markers for risk classification rule](#). *Lifetime Data Analysis* 2013;19(4):547–567.
61. **T Cai**, Y Zheng. [Resampling Procedures for Making Inference Under Nested Case-Control Studies](#). *Journal of the American Statistical Association* 2013;108(504):1532–1544.
62. Y Zheng, **T Cai**, MS Pepe. [Adopting nested case-control quota sampling designs for the evaluation of risk markers](#). *Lifetime Data Analysis* 2013;19(4):568–588.
63. D Liu, **T Cai**, Y Zheng. [Evaluating the Predictive Value of Biomarkers with Stratified Case-Cohort Design](#). *Biometrics* 2012;68(4):1219–1227.
64. H Uno, L Tian, **T Cai**, IS Kohane, LJ Wei. [A unified inference procedure for a class of measures to assess improvement in risk prediction systems with survival data](#). *Statistics in Medicine* 2012;32(14):2430–2442.
65. L Leon, **T Cai***. [Model checking techniques for assessing functional form specifications in censored linear regression models](#). *Statistica Sinica* 2012;22(2). (*: a PhD student thesis paper)
66. L Parast, S Cheng, **T Cai***. [Landmark Prediction of Long-Term Survival Incorporating Short-Term Event Time Information](#). *Journal of the American Statistical Association* 2012;107(500):1492–1501. (*: a PhD student thesis paper)
67. L Tian, **T Cai**, L Zhao, LJ Wei. [On the covariate-adjusted estimation for an overall treatment difference with](#)

- data from a randomized comparative clinical trial. *Biostatistics* 2012;13(2):256–273.
68. QM Zhou, Y Zheng, **T Cai**. Subgroup specific incremental value of new markers for risk prediction. *Lifetime Data Analysis* 2012;19(2):142–169.
 69. **T Cai**, X Lin, RJ Carroll. Identifying genetic marker sets associated with phenotypes via an efficient adaptive score test. *Biostatistics* 2012;13(4):776–790.
 70. Y Zheng, L Parast, **T Cai**, M Brown. Evaluating incremental values from new predictors with net reclassification improvement in survival analysis. *Lifetime Data Analysis* 2012;19(3):350–370.
 71. H Uno, **T Cai**, L Tian, LJ Wei. Graphical Procedures for Evaluating Overall and Subject-Specific Incremental Values from New Predictors with Censored Event Time Data. *Biometrics* 2011;67(4):1389–1396.
 72. H Uno, **T Cai**, MJ Pencina, RB D’Agostino, LJ Wei. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Statistics in Medicine* 2011;30(10):1105–1117.
 73. J Minnier, L Tian, **T Cai***. A Perturbation Method for Inference on Regularized Regression Estimates. *Journal of the American Statistical Association* 2011;106(496):1371–1382. (*: a PhD student thesis paper)
 74. L Parast, S Cheng, **T Cai***. Incorporating short-term outcome information to predict long-term survival with discrete markers. *Biometrical Journal* 2011;53(2):294–307. (*: a PhD student thesis paper)
 75. MC Wu, S Lee, **T Cai**, Y Li, M Boehnke, X Lin. Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. *The American Journal of Human Genetics* 2011;89(1):82–93.
 76. **T Cai**, G Tonini, X Lin. Kernel Machine Approach to Testing the Significance of Multiple Genetic Markers for Risk Prediction. *Biometrics* 2011;67(3):975–986.
 77. **T Cai**, L Tian, DM Lloyd-Jones. Comparing costs associated with risk stratification rules for t-year survival. *Biostatistics* 2011;12(4):597–609.
 78. **T Cai**, Y Zheng. Evaluating prognostic accuracy of biomarkers in nested case-control studies. *Biostatistics* 2011;13(1):89–100.
 79. **T Cai**, Y Zheng. Nonparametric Evaluation of Biomarker Accuracy Under Nested Case-Control Studies. *Journal of the American Statistical Association* 2011;106(494):569–580.
 80. X Lin, **T Cai**, MC Wu, Q Zhou, G Liu, DC Christiani, X Lin. Kernel machine SNP-set analysis for censored survival outcomes in genome-wide association studies. *Genetic Epidemiology* 2011;35(7):620–631.
 81. Y Zheng, **T Cai**, Y Jin, Z Feng. Evaluating Prognostic Accuracy of Biomarkers under Competing Risk. *Biometrics* 2011;68(2):388–396.
 82. L Tian, R Wang, **T Cai**, L Wei. The Highest Confidence Density Region and Its Usage for Joint Inferences about Constrained Parameters. *Biometrics* 2010;67(2):604–610.
 83. R Wang, L Tian, **T Cai**, LJ Wei. Nonparametric inference procedure for percentiles of the random effects distribution in meta-analysis. *The Annals of Applied Statistics* 2010;4(1):520–532.
 84. S McDaniel, J Minnier, RA Betensky, G Mohapatra, Y Shen, JF Gusella, DN Louis, **T Cai***. Assessing Population Level Genetic Instability via Moving Average. *Statistics in Biosciences* 2010;2(2):120–136. (*: a PhD student thesis paper)
 85. **T Cai**, L Parast, L Ryan. Meta-analysis for rare events. *Statistics in Medicine* 2010;29(20):2078–2089.
 86. **T Cai**, L Tian, H Uno, SD Solomon, LJ Wei. Calibrating parametric subject-specific risk estimation. *Biometrika* 2010;97(2):389–404.
 87. **T Cai**, L Tian, PH Wong, LJ Wei. Analysis of randomized comparative clinical trial data for personalized treatment selections. *Biostatistics* 2010;12(2):270–282.
 88. **T Cai**, TA Gerds, Y Zheng, J Chen. Robust Prediction of t-Year Survival with Data from Multiple Studies. *Biometrics* 2010;67(2):436–444.
 89. LF León, **T Cai**, LJ Wei. Robust Inferences for Covariate Effects on Survival Time with Censored Linear Regression Models. *Statistics in Biosciences* 2009;1:50–64.
 90. Y Zheng, **T Cai**, JL Stanford, Z Feng. Semiparametric Models of Time-Dependent Predictive Values of Prognostic Biomarkers. *Biometrics* 2009;66(1):50–60.
 91. B Jiang, X Zhang, **T Cai**. Estimating the Confidence Interval for Prediction Errors of Support Vector Machine Classifiers. *Journal of Machine Learning Research* 2008;9:521–540.
 92. L Tian, **T Cai**, LJ Wei. Identifying Subjects Who Benefit from Additional Information for Better Prediction of the Outcome Variables. *Biometrics* 2008;65(3):894–902.
 93. L Tian, **T Cai**, MA Pfeffer, N Piankov, PY Cremieux, LJ Wei. 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* 2008;10(2):275–281.
 94. **T Cai**, J Huang, L Tian. Regularized Estimation for the Accelerated Failure Time Model. *Biometrics* 2008;65(2):394–404.

95. **T Cai**, L Tian, SD Solomon, LJ Wei. [Predicting future responses based on possibly mis-specified working models](#). *Biometrika* 2008;95(1):75–92.
96. **T Cai**, LE Dodd. [Regression analysis for the partial area under the ROC curve](#). *Statistica Sinica* 2008;18(3):817–836.
