February 5, 2017

Curriculum Vitae

Yong Chen

Department of Biostatistics and Epidemiology
Center for Clinical Epidemiology and Biostatistics
Perelman School of Medicine
University of Pennsylvania
210 Blockley Hall, 423 Guardian Drive
Philadelphia, PA 19104-6021 USA

Office: 215-746-8155

Email: ychen123@upenn.edu

Professional Appointments

- Assistant Professor and Senior Scholar, Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania Perelman School of Medicine, September 2015 - present.
- Senior Scholar, Center for Evidence-based Practice, Perelman Center for Advanced Medicine, University of Pennsylvania, April 2016 - present.
- Assistant Professor, Division of Biostatistics, University of Texas School of Public Health, September 2010 - August 2015.

Education

- Ph.D. in Biostatistics, June 2010.
  Department of Biostatistics, The Johns Hopkins University, Baltimore, MD.
  Thesis Title: Pseudolikelihood methods: theory and its applications in genetic epidemiology.
  Advisors: Professors Kung-Yee Liang and Charles Rohde
- M.A. in Mathematics, June 2005.
  Department of Mathematics, The Johns Hopkins University, Baltimore, MD.
- B.Sc. in Mathematics, June 2003.
  University of Science and Technology of China, Hefei, Anhui, P.R. China.

Awards/Honors

- April, 2015. *Institute of Mathematical Statistics IMS Travel Award*
- May, 2010. *Margaret Merrell Award for excellence in research.*
  (recognizes outstanding research of a doctoral student at graduation)
  Department of Biostatistics, The Johns Hopkins University.
- 2005 – 2010. *Sommer Scholar.*
  `http://www.jhsph.edu/admissions/scholarships/institutional-scholarships/sommer-scholars/scholars/?scholar=43`
  Johns Hopkins School of Public Health.

Research Interests

I was trained as a mathematical statistician on statistical inference and the foundation of statistics by working with Professors Kung-Yee Liang and Charles Rohde. In the past, I have been working independently and collaboratively with my colleagues to advance rigorous statistical theory, and to develop statistical and computational methods that are truly motivated from small data (e.g., summary data in systematic reviews) and big data (e.g., healthcare data, self-reported vaccine/drug safety data, self-reported weight loss data). I am also interested in evidence-based medicine because I believe it is an area that needs deep thinking of the fundamental question "what is evidence in the data?", which is critical toward improved medical decision-making for patients and stakeholders, and is sometimes blurred by the use of 'standard' statistical analyses.

In general, I am interested in advancing general theory in all disciplines that can be broadly applicable, as well as developing tailored statistical models to specific applications by working with domain experts.

Areas that I am working on:

- Methods for large healthcare data (e.g., electronic health record data).
- Precision medicine using heterogeneous massive data.
- Comparative effectiveness research. Patient centered outcome research.
- Statistical inference.
- Philosophy of statistics. Statistical evidence.

Research Grants

• Active as Principal Investigator

1. *Dynamic learning for post-vaccine event prediction using temporal information in VAERS*
   Role: Principal Investigator (with Cui Tao at University of Texas)
   Funding Agency: NIAID (1R01AI130460)
   Period: 02/01/2017 - 01/31/2022
   Total cost: $3,122,444
2. *Statistical Methods and Software for Multivariate Meta-analysis*
   Role: Subcontract Principal Investigator (PI: Haitao Chu at University of Minnesota)
   Funding Agency: NLM (R21-LM012197)
   Period: 10/01/2015 - 07/31/2017
   Total cost: $421,086; Subcontract total cost: $180,000
3. Penn CTER Pilot Study Award on *Integrative Analysis in Large Distributed Research Networks*
   Role: Principal Investigator
   Funding Agency: Penn Center for Therapeutic Effectiveness Research (CTER)
   Period: 07/2016 - 06/2017
   Total cost: $20,000

4. Penn Undergraduate Research Mentoring Program Award on *Dynamic Traffic Network Data Stream Modeling and Simulation*
   Role: Principal Investigator
   Mentee: Karen Chi
   Funding Agency: Penn Undergraduate Research Mentoring Program
   Period: 06/2016 - 12/2016
   Total cost: $5,000

- **Impact score within fundable range**

1. *Innovative Methods Adjusting for Publication Bias in Multivariate and Network Meta-Analyses*
   Role: Principal Investigator (with Jing Ning at MD Anderson Cancer Center)
   Funding Agency: NLM
   Period: 04/01/2017 - 03/31/2019
   Total budget: $446,712; Impact score: 32 (first submission)

- **Active as Co-Investigator**

1. *Statistical methods for phenotype estimation and analysis using electronic health records data*
   Role: Co-Investigator
   Project PI: Rebecca Hubbard
   Funding Agency: Patient-Centered Outcomes Research Institute (PCORI)
   Period: 10/31/2016 - 10/30/2019
   Total cost: $1,059,241

2. *Biomedical Computing and Informatics Strategies for Infectious Disease Research*
   Role: Co-Investigator
   Project PI: Jason Moore
   Funding Agency: NIAID (R01-AI116794)
   Period: 03/01/2016 - 02/28/2021
   Total cost: $571,791 (annual cost in 2016)

3. *Statistical methods for evidence-based practice*
   Role: Lead Statistician and Senior Scholar
   Project PI: Craig Umscheid
   Funding Agency: Penn Medicine
   Period: 04/15/2016 - present (10% effort)

- **Pending**

1. *Likelihood Methods for Bias Correction in EHR Based Association Studies*
   Role: Principal Investigator
   Funding Agency: NIH
   Total budget: $2,078,018

2. *A general framework to account for outcome reporting bias in systematic reviews*
   Role: Principal Investigator
   Funding Agency: NIH
   Total budget: $1,583,915

3. *A Meta-Analytical Framework to Accelerate Comparative Effectiveness Research Conducted within Clinical Data Research Networks*
   Role: Principal Investigator
   Funding Agency: PCORI
   Total budget: $1,049,982

4. *Mathematical and Statistical Modeling of Traffic Flow based on Sensor Data*
   Role: Principal Investigator (Multiple PI grant with Jingmei Qiu at University of Houston and Benedetto Piccoli at Rutgers University)
   Total budget: $520,615

5. *Statistical Methods for Analysis of Electronic Health Record Data*
   Role: co-Investigator
   Project PI: Jinbo Chen
   Funding Agency: NIH
   Total budget: $1,100,141

6. *Race and Access to Knee Replacement: Intervention Effectiveness-Implementation Hybrid*
   Role: co-Director of Biostatistics and Data Management Core
   Project PI: Said Ibrahim
   Funding Agency: NIH
   Total budget: $3,500,000

7. *Transforming Mental Health Delivery Through Behavioral Economics and Implementation Science*
   Role: co-Investigator
   Funding Agency: NIH
   Project PI: Mandell, D/ Beidas, R

- **Completed**

1. *Multivariate meta-analysis of diagnostic tests*
   Role: Principal Investigator

Funding Agency: Agency for Healthcare Research and Quality (1R03HS022900)
Period: 04/01/2014 - 09/31/2016
Total cost: $97,000

2. *Novel statistical approaches in pharmacogenetic studies*
Role: Principal Investigator
Funding Agency: University of Texas School of Public Health (PRIME award)
Period: 09/01/2012 - 08/31/2013
Total cost: $25,000

Publications in Peer-reviewed Journals

underline: manuscript first-authored by a student advised/co-advised by Dr. Chen.
♣: corresponding author. †: equally contributed. ¶: senior statistician.

Statistical inference

1. Hong, C, Ning, Y, Wei, P, Cao, Y and **Chen, Y♣**. (2016) A semiparametric model for vQTL mapping, **Biometrics** (in press).

2. Hong, C, **Chen, Y♣**, Ning, Y, Wang, S, Wu, H, and Carroll, RJ. (2016) PLEMT: A novel pseudolikelihood based EM test for homogeneity in generalized exponential tilt mixture models, **Journal of the American Statistical Association** (in press)
(This paper won 2015 JSM Biometrics section Byar Awards)

3. **Chen, Y♣**, Huang, J, Ning, Y, Liang, K-Y and Lindsay, B. (2016) A conditional test for composite likelihood with boundary constraints. **Biometrika** (provisional acceptance).

4. **Chen, Y♣**, Ning, J, Ning, Y, Liang, K-Y and Bandeen-Roche, K. (2016) On the pseudolikelihood inference for semiparametric models with boundary problems. **Biometrika** (in press).

5. Ning, J, **Chen, Y**, Cai, C, Huang, X and Wang, MC. (2015) On the Dependence Structure of Bivariate Recurrent Event Processes: Inference and Estimation, **Biometrika** 102(2): 345-358.

6. **Chen, Y♣**, Ning, J and Cai, C. (2015) Regression analysis of longitudinal data with irregular and informative observation times, **Biostatistics**, 16(4): 727-739.

7. Ning, Y and **Chen, Y**. (2015) A class of pseudolikelihood ratio tests for homogeneity in exponential tilt mixture models, **Scandinavian Journal of Statistics** 42 (2), 504–517.

8. Nie, L†, **Chen, Y†** and Chu, H. (2011) Asymptotic Variances of Maximum Likelihood Estimator for the Correlation Coefficient from a BVN Distribution with One Variable Subject to Censoring, **Journal of Statistical Planning and Inference**, 141 (1), 392–401.

9. **Chen, Y♣** and Liang, KY. (2010) On the asymptotic behaviour of the pseudolikelihood ratio test statistic with boundary problems, **Biometrika**, 97 (3), 603–620.