97. TA Gerds, **T Cai**, M Schumacher. [The Performance of Risk Prediction Models](#). *Biometrical Journal* 2008;50(4):457–479.
98. Y Zheng, **T Cai**, MS Pepe, WC Levy. [Time-Dependent Predictive Values of Prognostic Biomarkers With Failure Time Outcome](#). *Journal of the American Statistical Association* 2008;103(481):362–368.
99. H Uno, **T Cai**, L Tian, LJ Wei. [Evaluating Prediction Rules for Year Survivors With Censored Regression Models](#). *Journal of the American Statistical Association* 2007;102(478):527–537.
100. JH Ware, **T Cai**. [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 \(DOI: 10.1002/sim.2929\)](#). *Statistics in Medicine* 2007;27(2):185–187.
101. L Tian, **T Cai**, E Goetghebeur, LJ Wei. [Model evaluation based on the sampling distribution of estimated absolute prediction error](#). *Biometrika* 2007;94(2):297–311.
102. **T Cai**, S Cheng. [Robust combination of multiple diagnostic tests for classifying censored event times](#). *Biostatistics* 2007;9(2):216–233.
103. **T Cai**, Y Zheng. [Model Checking for ROC Regression Analysis](#). *Biometrics* 2007;63(1):152–163.
104. L Tian, **T Cai**. [On the accelerated failure time model for current status and interval censored data](#). *Biometrika* 2006;93(2):329–342.
105. **T Cai**, PB Gilbert, SG Self. [Joint Inferences on Vaccine Efficacy Against Infection and Disease with Application to the First HIV Vaccine Efficacy Trial](#). *Journal of Biopharmaceutical Statistics* 2006;16(4):517–538.
106. MS Pepe, **T Cai**, G Longton. [Combining Predictors for Classification Using the Area under the Receiver Operating Characteristic Curve](#). *Biometrics* 2005;62(1):221–229.
107. **T Cai**. [The sensitivity and specificity of markers for event times](#). *Biostatistics* 2005;7(2):182–197.
108. **T Cai**, L Tian, LJ Wei. [Semiparametric Box-Cox power transformation models for censored survival observations](#). *Biometrika* 2005;92(3):619–632.
109. Y Zheng, **T Cai**, Z Feng. [Application of the Time-Dependent ROC Curves for Prognostic Accuracy with Multiple Biomarkers](#). *Biometrics* 2005;62(1):279–287.
110. MS Pepe, **T Cai**. [The Analysis of Placement Values for Evaluating Discriminatory Measures](#). *Biometrics* 2004;60(2):528–535.
111. **T Cai**. [Semi-parametric ROC regression analysis with placement values](#). *Biostatistics* 2004;5(1):45–60.
112. **T Cai**. [Semiparametric regression analysis for doubly censored data](#). *Biometrika* 2004;91(2):277–290.
113. **T Cai**, CS Moskowitz. [Semi-parametric estimation of the binormal ROC curve for a continuous diagnostic test](#). *Biostatistics* 2004;5(4):573–586.
114. **T Cai**, RA Betensky. [Hazard Regression for Interval-Censored Data with Penalized Spline](#). *Biometrics* 2003;59(3):570–579.
115. **T Cai**, MS Pepe. [Semiparametric Receiver Operating Characteristic Analysis to Evaluate Biomarkers for Disease](#). *Journal of the American Statistical Association* 2002;97(460):1099–1107.
116. **T Cai**, RJ Hyndman, MP Wand. [Mixed Model-Based Hazard Estimation](#). *Journal of Computational and Graphical Statistics* 2002;11(4):784–798.
117. **T Cai**, SC Cheng, LJ Wei. [Semiparametric Mixed-Effects Models for Clustered Failure Time Data](#). *Journal of the American Statistical Association* 2002;97(458):514–522.
118. **T Cai**, LJ Wei. [Regression Analysis for Multivariate Failure Time Observations](#). *Asymptotics in Statistics and Probability* 2000;33–46.
119. **T Cai**, LJ Wei, M Wilcox. [Semiparametric regression analysis for clustered failure time data](#). *Biometrika* 2000;87(4):867–878.
120. **T Cai**, J Shen. [Boundedness Is Redundant in a Theorem of Daubechies](#). *Applied and Computational Harmonic Analysis* 1999;6(3):400–404.

Collaborative Research

1. A Geva, M Liu, VA Panickan, P Avillach, **T Cai**, KD Mandl. [A high-throughput phenotyping algorithm is portable from adult to pediatric populations](#). *Journal of the American Medical Informatics Association* 2021.
2. IS Kohane, BJ Aronow, P Avillach, BK Beaulieu-Jones, R Bellazzi, RL Bradford, GA Brat, M Cannataro, JJ Cimino, N Garcia-Barrio, N Gehlenborg, M Ghassemi, A Gutierrez-Sacristan, DA Hanauer, JH Holmes, C Hong, JG Klann, NH Loh, Y Luo, KD Mandl, M Daniar, JH Moore, SN Murphy, A Neuraz, KY

- Ngiam, GS Omenn, N Palmer, LP Patel, M Pedrera-Jimenez, P Sliz, AM South, AL Tan, DM Taylor, BW Taylor, C Torti, AK Vallejos, KB Waghlikar, GM Weber, **T Cai**. [What Every Reader Should Know About Studies Using Electronic Health Record Data but May Be Afraid to Ask](#). *Journal of Medical Internet Research* 2021;23(3):e22219.
3. J Zhang, A Can, PM Lai, JS Mukundan, VM Castro, D Dligach, S Finan, V Gainer, N Shadick, G Savova, SN Murphy, **T Cai**, ST Weiss, R Du. [Geometric variations associated with posterior communicating artery aneurysms](#). *Journal of {NeuroInterventional} Surgery* 2021;neurintsurg-2020-017062.
 4. J Zhang, A Can, PM Lai, S Mukundan, VM Castro, D Dligach, S Finan, VS Gainer, NA Shadick, G Savova, SN Murphy, **T Cai**, ST Weiss, R Du. [Morphological variables associated with ruptured basilar tip aneurysms](#). *Scientific Reports* 2021;11(1).
 5. J Zhang, A Can, PM Lai, S Mukundan, VM Castro, D Dligach, S Finan, VS Gainer, NA Shadick, G Savova, SN Murphy, **T Cai**, ST Weiss, R Du. [Vascular Geometry Associated with Anterior Communicating Artery Aneurysm Formation](#). *World Neurosurgery* 2021;146:e1318–e1325.
 6. J Zhang, PM Lai, A Can, S Mukundan, VM Castro, D Dligach, S Finan, VS Gainer, NA Shadick, G Savova, SN Murphy, **T Cai**, ST Weiss, R Du. [Tobacco use and age are associated with different morphologic features of anterior communicating artery aneurysms](#). *Scientific Reports* 2021;11(1).
 7. LL Chepelev, X Wang, B Gold, C Bonzel, F Rybicki, JW Uyeda, A Sheikh, D Anderson, J Lindaman, G Mogel, D Mitsouras, MC Mahoney, **T Cai**, FJ Rybicki. [Improved Appropriateness of Advanced Diagnostic Imaging After Implementation of Clinical Decision Support Mechanism](#). *Journal of Digital Imaging* 2021.