Methods and Applications on Large Healthcare Data

10. Cai, Y, Du, J, Huang, J, Tao, C, and **Chen, Y**. (2016) Signal Detection Method for Temporal Variation of Adverse Effect with VAERS database. **International Conference on Intelligent Biology and Medicine,** (accepted).

11. Duan, R, Cao, M, Wu, Y, Huang, J, Denny, J, Xu, H and **Chen, Y**. (2016) An Empirical Study for Impacts of Measurement Errors on EHR based Association Studies, **AMIA annual symposium proceedings,** (accepted).
(This paper won the first prize of "Best of Student Papers in Knowledge Discovery and Data Mining (KDDM)"Awards)

12. Du, J, Cai, Y, **Chen, Y♣** and Tao, C. (2016) Trivalent influenza vaccine adverse symptoms analysis based on MedDRA terminology using VAERS data in 2011, **Journal of Biomedical Semantics**, 7–13.

13. Du, J, Cai, Y, **Chen, Y♣**, He, Y, and Tao, C. (2016) Analysis of Individual Differences in Vaccine Pharmacovigilance using VAERS Data and MedDRA System Organ Classes: A Use Case Study with Trivalent Influenza Vaccine, **Vaccine**, 2016.

14. Tao, C, Du, J, Cai, Y and **Chen, Y**. (2015) Trivalent Influenza Vaccine Adverse Event Analysis Based On MedDRA System Organ Classes Using VAERS Data, **Studies in health technology and informatics**, 2015. 216:1076.

15. Cao M, **Chen, Y**, Zhu M, Zhang J. (2015) Automated Evaluation of Medical Software Usage: Algorithm and Statistical Analyses, **Studies in health technology and informatics**, 2015. 216:965.

Comparative Effectiveness Research

16. Ning, J†, **Chen, Y†** and Piao, J (2016+) Maximum likelihood estimation and EM algorithm of Copas selection model for publication bias correction. **Biostatistics,** (in press).

17. Liu, Y, DeSantis, S and **Chen, Y** (2016+) Bayesian network meta-analysis of clinical trials with correlated outcomes subject to publication bias: application to a systematic review of alcohol dependence, **Journal of the Royal Statistical Society: Series C,** (in press).

18. Li, X, **Chen, Y**, and Li, R (2016) A frailty model for recurrent events during alternating restraint and non-restraint time periods, **Statistics in Medicine,** 20 February 2017.

19. **Chen, Y♣**, Liu, Y, Chu, H, Lee, M and Schmid, C (2016) A simple and robust method for multivariate meta-analysis of diagnostic test accuracy, **Statistics in Medicine,** (in press).

20. **Chen, Y♣**, Hong, C, Ning, Y and Su, X. (2016) Meta-analysis of studies with bivariate binary outcomes: a marginal beta-binomial model approach, **Statistics in Medicine,** (in press).

21. Chahoud, J, Semaan, A, **Chen, Y¶**, Cao, M, Rieber, A, Rady, P and Tyring, S. (2016) The Association between Beta-genus Human Papillomavirus and Cutaneous Squamous Cell Carcinoma in Immunocompetent Individuals: a Meta-analysis, **JAMA Dermatology,** (in press).

22. **Chen, Y♣**, Cai, Y, Hong, C, and Jackson, D. (2016) Inference for correlated effect sizes using multiple univariate meta-analyses, **Statistics in Medicine,** 35(9): 1405-1422.

23. Liu, Y†, **Chen, Y♣†** and Chu H. (2015) A unification of models for meta-analysis of diagnostic accuracy studies without a gold standard, **Biometrics,** 71(2):538–47.

24. **Chen, Y♣**, Hong, C and Riley, R. (2015) An alternative pseudolikelihood method for multivariate random-effects meta-analysis, **Statistics in Medicine** 34 (3): 361-380.

25. **Chen, Y♣**, Liu, Y, Ning, J, Cormier J and Chu H. (2015) A model for combining case-control and cohort studies in systematic reviews of diagnostic tests, **Journal of the Royal Statistical Society: Series C,** 64(3): 469-489.

26. **Chen, Y♣**, Liu, Y, Ning, J, Nie, L, Zhu, H and Chu H. (2014) A composite likelihood method for bivariate analysis of sensitivity and specificity in diagnostic reviews, **Statistical Methods in Medical Research.**

27. **Chen, Y♣**, Chu, H, Luo, S, Nie L and Chen S. (2014) Bayesian analysis on meta-analysis of case-control studies accounting for within-study correlation, **Statistical Methods in Medical Research.**

28. **Chen, Y♣**, Luo, S, Chu, H, Su, X and Nie, L. (2014) An Empirical Bayes Method for Multivariate Meta-analysis with Application in Clinical Trials, **Communications in Statistics - Theory and Methods,** 43(16), 3536–3551.

29. Ma, X, **Chen, Y**, Cole, S and Chu, H. (2014) A hybrid Bayesian hierarchical model combining cohort and case-control studies for meta-analysis of diagnostic tests: accounting for partial verification bias, **Statistical Methods in Medical Research.**

30. Luo, S, **Chen, Y**, Su, X and Chu, H. (2014) mmeta: An R package for multivariate meta-analysis. **Journal of Statistical Software,** 56 (11).

31. **Chen, Y♣**, Luo, S, Chu, H and Wei, P. (2013) Bayesian inference on risk differences: an application to multivariate meta-analysis of adverse events in clinical trials, **Statistics in Biopharmaceutical Research,** 5 (2): 142-155.

32. Nie, L, Soon, G, Qi, K, **Chen, Y** and Chu, HT. (2013) A note on partial covariate-adjustment and design considerations in noninferiority trials when patient-level data are not available, **Journal of Biopharmaceutical Statistics,** 23 (5), 1042–1053.

33. Chu, H, Nie, L, **Chen, Y**, Huang, Y and Sun, W. (2012) Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk, **Statistical Methods in Medical Research,** 21 (6): 621–633.

Statistical Genetics and Bioinformatics

34. Sun, H, Wang, Y, **Chen, Y**, Lu, Y and Wang, S. (2016+) pETM: a penalized Exponential Tilt Model for analysis of correlated high-dimensional DNA methylation data. **Bioinformatics,** (in press).

35. Liu, Y, **Chen, Y♣** and Scheet, P♣ (2016), A Meta-Analytic Framework for Detection of Genetic Interactions, **Genetic Epidemiology,** 40 (7), 534 –543.

36. **Chen, Y♣**, Ning, Y, Hong, C and Wang, S. (2013) Semiparametric tests for identifying differentially methylated loci with case-control designs using Illumina arrays, **Genetic Epidemiology,** 38 (1), 42–50.

37. Tong P, **Chen, Y**, Su X and Coombes K (2013) SIBER: Systematic Identication of Bimodally Expressed Genes Using RNAseq Data. **Bioinformatics,** 29 (5), 605–613.

38. Huang, J, **Chen, Y♣**, Swartz, M and Ionita-Laza, I (2014) Family-Based Association Test for Sequence Data with Applications in the GAW 18 Simulated data, **BMC proceedings 8 (Suppl 1)**: S27, 2014.

Letters

39. **Chen, Y♣** and Luo, S. (2011) A few remarks on "Statistical distribution of the difference of two proportions", **Statistics in Medicine,** 30 (15), 1913–1915.

Book reviews

40. **Chen, Y♣** and Li, T. (2015) Book review on "Applied Meta-analysis for Social Science Research", **The American Statistician**.

Book chapters

41. Ma, X, Liu, YL, **Chen, Y**, Chu, HT. (2015) Meta-analysis of diagnostic tests in handbook of meta-analysis. Editors: Schmid, C, Stijnen,T and White, I. Handbooks of Modern Statistical Methods series, Chapman and Hall.

Manuscripts under revision (selected)

42. **Chen, Y**$^{\clubsuit}$, Liang, K-Y, Tong, P, Beaty, T, Barnes, K. and Kao, L (2016+) A pseudolikelihood approach for assessing genetic association in case-control studies with unmeasured population structure, **Statistical Methods in Medical Research,** (provisional acceptance).

43. Hong, C, Riley, R and **Chen, Y** (2016+) Robust variance estimator for Riley method of the multivariate meta-analysis when within-study correlations are unknown, **Research Synthesis Methods,** (minor revision).

44. Ma, X, Chu, H, Ibrahim, J, and **Chen, Y** (2016+) A Bayesian hierarchical model for network meta-analysis of diagnostic tests, **Biostatistics,** (under revision)

Collaborative Papers

45. Lu, G, Changb, J, Liu, Z, Chen, Y, Li, M, and Zhu, J. (2016) Phospholipase C Beta 1: A Candidate Signature Gene for Proneural Subtype High-Grade Glioma, **Molecular Neurobiology** 53:6511-6525.

46. Liu, L, Xu, H, Wang, W, Wu, C, **Chen, Y**, Yang, J, Cen, P, Xu, J, Liu, C, Long, J, Guha, S, Fu, D, Ni, Q, Jatoi, A, Chari, S, McCleary-Wheeler, A, Fernandez-Zapico, M, Li, M, Yu, X, (2015) A Preoperative Serum Signature of CEA+/CA125+/CA19-9 $\geq$ 1,000 U/mL Indicates Poor Outcome to Pancreatectomy for Pancreatic Cancer, **International Journal of Cancer** (9), 2216-2227.

47. Xu, C, Wallace, M, Yang, J, Jiang, L, Zhai, Q, Zhang, Y, Hong, C, **Chen, Y**, Frank, T, Stauffer, J, Asbun, H, Raimondo, M, Woodward, T, Li, Z, Guha, S, Zheng, L and Li, M (2014) ZIP4 is a Novel Diagnostic and Prognostic Marker in Human Pancreatic Cancer: A Systemic Comparison Between EUS-FNA and Surgical Specimens, **Current Molecular Medicine**, 14, 1–7.