 8. SK Tedeschi, T Cai, Z He, Y Ahuja, C Hong, KA Yates, K Dahal, C Xu, H Lyu, K Yoshida, DH Solomon, **T Cai**, KP Liao. [Classifying Pseudogout Using Machine Learning Approaches With Electronic Health Record Data](#). *Arthritis Care & Research* 2021;73(3):442–448.
 9. TT Le, A Gutiérrez-Sacristán, J Son, C Hong, AM South, BK Beaulieu-Jones, NH Loh, Y Luo, M Morris, KY Ngiam, LP Patel, MJ Samayamuthu, E Schriver, AL Tan, J Moore, **T Cai**, GS Omenn, P Avillach, IS Kohane, S Visweswaran, DL Mowery, Z Xia. [Multinational Prevalence of Neurological Phenotypes in Patients Hospitalized with {COVID}-19](#). ____ 2021.
 10. W Yuan, BK Beaulieu-Jones, K Yu, SL Lipnick, N Palmer, J Loscalzo, **T Cai**, IS Kohane. [Temporal bias in case-control design: preventing reliable predictions of the future](#). *Nature Communications* 2021;12(1).
 11. Y Ahuja, N Kim, L Liang, T Cai, K Dahal, T Seyok, C Lin, S Finan, K Liao, G Savova, T Chitnis, **T Cai**, Z Xia. [Leveraging electronic health records data to predict multiple sclerosis disease activity](#). *Annals of Clinical and Translational Neurology* 2021.
 12. AL Beam, B Kompa, A Schmaltz, I Fried, G Weber, N Palmer, X Shi, **T Cai**, IS Kohane. [Clinical Concept Embeddings Learned from Massive Sources of Multimodal Medical Data](#). *Pac Symp Biocomput* 2020;25:295–306.
 13. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, G Savova, SY Shaw, S Churchill, KE Burke, EW Karlson, SN Murphy, I Kohane, KP Liao, RJ Xavier. [Use of Narrative Concepts in Electronic Health Records to Validate Associations Between Genetic Factors and Response to Treatment of Inflammatory Bowel Diseases](#). *Clinical Gastroenterology and Hepatology* 2020;18(8):1890–1892.
 14. AW Sonabend, AM Pellegrini, S Chan, HE Brown, JN Rosenquist, PJ Vuijk, AE Doyle, RH Perlis, **T Cai**. [Integrating questionnaire measures for transdiagnostic psychiatric phenotyping using word2vec](#). *PLOS ONE* 2020;15(4):e0230663.
 15. D Cheng, KC Cushing, **T Cai**, AN Ananthakrishnan. [Safety and Efficacy of Tumor Necrosis Factor Antagonists in Older Patients With Ulcerative Colitis: Patient-Level Pooled Analysis of Data From Randomized Trials](#). *Clinical Gastroenterology and Hepatology* 2020.
 16. FM Sacks, L Liang, JD Furtado, **T Cai**, WS Davidson, Z He, RL McClelland, EB Rimm, MK Jensen. [Protein-Defined Subspecies of {HDLs} \(High-Density Lipoproteins\) and Differential Risk of Coronary Heart Disease in 4 Prospective Studies](#). *Arteriosclerosis, Thrombosis, and Vascular Biology* 2020;40(11):2714–2727.
 17. GA Brat, GM Weber, N Gehlenborg, P Avillach, NP Palmer, L Chiovato, J Cimino, LR Waitman, GS Omenn, A Malovini, JH Moore, BK Beaulieu-Jones, V Tibollo, SN Murphy, SL Yi, MS Keller, R Bellazzi, DA Hanauer, A Serret-Larmande, A Gutierrez-Sacristan, JJ Holmes, DS Bell, KD Mandl, RW Follett, JG Klann, DA Murad, L Scudeller, M Bucalo, K Kirchoff, J Craig, J Obeid, V Jouhet, R Griffier, S Cossin, B Moal, LP Patel, A Bellasi, HU Prokosch, D Kraska, P Sliz, AL Tan, KY Ngiam, A Zambelli, DL Mowery, E Schiver, B Devkota, RL Bradford, M Daniar, C Daniel, V Benoit, R Bey, N Paris, P Serre, N Orlova, J Dubiel, M Hilka, AS Jannot, S Breant, J Leblanc, N Griffon, A Burgun, M Bernaux, A Sandrin, E Salamanca, S Cormont, T Ganslandt, T Gradinger, J Champ, M Boeker, P Martel, L Esteve, A Gramfort, O Grisel, D Leprovost, T Moreau, G Varoquaux, J Vie, D Wassermann, A Mensch, C Caucheteux, C Haverkamp, G Lemaitre, S Bosari,

- ID Krantz, A South, **T Cai**, IS Kohane. [International electronic health record-derived {COVID}-19 clinical course profiles: the 4CE consortium.](#) *npj Digital Medicine* 2020;3(1).
18. GM Weber, C Hong, NP Palmer, P Avillach, SN Murphy, A Guti{'e}rrez-Sacrist{'a}n, Z Xia, A Serret-Larmande, A Neuraz, GS Omenn, S Visweswaran, JG Klann, AM South, NH Loh, M Cannataro, BK Beaulieu-Jones, R Bellazzi, G Agapito, M Alessiani, BJ Aronow, DS Bell, A Bellasi, V Benoit, M Beraghi, M Boeker, J Booth, S Bosari, FT Bourgeois, NW Brown, M Bucalo, L Chiovato, L Chiudinelli, A Dagliati, B Devkota, SL DuVall, RW Follett, T Ganslandt, NG Barrio, T Gradinger, R Griffier, DA Hanauer, JH Holmes, P Horki, KM Huling, RW Issitt, V Jouhet, MS Keller, D Kraska, M Liu, Y Luo, KE Lynch, A Malovini, KD Mandl, C Mao, A Maram, ME Matheny, T Maulhardt, M Mazzitelli, M Milano, JH Moore, JS Morris, M Morris, DL Mowery, TP Naughton, KY Ngiam, JB Norman, LP Patel, MP Jimenez, RB Ramoni, ER Schriver, L Scudeller, NJ Sebire, PS Balazote, A Spiridou, AL Tan, BW Tan, V Tibollo, C Torti, EM Trecarichi, M Vitacca, A Zambelli, C Zucco, IS Kohane, **T Cai**, GA Brat. [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.](#) ____ 2020.
 19. IS Beam. [Clinical Concept Embeddings Learned from Massive Sources of Multimodal Medical Data.](#) *Pac Symp Biocomput* 2020;25:295–306.
 20. J Zhang, A Can, PM Lai, S Mukundan, VM Castro, D Dligach, S Finan, S Yu, VS Gainer, NA Shadick, G Savova, SN Murphy, **T Cai**, ST Weiss, R Du. [Age and morphology of posterior communicating artery aneurysms.](#) *Scientific Reports* 2020;10(1).
 21. J Zhang, A Can, PM Lai, S Mukundan, VM Castro, D Dligach, S Finan, VS Gainer, NA Shadick, G Savova, SN Murphy, **T Cai**, ST Weiss, R Du. [Surrounding vascular geometry associated with basilar tip aneurysm formation.](#) *Scientific Reports* 2020;10(1).