48. Lin, Y†, **Chen, Y**†, Wang, Y†, Yang, J†, Zhu, V, Liu, Y, Cui, X, Yan, W, Jiang, T, Hergenroeder, G, Fletcher, S, Levine, J, Kim, D, Tandon, N, Zhu, J and Li, M, (2013) ZIP4 is a novel molecular marker for glioma, **Neuro-Oncology**, 15 (8), 1008–1016. PMID: 21191590.

49. Zhang Y, Yang, J, Cui, X, **Chen, Y**¶, Zhu, V, Hagan, J, Wang, H, Yu, X, Hodges, S, Fang, J, Chiao, P, Logsdon, C, Fisher, W, Brunicardi, F, Chen, C, Yao, Q, Fernandez-Zapico, M, and Li, M, (2013), A Novel epigenetic CREB-miR 373 Axis Mediates ZIP4-Induced Pancreatic Cancer Growth, **EMBO Molecular Medicine**, 5 (9), 1322–1334.

50. Lin Y, Zhang G, Zhang J, Gao G, Li M, **Chen, Y**, Wang J, Li G, Song S, Wang Y and Jiang T, (2013) A panel of four cytokines predicts the prognosis of patients with malignant gliomas, **Journal of Neuro-Oncology**, 114 (2), 199–208.

51. Go, VF, Frangakis, C, Minh, N, Latkin, C, Ha, T, Mo, T, Sripaipan, T, Davis, W, Zelaya, C, **Chen, Y**, Celentano, D and Quan, V, (2013) Effects of an HIV peer prevention intervention on sexual and injecting risk behaviors among injecting drug users and their risk partners in Thai Nguyen, Vietnam: a randomized controlled trial, **Social Science & Medicine**, 96 154–164.

52. Nestadt, G,Wang, Y, Grados, MA, Riddle, MA, Greenberg, BD, Knowles, JA, Fyer, AJ, McCracken, JT, Rauch, SL, Murphy, DL, Rasmussen, SA, Cullen, B, Piacentini, J, Geller, D, Pauls, D, Bienvenu, OJ, **Chen, Y**, Liang, KY, Goes, FS, Maher, B, Pulver, AE, Shugart, YY, Valle, D, Samuels, JF and Chang, YC (2011) Homeobox genes in obsessive-compulsive disorder, **American Journal of Medical Genetics Part B: Neuropsychiatric Genetics**, 159 (1), 53–60.

53. Blom, RM, Samuels, JF, Grados, MA, **Chen, Y**, Bienvenu, OJ, Riddle, MA, Liang, KY, Brandt, J and Nestadt, G (2011), Cognitive functioning in compulsive hoarding, **Journal of Anxiety Disorders**, 25 (8), 1139–1144.

54. Frangakis, CE, Geschwind J, Kim, D, **Chen, Y**, Koteish, A, Hong, K, Liapi, E, Georgiades, CS. (2010). Chemoembolization decreases drop-off risk of hepatocellular carcinoma patients on the liver transplant list **Cardiovascular and interventional radiology**, 34 (6), 1254–1261.

55. Wible, BC, Rilling, WS, Drescher, P, Hieb, RA, Saeian, K, Frangakis, CE, **Chen, Y**, Eastwood, D., Kim, HS. (2010). Longitudinal Quality of Life Assessment of Patients with Hepatocellular Carcinoma After Primary Transarterial Chemoembolization, **Journal of Vascular and Interventional Radiology**, 21(7), 1024–1030.

56. Buijs, M, Vossen, JA, Frangakis, CE, Hong, K, Georgiades, C, **Chen, Y**, Liapi, E, and Geschwind, JF. (2008). Nonresectable Hepatocellular Carcinoma: Long-term Toxicity in Patients Treated with Transarterial Chemoembolization – Single-Center Experience. **Radiology**, 249, 346–354.

Statistical software

♣*: main author of the software.*

1. R package *'xmeta'*: a comprehensive collection of functions for multivariate meta-analyses of continuous or binary outcomes. This package also includes functions and visualization tools for detection and correction of publication bias in multivariate meta-analysis.
   `https://cran.rstudio.com/web/packages/xmeta/index.html`
   Role: designer, co-author and maintainer (with Chuan Hong♣, Haitao Chu and Yulun Liu)

2. R package *'robustETM'*: Testing homogeneity for generalized exponential tilt model. This package includes a collection of functions for (1) implementing methods for

testing homogeneity for generalized exponential tilt model; and (2) implementing existing methods under comparison.
`https://cran.r-project.org/web/packages/robustETM/index.html`
Role: co-designer and co-author (with Chuan Hong♣, Yang Ning and Hao Wu)

3. R package *'mmeta'*: This package implements bivariate meta-analysis for binary outcomes.
`https://cran.rstudio.com/web/packages/mmeta/index.html`
Role: co-designer and co-author (with Xiao Su♣, Sheng Luo and Haitao Chu)

4. R codes for YETI in "*A meta-analytic framework for detection of genetic interactions*" *Genetic Epidemiology*.
`http://www.yulunliu.com/software.html`
Role: co-author (with Yulun Liu♣ and Paul Scheet)

Students and Postdoc Fellows Advising

- Postdoc Fellow: Dr. Jing Huang (May 2016 - present)
Research topic: Statistical and computational methods for analyzing electronic health records data.

- Postdoc Fellow: Dr. Yulun Liu (July 2016 - present)
Research topic: Integrative analysis of heterogeneous and massive datasets.

- Ph.D. Students Advising or Co-Advising
  - Ph.D. Candidate: *Rui Duan* (sole advisor)
  Thesis title: TBA
  - Ph.D. Candidate: *Le Wang* (co-advisor: Dr. Jinbo Chen)
  Thesis title: TBA
  - Ph.D. Candidate: *Ming Cao* (co-advisor: Dr. Kayo Fujimoto)
  Thesis title: Statistical inference in networks under hierarchical exponential random graph models and its longitudinal extension

- Ph.D. Students Advised (primary advisor)
  - *Dr. Chuan Hong*
  Thesis title: Statistical tests for homogeneity in parametric and semiparametric models
  (graduated in 2016; now Postdoc fellow at Harvard working with Professor Tianxi Cai)
    * **2015 Institute of Mathematical Statistics IMS Travel Award**
    * **2015 JSM Biometrics section Byar Awards**
    * **2015 ENAR Distinguished Student Paper Awards**
    * **2014 SRCOS Boyd Harshbarger Travel Award**

    * Girardi/Keese Scholarship in Biostatistics at UTSPH
  - *Dr. Yulun Liu* (co-advisor: Dr. Paul Scheet)
    Thesis title: Meta-analysis and its application in biomedical studies
    (graduated in 2016; now Postdoc fellow at University of Pennsylvania)
    * **2014 SRCOS Boyd Harshbarger Travel Award**
    * **Dean's Student Travel Awards at UTSPH**
  - *Dr. Yi Cai* (co-advisor: Dr. Cui Tao)
    Thesis title: Statistical methods for signal detection of adverse effects using VAERS data
    (graduated in 2016; now data scientist at Pieces Technologies Inc.)

- Ph.D. Students co-advised
  - *Dr. Xiaoqi Li* (with Dr. Ruosha Li)
    Thesis title: Statistical models for recurrent events during alternating restraint and non-restraint periods
    (graduated in 2016; now Senior Research Scientist at Eli Lilly and Company)

- Ph.D. Thesis committee chaired
  - Dr. Xueming Gu
    (graduated in 2011; now statistician at Eli Lilly and Company)
  - Dr. Suyu Liu
    (graduated in 2012; now tenure-track assistant professor at MD Anderson Cancer Center)
  - Dr. Jun Liu
    (graduated in 2012; now statistician at MD Anderson Cancer Center)

- Ph.D. Thesis committee served
  - Xiao Su (2011-2012); Wei Qiao (2011-2013)

- Master Thesis Committee served
  - Cindy Ma (graduated in 2012), Yan Shi (graduated in 2012)

Professional Activities

- Editorial Activities
  - Associate Editor for *International Journal of Biostatistics*.
    January 2013 - present
  - Statistical Advisory Board for *PLOS ONE* .
    April 2015 - present

- Scientific Merit Reviewer

- *Patient-Centered Outcomes Research Institute*(PCORI): Large Pragmatic Trial Panel. March 2015 - present

- Reviewer of
  Journal of the American Statistical Association (theory and methods; case studies and applications), Biometrika, Statistical Science, Biometrics, Statistica Sinica, Biostatistics, Journals of the Royal Statistical Society: Series C, Statistics in Medicine, Genetic Epidemiology, Psychometrika, Biometrical journal, PLOS ONE, Computational Statistics & Data Analysis, Scientific Reports, Epidemiological Methods, Pharmacoepidemiology and Drug Safety, Epidemiologic Perspectives and Innovations

- Section organizer of
  - *Eastern North American Regional Meeting of International Biometric Society (ENAR) (Invited Session; March 12-15, 2017, Washington, DC).*
  - *International Chinese Statistical Association (Invited Session; June 12-15, 2016; Atlanta, GA).*
  - *ENAR (Invited Session; March 6-9, 2016; Austin, TX).*
  - *ENAR (Invited Session; March 16-19, 2014; Baltimore, MD).*
  - *Southern Regional Council on Statistics*
    (Invited Session; June 5-8, 2011; McCormick, SC).