 22. M Koch, ST DeKosky, M Goodman, J Sun, JD Furtado, AL Fitzpatrick, RH Mackey, **T Cai**, OL Lopez, LH Kuller, KJ Mukamal, MK Jensen. [Association of Apolipoprotein E in Lipoprotein Subspecies With Risk of Dementia.](#) *JAMA Network Open* 2020;3(7):e209250.
 23. M Koch, ST DeKosky, M Goodman, J Sun, JD Furtado, AL Fitzpatrick, RH Mackey, **T Cai**, OL Lopez, LH Kuller, KJ Mukamal, MK Jensen. [High density lipoprotein and its apolipoprotein-defined subspecies and risk of dementia.](#) *Journal of Lipid Research* 2020;61(3):445–454.
 24. S Huang, J Huang, T Cai, KP Dahal, A Cagan, Z He, J Stratton, I Gorelik, C Hong, **T Cai**, KP Liao. [Impact of {ICD}10 and secular changes on electronic medical record rheumatoid arthritis algorithms.](#) *Rheumatology* 2020;59(12):3759–3766.
 25. SK Tedeschi, T Cai, Z He, Y Ahuja, C Hong, KA Yates, K Dahal, C Xu, H Lyu, K Yoshida, DH Solomon, **T Cai**, KP Liao. [Classifying Pseudogout using Machine Learning Approaches with Electronic Health Record Data.](#) *Arthritis Care & Research* 2020.
 26. T Zhang, M Goodman, F Zhu, B Healy, R Carruthers, T Chitnis, H Weiner, **T Cai**, P De Jager, H Tremlett, Z Xia. [Phenome-wide examination of comorbidity burden and multiple sclerosis disease severity.](#) *Neurology - Neuroimmunology Neuroinflammation* 2020;7(6):e864.
 27. Y Wang, J Zhu, S Aroner, K Overvad, **T Cai**, M Yang, A Tjønneland, A Handberg, MK Jensen. [Plasma CD36 and Incident Diabetes: A Case-Cohort Study in Danish Men and Women.](#) *Diabetes & Metabolism Journal* 2020;44(1):134.
 28. A Can, PM Lai, VM Castro, S Yu, D Dligach, S Finan, V Gainer, NA Shadick, G Savova, S Murphy, **T Cai**, ST Weiss, R Du. [Decreased Total Iron Binding Capacity May Correlate with Ruptured Intracranial Aneurysms.](#) *Scientific Reports* 2019;9(1):6054.
 29. A Jorge, VM Castro, A Barnado, V Gainer, C Hong, **T Cai**, T Cai, R Carroll, JC Denny, L Crofford, KH Costenbader, KP Liao, EW Karlson, CH Feldman. [Identifying lupus patients in electronic health records: Development and validation of machine learning algorithms and application of rule-based algorithms.](#) *Seminars in Arthritis and Rheumatism* 2019;49(1):84–90.
 30. M Koch, ST DeKosky, M Goodman, J Sun, JD Furtado, AL Fitzpatrick, RH Mackey, **T Cai**, OL Lopez, LH Kuller, KJ Mukamal, MK Jensen. [High density lipoprotein and its apolipoprotein-defined subspecies and risk of dementia.](#) *Journal of Lipid Research* 2019;61(3):445–454.
 31. Q Zhong, B Gelaye, EW Karlson, P Avillach, JW Smoller, **T Cai**, MA Williams. [Associations of antepartum suicidal behaviour with adverse infant and obstetric outcomes.](#) *Paediatric and Perinatal Epidemiology* 2019;33(2):137–144.
 32. SS Zhao, C Hong, T Cai, C Xu, J Huang, J Ermann, NJ Goodson, DH Solomon, **T Cai**, KP Liao. [Incorporating natural language processing to improve classification of axial spondyloarthritis using electronic health records.](#) *Rheumatology* 2019;59(5):1059–1065.

33. SS Zhao, C Hong, T Cai, C Xu, J Huang, J Ermann, NJ Goodson, DH Solomon, **T Cai**, KP Liao. [Incorporating natural language processing to improve classification of axial spondyloarthritis using electronic health records.](#) *Rheumatology* 2019;59(5):1059–1065.
34. T Cai, L Zhang, N Yang, KK Kumamaru, FJ Rybicki, **T Cai**, KP Liao. [EXtraction of EMR numerical data: an efficient and generalizable tool to EXTEND clinical research.](#) *BMC Medical Informatics and Decision Making* 2019;19(1):226.
35. A Aghayev, AA Giannopoulos, J Gronsbell, E George, **T Cai**, ML Steigner, D Mitsouras, FJ Rybicki. [Common First-Pass CT Angiography Findings Associated With Rapid Growth Rate in Abdominal Aorta Aneurysms Between 3 and 5 cm in Largest Diameter.](#) *American Journal of Roentgenology* 2018;210(2):431–437.
36. A Can, RF Rudy, VM Castro, S Yu, D Dligach, S Finan, V Gainer, NA Shadick, G Savova, S Murphy, **T Cai**, ST Weiss, R Du. [Association between aspirin dose and subarachnoid hemorrhage from saccular aneurysms.](#) *Neurology* 2018;91(12):e1175–e1181.
37. A Can, VM Castro, S Yu, D Dligach, S Finan, VS Gainer, NA Shadick, G Savova, S Murphy, **T Cai**, ST Weiss, R Du. [Antihyperglycemic Agents Are Inversely Associated With Intracranial Aneurysm Rupture.](#) *Stroke* 2018;49(1):34–39.
38. B Qian, RF Rudy, **T Cai**, R Du. [Cerebral Artery Diameter in Inbred Mice Varies as a Function of Strain.](#) *Frontiers in Neuroanatomy* 2018;12.
39. BP Hejblum, J Cui, LJ Lahey, A Cagan, JA Sparks, J Sokolove, **T Cai**, KP Liao. [Association Between Anti-Citrullinated Fibrinogen Antibodies and Coronary Artery Disease in Rheumatoid Arthritis.](#) *Arthritis Care & Research* 2018;70(7):1113–1117.
40. J Zhang, A Can, S Mukundan, M Steigner, VM Castro, D Dligach, S Finan, S Yu, V Gainer, NA Shadick, G Savova, S Murphy, **T Cai**, Z Wang, ST Weiss, R Du. [Morphological Variables Associated With Ruptured Middle Cerebral Artery Aneurysms.](#) *Neurosurgery* 2018;85(1):75–83.
41. JD Furtado, R Yamamoto, JT Melchior, AB Andraski, M Gamez-Guerrero, P Mulcahy, Z He, **T Cai**, WS Davidson, FM Sacks. [Distinct Proteomic Signatures in 16 HDL \(High-Density Lipoprotein\) Subspecies.](#) *Arteriosclerosis, Thrombosis, and Vascular Biology* 2018;38(12):2827–2842.
42. Q Zhong, B Gelaye, JW Smoller, P Avillach, **T Cai**, MA Williams. [Adverse obstetric outcomes during delivery hospitalizations complicated by suicidal behavior among US pregnant women.](#) *PLOS ONE* 2018;13(2):e0192943.