- Short Course and Guest Lecture
  - 2016
    - (Guest Lecture): Methylation data and mixture models. Penn Biostatistics Big data working group. October 21, 2016.
    - (Course instructor for short course; co-teach with Dr. Haitao Chu): Statistical methods and software for multivariate meta-analysis, 2016 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop, September 28-30, 2016, Washington DC. (half-day course)
    - (Guest Lecture): The Development of Meta-Analysis, University of Oxford, July 4, 2016, Oxford, UK.
      The 3rd Course on Network Meta-Analysis
    - (Course instructor for short course; co-teach with Dr. Haitao Chu): Statistical Methods and Software for Multivariate Meta-Analysis, June 12, 2016, Atlanta, GA. (half-day course)
    - (Guest Lecture): Longitudinal data under biased sampling, BSTA 656 on longitudinal data analysis, Department of Biostatistics and Epidemiology, University of Pennsylvania, April 12, 2016, Philadelphia, PA
  - 2010
    - (Guest Lecture) A pseudolikelihood approach for population substructure, 2010, SLAM working group, Department of Biostatistics, Johns Hopkins School of Public Health, Baltimore, MD.

    - (Guest Lecture) A lecture on Pseudolikelihood, Asymptotic properties of likelihood ratio tests and pseudolikelihood ratio tests under nonstandard condition, Graduate Course on Advanced Statistical Theory 2010, Department of Biostatistics, Johns Hopkins University, Baltimore, MD.

- Invited talk
  - 2016
    - Bias reduction methods in self-reported longitudinal data, Pennsylvania State University, September 22, 2016, State College, PA.
    - Publication Bias Problem in Multivariate Meta-Analysis, SRSM, July 13, 2016, Florence, Italy.
    - Multivariate Meta-Analysis: estimation and bias correction, MRC Biostatistics Unit, University of Cambridge, July 8, 2016, Cambridge, UK.
    - Bias Correction in Multivariate and Network Meta-Analysis, Department of Psychiatry, University of Oxford, July 7, 2016, Oxford, UK.
    - Misclassification and measurement error in EHR based association studies, ICSA, June 15, 2016, Atlanta, GA.
    - On the asymptotic behavior of the (pseudo)likelihood ratio test statistic under nonstandard conditions, Department of Mathematics, University of Maryland at College Park, April 7, 2016, College Park, MD.
    - Detecting and correcting Publication Bias in Multivariate Meta-Analysis, March 9, 2016, ENAR, Austin, TX
  - 2015
    - Analysis of Complex Data Under Biased Sampling, CCEB, University of Pennsylvania, October 22, 2015, Philadelphia, PA
    - Analysis of longitudinal data under biased sampling, SLAM working group, Department of Biostatistics, Johns Hopkins School of Public Health, September 4, 2015, Baltimore, MD
    - A simple test for detecting publication bias in multivariate meta-analysis, August 9, 2015, Joint Statistical Meetings, Seattle, WA
    - Analysis of longitudinal data under biased sampling, March 23, 2015, Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA.
    - Analysis of longitudinal data under biased sampling, March 9, 2015, Department of Biostatistics and Computational Biology, University of Rochester, Rochester, NY.
    - Analysis of longitudinal data under biased sampling, February 24, 2015, Department of Statistics, Michigan State University, East Lansing, MI.
    - Analysis of longitudinal data under biased sampling, February 11, 2015, Department of Epidemiology and Biostatistics, Indiana University, Bloomington, IN.
    - Analysis of longitudinal data under biased sampling, February 3, 2015, Department of Biostatistics, Brown University, Providence, RI.

* Analysis of longitudinal data under biased sampling, January 19, 2015, Biostatistics and Biomathematics, Fred Hutchinson Cancer Research Center, Seattle, WA.
* Analysis of longitudinal data under biased sampling, January 15, 2015, Department of Biostatistics, University of Pittsburg, Pittsburg, PA.

– 2014

* Analysis of longitudinal data under biased sampling, September 19, 2014, University of Texas School of Public Health, Houston, TX
* Regression analysis of longitudinal data with irregular and informative observation times, August 2-7, 2014, Boston, MA.
* On the pseudolikelihood inference for semiparametric models with boundary problems, June 15-18, 2014, Honolulu, HI.
* Bayesian inference on risk differences: an application to multivariate meta-analysis of adverse events in clinical trials, May, 2014, Muncie, IN.
* Bias reduction methods for longitudinal data analysis, Department of Biostatistics, University of Kentucky, February 8, 2014, Lexington, KY.
* Bias reduction methods for longitudinal data analysis, Department of Epidemiology and Biostatistics, University of Maryland at College Park, January 31, 2014, College Park, MD.
* Novel statistical tests for homogeneity in semiparametric mixture models with application to methylation data, January 21, 2014, University of Minnesota, MN.

– 2013

* Regression analysis of longitudinal data with irregular and informative observation times, October 2, 2013, University of Minnesota, MN.
* Regression analysis of longitudinal data with irregular and informative observation times, June 19, 2013, Xiamen University, Xiamen, China.
* Regression analysis of longitudinal data with irregular and informative observation times, June 25, 2013, Southwestern University of Finance and Economics, Chengdu, China.

– 2012

* Regression analysis of longitudinal data with irregular and informative observation times, November 30, 2012, University of Texas School of Public Health, Houston, TX.

– 2011

* Assessing Genetic Association in Case-Control Studies with Unmeasured Population Substructure, June 7, 2011, Southern Regional Council on Statistics (SRCOS) Summer Research Conference, McCormick, SC

– 2010

* Assessing genetic association in case-control studies with unmeasured population substructure, 2010, Center for Devices and Radiological Health, Food and Drug Administration, Silver Spring, MD.

    - Assessing genetic association in case-control studies with unmeasured population substructure, 2010, The University of Texas School of Public Health, Houston, TX.
    - Assessing genetic association in case-control studies with unmeasured population substructure, 2010, University of Michigan, Ann Arbor, MI.
    - Assessing genetic association in case-control studies with unmeasured population substructure, 2010, The University of Texas MD Anderson Cancer Center, Houston, TX.

- Contributed talk (2007 - 2013)
  - On the asymptotic behavior of the pseudolikelihood ratio test statistic with boundary problems in semiparametric models, 2013, Joint Statistical Meetings, Montreal, Canada.
  - Regression analysis of longitudinal data with irregular and informative observation times, 2012, Joint Statistical Meetings, San Diego, CA.
  - On the asymptotic behavior of the pseudolikelihood ratio test statistic with boundary problems, 2011, Joint Statistical Meetings, Miami, FL.
  - Assessing genetic association in case-control studies with unmeasured population substructure, 2010, American Public Health Association Annual Meeting, Denver, Colorado, 80211.
  - Assessing genetic association in case-control studies with unmeasured population substructure, 2010, ENAR, New Orleans, LA.
  - Assessing genetic association in case-control studies with unmeasured population substructure, 2009, ENAR, Washington DC.
  - Exact Bayesian Inference in 2×2 Tables, 2008, Joint Statistical Meetings, Denver, CO.
  - Exact Bayesian inference in matched pair designs, 2008, ENAR, Arlington, VA.
  - A Likelihood Approach to Sample Size Calculation, 2007, ENAR, Atlanta, GA.

Teaching

- Courses Taught in University of Pennsylvania and Student Evaluations

Time	Catalog Number	Course Title	Number of Resp/Stud	Average/Full Ratings
Fall 2016	BSTA-622	Inference II	10/11	3.4/4
Fall 2015	EPID-621	Longitudinal and clustered data (all students are medical doctors)	5/7	4.0/5

The last two columns summarize 'number of responses vs. number of students' and 'the average rating vs. the full (highest possible) ratings' of teaching evaluations for

classes that I have taught at UPenn. The average rating corresponds to 'Overall rating of instructor effectiveness of instructors for this course'. Here 5 is the highest possible rating, while 1 is the lowest possible rating.

- Courses Taught in University of Texas and Student Evaluations

Time	Catalog Number	Course Title	Number of Resp/Stud	Average/Full Ratings
Fall 2014	PH1916	Generalized Linear Models	9/9	5.00/5
Spring 2014	PH1918	Methods for Correlated Data	16/20	4.69/5
Fall 2013	PH1916	Generalized Linear Models	10/13	4.80/5
Spring 2013	PH1690	Foundations of Biostatistics	20/48	4.00/5
Fall 2012	PH1916	Generalized Linear Models	9/12	4.67/5
Spring 2012	PH1918	Methods for Correlated Data	13/18	4.69/5
Spring 2012	PH1999	Longitudinal Data Analysis	13/18	4.69/5
Spring 2012	PH1916	Generalized Linear Models	12/17	4.50/5

The last two columns summarize 'number of responses vs. number of students' and 'the average rating vs. the full (highest possible) ratings' of teaching evaluations for classes that I have taught at UT. *The average rating corresponds to 'Overall rating of instructor effectiveness of instructors for this course'. Here 5 is the highest possible rating, while 1 is the lowest possible rating.*

- Courses Taught in Department of Mathematics, the Johns Hopkins University

Time	Course Title	Number of Students
Fall 2005	Qualify Exams in Mathematics	9

Service at University of Pennsylvania

- Biostatistics Faculty Recruitment Committee, September 2016 - present
- Graduate Program Committees, February 2016 - present
- Seminar Series Committee, February 2016 - present
- Biostatistics Junior Faculty Committee, February 2016 - present

Memberships

- American Statistical Association
- American Medical Informatics Association
- International Biometric Society
- American Society of Human Genetics
- International Chinese Statistical Association
- American Mathematical Society
- American Public Health Association
- Global Health Council