43. Q Zhong, LP Mittal, MD Nathan, KM Brown, DK González, T Cai, S Finan, B Gelaye, P Avillach, JW Smoller, EW Karlson, **T Cai**, MA Williams. [Use of natural language processing in electronic medical records to identify pregnant women with suicidal behavior: towards a solution to the complex classification problem.](#) *European Journal of Epidemiology* 2018;34(2):153–162.
44. **T Cai**, Y Zhang, Y Ho, N Link, J Sun, J Huang, TA Cai, S Damrauer, Y Ahuja, J Honerlaw, J Huang, L Costa, P Schubert, C Hong, D Gagnon, YV Sun, JM Gaziano, P Wilson, K Cho, P Tsao, CJ O'Donnell, KP and. [Association of Interleukin 6 Receptor Variant With Cardiovascular Disease Effects of Interleukin 6 Receptor Blocking Therapy.](#) *JAMA Cardiology* 2018;3(9):849–857.
45. TF Imran, D Posner, J Honerlaw, J Vassy, R Song, Y Ho, S Kittner, K Liao, **T Cai**, CJ O'Donnell, L Djousse, DR Gagnon, JM Gaziano, PW Wilson, K Cho. [A phenotyping algorithm to identify acute ischemic stroke accurately from a national biobank: the Million Veteran Program.](#) *Clinical Epidemiology* 2018;10:1509–1521.
46. TH McCoy, S Yu, KL Hart, VM Castro, HE Brown, JN Rosenquist, AE Doyle, PJ Vuijk, **T Cai**, RH Perlis. [High Throughput Phenotyping for Dimensional Psychopathology in Electronic Health Records.](#) *Biological Psychiatry* 2018;83(12):997–1004.
47. TH McCoy, VM Castro, KL Hart, AM Pellegrini, S Yu, **T Cai**, RH Perlis. [Genome-wide Association Study of Dimensional Psychopathology Using Electronic Health Records.](#) *Biological Psychiatry* 2018;83(12):1005–1011.
48. V Gamerman, **T Cai**, A Elsässer. [Pragmatic randomized clinical trials: best practices and statistical guidance.](#) *Health Services and Outcomes Research Methodology* 2018;19(1):23–35.
49. A Can, VM Castro, YH Ozdemir, S Dagen, D Dligach, S Finan, S Yu, V Gainer, NA Shadick, G Savova, S Murphy, **T Cai**, ST Weiss, R Du. [Alcohol Consumption and Aneurysmal Subarachnoid Hemorrhage.](#) *Translational Stroke Research* 2017;9(1):13–19.
50. A Can, VM Castro, YH Ozdemir, S Dagen, S Yu, D Dligach, S Finan, V Gainer, NA Shadick, S Murphy, **T Cai**, G Savova, R Dammers, ST Weiss, R Du. [Association of intracranial aneurysm rupture with smoking duration, intensity, and cessation.](#) *Neurology* 2017;89(13):1408–1415.
51. A Geva, JL Gronsbell, **T Cai**, T Cai, SN Murphy, JC Lyons, MM Heinz, MD Natter, N Patibandla, J Bickel, MP Mullen, KD Mandl, S Abman, I Adatia, ED Austin, J Feinstein, J Fineman, B Hanna, R Hopper, D Ivy, R Keller, U Krishnan, T Kulik, M Mullen, U Raj, EB Rosenzweig. [A Computable Phenotype Improves Cohort Ascertainment in a Pediatric Pulmonary Hypertension Registry.](#) *The Journal of Pediatrics*

- 2017;188:224–231.e5.
52. C Chen, M Stein, R Ursano, **T Cai**, J Gelernter, S Heeringa, S Jain, K Jensen, A Maihofer, C Mitchell, C Nievergelt, M Nock, B Neale, R Polimanti, S Ripke, M Thomas, Q Wang, E Ware, S Borja, R Kessler, J Smoller. [Genome-Wide Association Study Of Posttraumatic Stress Disorder Symptoms In Two Cohorts Of United States Army Soldiers](#). *European Neuropsychopharmacology* 2017;27:S472.
 53. K Naxerova, JG Reiter, E Brachtel, JK Lennerz, M van de Wetering, A Rowan, **T Cai**, H Clevers, C Swanton, MA Nowak, SJ Elledge, RK Jain. [Origins of lymphatic and distant metastases in human colorectal cancer](#). *Science* 2017;357(6346):55–60.
 54. KP Liao, JA Sparks, BP Hejblum, I Kuo, J Cui, LJ Lahey, A Cagan, VS Gainer, W Liu, TT Cai, J Sokolove, **T Cai**. [Phenome-Wide Association Study of Autoantibodies to Citrullinated and Noncitrullinated Epitopes in Rheumatoid Arthritis](#). *Arthritis & Rheumatology* 2017;69(4):742–749.
 55. M Koch, JD Furtado, GZ Jiang, BE Gray, **T Cai**, F Sacks, A Tjønneland, K Overvad, MK Jensen. [Associations of anthropometry and lifestyle factors with HDL subspecies according to apolipoprotein C-III](#). *Journal of Lipid Research* 2017;58(6):1196–1203.
 56. M Ong, MP Mullen, ED Austin, P Szolovits, MD Natter, A Geva, **T Cai**, SW Kong, KD Mandl. [Learning a Comorbidity-Driven Taxonomy of Pediatric Pulmonary Hypertension](#). *Circulation Research* 2017;121(4):341–353.
 57. MB Stein, EB Ware, C Mitchell, C Chen, S Borja, **T Cai**, CL Dempsey, CS Fullerton, J Gelernter, SG Heeringa, S Jain, RC Kessler, JA Naifeh, MK Nock, S Ripke, X Sun, JC Beckham, NA Kimbrel, RJ Ursano, JW and. [Genomewide association studies of suicide attempts in US soldiers](#). *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 2017;174(8):786–797.
 58. SA Aroner, M Yang, J Li, JD Furtado, FM Sacks, A Tjønneland, K Overvad, **T Cai**, MK Jensen. [Apolipoprotein C-III and High-Density Lipoprotein Subspecies Defined by Apolipoprotein C-III in Relation to Diabetes Risk](#). *American Journal of Epidemiology* 2017;186(6):736–744.
 59. W Xie, D Agniel, A Shevchenko, SV Malov, A Svitin, N Cherkasov, MK Baum, A Campa, S Gaseitsiwe, H Bussmann, J Makhema, R Marlink, V Novitsky, T Lee, **T Cai**, SJ O’Brien, M Essex. [Genome-Wide Analyses Reveal Gene Influence on HIV Disease Progression and HIV-1C Acquisition in Southern Africa](#). *AIDS Research and Human Retroviruses* 2017;33(6):597–609.
 60. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, SY Shaw, G Savova, S Churchill, EW Karlson, I Kohane, KP Liao, SN Murphy. [Comparative Effectiveness of Infliximab and Adalimumab in Crohn’s Disease and Ulcerative Colitis](#). *Inflammatory Bowel Diseases* 2016;22(4):880–885.
 61. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, SY Shaw, G Savova, S Churchill, EW Karlson, SN Murphy, KP Liao, I Kohane. [Identification of Nonresponse to Treatment Using Narrative Data in an Electronic Health Record Inflammatory Bowel Disease Cohort](#). *Inflammatory Bowel Diseases* 2016;22(1):151–158.
 62. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, SY Shaw, S Churchill, EW Karlson, SN Murphy, KP Liao, I Kohane. [Statin Use Is Associated With Reduced Risk of Colorectal Cancer in Patients With Inflammatory Bowel Diseases](#). *Clinical Gastroenterology and Hepatology* 2016;14(7):973–979.
 63. VM Castro, D Dligach, S Finan, S Yu, A Can, M Abd-El-Barr, V Gainer, NA Shadick, S Murphy, **T Cai**, G Savova, ST Weiss, R Du. [Large-scale identification of patients with cerebral aneurysms using natural language processing](#). *Neurology* 2016;88(2):164–168.
 64. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Diabetes and the risk of infections with immunomodulator therapy in inflammatory bowel diseases](#). *Alimentary Pharmacology & Therapeutics* 2015;41(11):1141–1148.
 65. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Colonoscopy Is Associated With a Reduced Risk for Colon Cancer and Mortality in Patients With Inflammatory Bowel Diseases](#). *Clinical Gastroenterology and Hepatology* 2015;13(2):322–329.e1.
 66. AN Ananthakrishnan, A Cagan, **T Cai**, VS Gainer, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao, RJ Xavier. [Common Genetic Variants Influence Circulating Vitamin D Levels in Inflammatory Bowel Diseases](#). *Inflammatory Bowel Diseases* 2015;21(11):2507–2514.
 67. H Uno, J Wittes, H Fu, SD Solomon, B Claggett, L Tian, **T Cai**, MA Pfeffer, SR Evans, L Wei. [Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies](#). *Annals of Internal Medicine* 2015;163(2):127.
 68. KP Liao, **T Cai**, GK Savova, SN Murphy, EW Karlson, AN Ananthakrishnan, VS Gainer, SY Shaw, Z Xia, P Szolovits, S Churchill, I Kohane. [Development of phenotype algorithms using electronic medical records and incorporating natural language processing](#). *BMJ* 2015;350:h1885–h1885.
 69. Q Zhong, B Gelaye, M Miller, GL Fricchione, **T Cai**, PA Johnson, DC Henderson, MA Williams. [Suicidal](#)

- behavior-related hospitalizations among pregnant women in the USA, 2006-2012. *Archives of Women's Mental Health* 2015;19(3):463-472.
70. RC Kessler, CH Warner, C Ivany, MV Petukhova, S Rose, EJ Bromet, M Brown, **T Cai**, LJ Colpe, KL Cox, CS Fullerton, SE Gilman, MJ Gruber, SG Heeringa, L Lewandowski-Romps, J Li, AM Millikan-Bell, JA Naifeh, MK Nock, AJ Rosellini, NA Sampson, M Schoenbaum, MB Stein, S Wessely, AM Zaslavsky, RJ Ursano. [Predicting Suicides After Psychiatric Hospitalization in US Army Soldiers.](#) *JAMA Psychiatry* 2015;72(1):49-57.
 71. VM Castro, J Minnier, SN Murphy, I Kohane, SE Churchill, V Gainer, **T Cai**, AG Hoffnagle, Y Dai, S Block, SR Weill, M Nadal-Vicens, AR Pollastri, JN Rosenquist, S Goryachev, D Ongur, P Sklar, RH Perlis, JW Smoller, JW Smoller, RH Perlis, PH Lee, VM Castro, AG Hoffnagle, P Sklar, EA Stahl, SM Purcell, DM Ruderfer, AW Charney, P Roussos, C Pato, M Pato, H Medeiros, J Sobel, N Craddock, I Jones, L Forty, A DiFlorio, E Green, L Jones, K Dunjewski, M Landén, C Hultman, A Juréus, S Bergen, O Svantesson, S McCarroll, J Moran, JW Smoller, K Chambert, RA and. [Validation of Electronic Health Record Phenotyping of Bipolar Disorder Cases and Controls.](#) *American Journal of Psychiatry* 2015;172(4):363-372.
 72. AN Ananthakrishnan, A Cagan, VS Gainer, S Cheng, **T Cai**, E Scoville, GG Konijeti, P Szolovits, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Thromboprophylaxis Is Associated With Reduced Post-hospitalization Venous Thromboembolic Events in Patients With Inflammatory Bowel Diseases.](#) *Clinical Gastroenterology and Hepatology* 2014;12(11):1905-1910.
 73. AN Ananthakrishnan, A Cagan, VS Gainer, S Cheng, **T Cai**, P Szolovits, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Mortality and extraintestinal cancers in patients with primary sclerosing cholangitis and inflammatory bowel disease.](#) *Journal of Crohn's and Colitis* 2014;8(9):956-963.
 74. AN Ananthakrishnan, A Cagan, VS Gainer, SC Cheng, **T Cai**, P Szolovits, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Higher plasma vitamin D is associated with reduced risk of Clostridium difficile infection in patients with inflammatory bowel diseases.](#) *Alimentary Pharmacology & Therapeutics* 2014;39(10):1136-1142.
 75. AN Ananthakrishnan, A Cheng, A Cagan, **T Cai**, VS Gainer, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Mode of Childbirth and Long-Term Outcomes in Women with Inflammatory Bowel Diseases.](#) *Digestive Diseases and Sciences* 2014;60(2):471-477.
 76. AN Ananthakrishnan, S Cheng, **T Cai**, A Cagan, VS Gainer, P Szolovits, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Serum Inflammatory Markers and Risk of Colorectal Cancer in Patients With Inflammatory Bowel Diseases.](#) *Clinical Gastroenterology and Hepatology* 2014;12(8):1342-1348.e1.
 77. AN Ananthakrishnan, S Cheng, **T Cai**, A Cagan, VS Gainer, P Szolovits, SY Shaw, S Churchill, EW Karlson, SN Murphy, I Kohane, KP Liao. [Association Between Reduced Plasma 25-Hydroxy Vitamin D and Increased Risk of Cancer in Patients With Inflammatory Bowel Diseases.](#) *Clinical Gastroenterology and Hepatology* 2014;12(5):821-827.
 78. HM van Loo, **T Cai**, MJ Gruber, J Li, P de Jonge, M Petukhova, S Rose, NA Sampson, RA Schoevers, KJ Wardenaar, MA Wilcox, AO Al-Hamzawi, LH Andrade, EJ Bromet, B Bunting, J Fayyad, SE Florescu, O Gureje, C Hu, Y Huang, D Levinson, ME Medina-Mora, Y Nakane, J Posada-Villa, KM Scott, M Xavier, Z Zarkov, RC Kessler. [Major depressive disorder subtypes to predict long-term course.](#) *Depression and Anxiety* 2014;31(9):765-777.
 79. KJ Wardenaar, HM van Loo, **T Cai**, M Fava, MJ Gruber, J Li, P de Jonge, AA Nierenberg, MV Petukhova, S Rose, NA Sampson, RA Schoevers, MA Wilcox, J Alonso, EJ Bromet, B Bunting, SE Florescu, A Fukao, O Gureje, C Hu, YQ Huang, AN Karam, D Levinson, ME Mora, J Posada-Villa, KM Scott, NI Taib, MC Viana, M Xavier, Z Zarkov, RC Kessler. [The effects of co-morbidity in defining major depression subtypes associated with long-term course and severity.](#) *Psychological Medicine* 2014;44(15):3289-3302.
 80. M Ong, IS Kohane, **T Cai**, MP Gorman, KD Mandl. [Population-Level Evidence for an Autoimmune Etiology of Epilepsy.](#) *JAMA Neurology* 2014;71(5):569-574.
 81. S Yu, KK Kumamaru, E George, RM Dunne, A Bedayat, M Neykov, AR Hunsaker, KE Dill, **T Cai**, FJ Rybicki. [Classification of CT pulmonary angiography reports by presence, chronicity, and location of pulmonary embolism with natural language processing.](#) *Journal of Biomedical Informatics* 2014;52:386-393.
 82. AN Ananthakrishnan, **T Cai**, G Savova, S Cheng, P Chen, RG Perez, VS Gainer, SN Murphy, P Szolovits, Z Xia, S Shaw, S Churchill, EW Karlson, I Kohane, RM Plenge, KP Liao. [Improving Case Definition of Crohn's Disease and Ulcerative Colitis in Electronic Medical Records Using Natural Language Processing.](#) *Inflammatory Bowel Diseases* 2013;19(7):1411-1420.
 83. AN Ananthakrishnan, VS Gainer, RG Perez, **T Cai**, SC Cheng, G Savova, P Chen, P Szolovits, Z Xia, PL De Jager, SY Shaw, S Churchill, EW Karlson, I Kohane, RH Perlis, RM Plenge, SN Murphy, KP Liao. [Psychiatric](#)

- co-morbidity is associated with increased risk of surgery in Crohn's disease. *Alimentary Pharmacology & Therapeutics* 2013;37(4):445–454.
84. AN Ananthakrishnan, VS Gainer, **T Cai**, RG Perez, S Cheng, G Savova, P Chen, P Szolovits, Z Xia, PL De Jager, S Shaw, S Churchill, EW Karlson, I Kohane, RH Perlis, RM Plenge, SN Murphy, KP Liao. [Similar Risk of Depression and Anxiety Following Surgery or Hospitalization for Crohn's Disease and Ulcerative Colitis](#). *American Journal of Gastroenterology* 2013;108(4):594–601.
 85. GM Altschuler, O Hofmann, I Kalatskaya, R Payne, SJ Sui, U Saxena, AV Krivtsov, SA Armstrong, **T Cai**, L Stein, WA Hide. [Pathprinting: An integrative approach to understand the functional basis of disease](#). *Genome Medicine* 2013;5(7):68.
 86. KP Liao, D Diogo, J Cui, **T Cai**, Y Okada, VS Gainer, SN Murphy, N Gupta, D Mirel, AN Ananthakrishnan, P Szolovits, SY Shaw, S Raychaudhuri, S Churchill, I Kohane, EW Karlson, RM Plenge. [Association between low density lipoprotein and rheumatoid arthritis genetic factors with low density lipoprotein levels in rheumatoid arthritis and non-rheumatoid arthritis controls](#). *Annals of the Rheumatic Diseases* 2013;73(6):1170–1175.
 87. KP Liao, F Kurreeman, G Li, G Duclos, S Murphy, R Guzman, **T Cai**, N Gupta, V Gainer, P Schur, J Cui, JC Denny, P Szolovits, S Churchill, I Kohane, EW Karlson, RM Plenge. [Associations of autoantibodies, autoimmune risk alleles, and clinical diagnoses from the electronic medical records in rheumatoid arthritis cases and non-rheumatoid arthritis controls](#). *Arthritis & Rheumatism* 2013;65(3):571–581.
 88. KP Liao, **T Cai**, VS Gainer, A Cagan, SN Murphy, C Liu, S Churchill, SY Shaw, I Kohane, DH Solomon, RM Plenge, EW Karlson. [Lipid and Lipoprotein Levels and Trend in Rheumatoid Arthritis Compared to the General Population](#). *Arthritis Care & Research* 2013;65(12):2046–2050.
 89. PS Lai, JQ Hang, FY Zhang, X Lin, BY Zheng, HL Dai, L Su, **T Cai**, DC Christiani. [Gender differences in the effect of occupational endotoxin exposure on impaired lung function and death: the Shanghai Textile Worker Study](#). *Occupational and Environmental Medicine* 2013;71(2):118–125.
 90. RT Gandhi, .. , RJ Bosch, E Aga, MA Bedison, B Bastow, JL Schmitz, JD Siliciano, RF Siliciano, JJ Eron, JW Mellors. [Residual plasma viremia and infectious HIV-1 recovery from resting memory CD4 cells in patients on antiretroviral therapy: results from ACTG A5173](#). *Antiviral Therapy* 2013.
 91. Z Xia, E Secor, LB Chibnik, RM Bove, S Cheng, T Chitnis, A Cagan, VS Gainer, PJ Chen, KP Liao, SY Shaw, AN Ananthakrishnan, P Szolovits, HL Weiner, EW Karlson, SN Murphy, GK Savova, **T Cai**, SE Churchill, RM Plenge, IS Kohane, PL De Jager. [Modeling Disease Severity in Multiple Sclerosis Using Electronic Health Records](#). *PLoS ONE* 2013;8(11):e78927.
 92. AN Ananthakrishnan, R Guzman-Perez, V Gainer, **T Cai**, S Churchill, I Kohane, RM Plenge, S Murphy. [Predictors of severe outcomes associated with Clostridium difficile infection in patients with inflammatory bowel disease](#). *Alimentary Pharmacology & Therapeutics* 2012;35(7):789–795.
 93. MT Lu, S Demehri, **T Cai**, L Parast, AR Hunsaker, SZ Goldhaber, FJ Rybicki. [Axial and Reformatted Four-Chamber Right Ventricle-to-Left Ventricle Diameter Ratios on Pulmonary CT Angiography as Predictors of Death After Acute Pulmonary Embolism](#). *American Journal of Roentgenology* 2012;198(6):1353–1360.
 94. RJ Carroll, WK Thompson, AE Eyler, AM Mandelin, **T Cai**, RM Zink, JA Pacheco, CS Boomershine, TA Lasko, H Xu, EW Karlson, RG Perez, VS Gainer, SN Murphy, EM Ruderman, RM Pope, RM Plenge, AN Kho, KP Liao, JC Denny. [Portability of an algorithm to identify rheumatoid arthritis in electronic health records](#). *Journal of the American Medical Informatics Association* 2012;19(e1):e162–e169.
 95. S Wassertheil-Smoller, A McGinn, M Allison, **T Cai**, D Curb, C Eaton, S Hendrix, R Kaplan, M Ko, LW Martin, X Xue. [Improvement in Stroke Risk Prediction: Role of C-Reactive Protein and Lipoprotein-Associated Phospholipase A2 in the Women's Health Initiative](#). *International Journal of Stroke* 2012;9(7):902–909.
 96. F Kurreeman, K Liao, L Chibnik, B Hickey, E Stahl, V Gainer, G Li, L Bry, S Mahan, K Ardlie, B Thomson, P Szolovits, S Churchill, SN Murphy, **T Cai**, S Raychaudhuri, I Kohane, E Karlson, RM Plenge. [Genetic Basis of Autoantibody Positive and Negative Rheumatoid Arthritis Risk in a Multi-ethnic Cohort Derived from Electronic Health Records](#). *The American Journal of Human Genetics* 2011;88(1):57–69.
 97. L Qi, L Parast, **T Cai**, C Powers, EV Gervino, TH Hauser, FB Hu, A Doria. [Genetic Susceptibility to Coronary Heart Disease in Type 2 Diabetes](#). *Journal of the American College of Cardiology* 2011;58(25):2675–2682.
 98. RH Perlis, DV Iosifescu, VM Castro, SN Murphy, VS Gainer, J Minnier, **T Cai**, S Goryachev, Q Zeng, PJ Gallagher, M Fava, JB Weilburg, SE Churchill, IS Kohane, JW Smoller. [Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model](#). *Psychological Medicine* 2011;42(1):41–50.
 99. TJ Love, **T Cai**, EW Karlson. [Validation of Psoriatic Arthritis Diagnoses in Electronic Medical Records Using Natural Language Processing](#). *Seminars in Arthritis and Rheumatism* 2011;40(5):413–420.
 100. AR Hunsaker, IB Oliva, **T Cai**, B Trotman-Dickenson, RR Gill, H Hatabu, FJ Rybicki. [Contrast Opacification](#)

- Using a Reduced Volume of Iodinated Contrast Material and Low Peak Kilovoltage in Pulmonary CT Angiography: Objective and Subjective Evaluation. *American Journal of Roentgenology* 2010;195(2):W118–W124.
101. C Wu, A Sodickson, **T Cai**, NA Levit, ML Steigner, FJ Rybicki, SM Ledbetter. [Comparison of Respiratory Motion Artifact From Craniocaudal Versus Caudocranial Scanning With 64-MDCT Pulmonary Angiography.](#) *American Journal of Roentgenology* 2010;195(1):155–159.
 102. FJ Rybicki, D Mitsouras, CD Owens, A Whitmore, M Gerhard-Herman, N Wake, **T Cai**, Q Zhou, MS Conte, MA Creager, RV Mulkern. [Multi-contrast high spatial resolution black blood inner volume three-dimensional fast spin echo MR imaging in peripheral vein bypass grafts.](#) *The International Journal of Cardiovascular Imaging* 2010;26(6):683–691.
 103. KP Liao, **T Cai**, V Gainer, S Goryachev, Q Zeng-treitler, S Raychaudhuri, P Szolovits, S Churchill, S Murphy, I Kohane, EW Karlson, RM Plenge. [Electronic medical records for discovery research in rheumatoid arthritis.](#) *Arthritis Care & Research* 2010;62(8):1120–1127.
 104. ML Steigner, D Mitsouras, AG Whitmore, HJ Otero, C Wang, O Buckley, NA Levit, AZ Hussain, **T Cai**, RT Mather, O Smedby, MF DiCarli, FJ Rybicki. [Iodinated Contrast Opacification Gradients in Normal Coronary Arteries Imaged With Prospectively ECG-Gated Single Heart Beat 320-Detector Row Computed Tomography.](#) *Circulation: Cardiovascular Imaging* 2010;3(2):179–186.
 105. X Gao, CQ Lai, T Scott, J Shen, **T Cai**, JM Ordovas, KL Tucker. [Urinary 8-Hydroxy-2-deoxyguanosine and Cognitive Function in Puerto Rican Adults.](#) *American Journal of Epidemiology* 2010;172(3):271–278.
 106. C Wu, **T Cai**, K Rikova, D Merberg, S Kasif, M Steffen. [A Predictive Phosphorylation Signature of Lung Cancer.](#) *PLoS ONE* 2009;4(11):e7994.
 107. MC Cornelis, L Qi, C Zhang, P Kraft, J Manson, **T Cai**, DJ Hunter, FB Hu. [Joint Effects of Common Genetic Variants on the Risk for Type 2 Diabetes in U.S. Men and Women of European Ancestry.](#) *Annals of Internal Medicine* 2009;150(8):541–550.
 108. FJ Rybicki, HJ Otero, ML Steigner, G Vorobiof, L Nallamshetty, D Mitsouras, H Ersoy, RT Mather, PF Judy, **T Cai**, K Coyner, K Schultz, AG Whitmore, MF Di Carli. [Initial evaluation of coronary images from 320-detector row computed tomography.](#) *The International Journal of Cardiovascular Imaging* 2008;24(5):535–546.
 109. K Chien, **T Cai**, H Hsu, T Su, W Chang, M Chen, Y Lee, FB Hu. [A prediction model for type 2 diabetes risk among Chinese people.](#) *Diabetologia* 2008;52(3):443–450.
 110. ML Steigner, HJ Otero, **T Cai**, D Mitsouras, L Nallamshetty, AG Whitmore, H Ersoy, NA Levit, MF Di Carli, FJ Rybicki. [Narrowing the phase window width in prospectively ECG-gated single heart beat 320-detector row coronary CT angiography.](#) *The International Journal of Cardiovascular Imaging* 2008;25(1):85–90.
 111. MT Lu, **T Cai**, H Ersoy, AG Whitmore, NA Levit, SZ Goldhaber, FJ Rybicki. [Comparison of ECG-gated versus non-gated CT ventricular measurements in thirty patients with acute pulmonary embolism.](#) *The International Journal of Cardiovascular Imaging* 2008;25(1):101–107.
 112. MT Lu, **T Cai**, H Ersoy, AG Whitmore, R Quiroz, SZ Goldhaber, FJ Rybicki. [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;246(1):281–287.
 113. H Ersoy, SZ Goldhaber, **T Cai**, T Luu, J Rosebrook, R Mulkern, F Rybicki. [Time-Resolved MR Angiography: A Primary Screening Examination of Patients with Suspected Pulmonary Embolism and Contraindications to Administration of Iodinated Contrast Material.](#) *American Journal of Roentgenology* 2007;188(5):1246–1254.
 114. JD MacKenzie, J Nazario-Larrieu, **T Cai**, MS Ledbetter, MA Duran-Mendicuti, PF Judy, FJ Rybicki. [Reduced-Dose CT: Effect on Reader Evaluation in Detection of Pulmonary Embolism.](#) *American Journal of Roentgenology* 2007;189(6):1371–1379.
 115. SK Chuang, **T Cai**. [Predicting Clustered Dental Implant Survival Using Frailty Methods.](#) *Journal of Dental Research* 2006;85(12):1147–1151.
 116. SK Chuang, **T Cai**, CW Douglass, LJ Wei, TB Dodson. [Frailty Approach for the Analysis of Clustered Failure Time Observations in Dental Research.](#) *Journal of Dental Research* 2005;84(1):54–58.
 117. JJ Rogus, **T Cai**, L Wei. [Issues in genomic screening: Critical values, sample sizes, and the ability to detect linkage.](#) *Genetic Epidemiology* 1999;17(S1):S697–S701.

Recent Invited Talks

- 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 Annual Statistical Conference (Nov 2012)
- Division of Biostatistics, Albert Einstein 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, Bangkok 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, Germany (Jan 2010)
- 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)