Jiayi (Jessie) Tong

423 Guardian Drive, Philadelphia, PA 19104
Jiayi.Tong@pennmedicine.upenn.edu; (858)-249-8559

Education & Professional Experiences

- ***Ph.D. in Biostatistics, Sep 2019 - present***
  Department of Biostatistics, Epidemiology and Informatics
  Perelman School of Medicine at the University of Pennsylvania
  Advisor: Dr. Yong Chen
- ***M.S. in Biostatistics, Sep 2019 - June 2021***
  Department of Biostatistics, Epidemiology and Informatics
  Perelman School of Medicine at the University of Pennsylvania
  Advisors: Dr. Yong Chen, Dr. Adam Cuker
- ***Research Assistant, Oct 2017 - May 2019***
  Department of Biostatistics, Epidemiology and Informatics
  Perelman School of Medicine at the University of Pennsylvania
  Advisor: Dr. Yong Chen
- ***B.S. with High Honors in Applied Mathematics, Sep 2014 - June 2017***
  Cum Laude
  University of California, San Diego
  Thesis title: Strongly Correlated Band Random Matrices
  Thesis paper: Random Matrices with Blocks of Intermediate Scale: Strongly Correlated Band Random Matrices
  Advisor: Dr. Todd Kemp

Awards & Honors

- ***ASA Philadelphia Chapter Exceptional Achievement Award for Graduate Students (2022)***
  American Statistical Association (ASA) Phidelphia Chapter
- ***Award for Outstanding Data Science Communication and Innovation (2021)***
  OptumLabs UnitedHealth Group, Minnetonka
- ***Department Honors in Mathematics (Applied) with High Distinction (2017)***
  `https://math.ucsd.edu/_files/undergraduate/honors-program/honors-program-presentations/2016-2017/Jiayi_Tong_Honors_Thesis.pdf`
  University of California, San Diego
- ***Honor society: Phi Beta Kappa (2017)***
  University of California, San Diego
- ***Provost's Honors (2015, 2016, 2017)***
  University of California, San Diego

Publications

Published

1. Yin, Z., **Tong, J.**, Chen, Y., Hubbard, R.A. and Tang, C.Y., (2021). A cost-effective chart review sampling design to account for phenotyping error in electronic health records (EHR) data. *Journal of the American Medical Informatics Association.* 29(1), pp.52-61

2. Luo, C., Jiang, Y., Du, J., **Tong, J.**, Huang, J., Lo Re III, V., Ellenberg, S.S., Poland, G.A., Tao, C. and Chen, Y., (2021). Prediction of post-vaccination Guillain-Barré syndrome using data from a passive surveillance system. *Pharmacoepidemiology and Drug Safety* (30(5), pp.602-609).

3. Fan, R., Zhang, Y., Xu, Y., **Tong, J.,** Chen, Z., Gu, M., Fan, W., Chen, Y., Peng, F., Jiang, Y., (2021) Serum antinuclear antibodies associate with worse prognosis in AQP4-positive neuromyelitis optica spectrum disorder. *Brain and Behavior*, p.e01865.

4. **Tong, J.**, Chen, Z., Duan, R., Lo-Ciganic, W., Lyu, T., Tao, C., Merkel, P., Kranzler, H., Bian, J., Chen, Y.. (2020) Identifying Clinical Risk Factors of Opioid Use Disorder using a Distributed Algorithm to Combine Real-World Data from a Large Clinical Data Research Network. *AMIA Annu Symp Proc.* (Vol. 2020, p. 1220). American Medical Informatics Association.

5. **Tong, J.**, Duan, R., Li, R., Scheuemie, M. J., Moore, J. H., and Chen, Y. (2020). Robust-ODAL: Learning from heterogeneous health systems without sharing patient-level data. *In Pacific Symposium on Biocomputing* (Vol. 25, pp. 695-706).

6. Paydary, K., Banwell, E., **Tong, J.**, Chen, Y. and Cuker, A., (2020). Diagnostic accuracy of the PLASMIC score in patients with suspected thrombotic thrombocytopenic purpura: A systematic review and meta-analysis. *Transfusion*, 60(9), pp.2047-2057.

7. Duan, R., Chen Z., **Tong, J.**, Luo, C., Lyu, T., Tao, C., Maraganore, D., Bian, J. and Chen, Y.. (2020) Leverage real-world longitudinal data in large clinical research networks for Alzheimer's disease and related dementia. *AMIA Annu Symp Proc.* (Vol. 2020, p. 393). American Medical Informatics Association.

8. Duan, R., Luo, C., Schuemie, M.J., **Tong, J.**, Liang, J., Boland, M.R., Bian, J., Xu, H., Berlin, J.A., Moore, J.H., Mahoney, K.B. and Chen, Y., (2020). Learning from local to global - an efficient distributed algorithm for modeling time to event data. *Journal of the American Medical Informatics Association*, 27(7), pp.1028-1036.

9. Hubbard, R. A., **Tong, J.**, Duan, R., Chen, Y. (2020). Reducing Bias Due to Outcome Misclassification for Epidemiologic Studies Using EHR-derived Probabilistic Phenotypes. *Epidemiology*, 31(4), pp.542-550.

10. **Tong, J.**, Huang, J., Chubak, J., Wang, X., Hubbard, R., and Chen, Y. (2019) An Augmented Estimation Procedure for EHR-based Association Studies Accounting for Differential Misclassification, *Journal of the American Medical Informatics Association.* 27(2), pp.244-253.

11. Li, R., **Tong, J.**, Duan, R., Chen, Y. and Moore, J.H., (2019). Evaluation of phenotyping errors on polygenic risk score predictions. *International Conference on Bioinformatics Models, Methods and Algorithms* (in press).

12. Liang, J., Liu, J., Fan, R., Chen, Z., Chen, X., **Tong, J.**, Chen, Y., Peng, F. and Jiang, Y., (2019). Plasma Homocysteine Level Is Associated with the Expanded Disability Status Scale in Neuromyelitis Optica Spectrum Disorder. *Neuroimmunomodulation*, 26(5), pp.258-264.

13. Chiasakul, T., Jesus, E., **Tong, J.**, Chen, Y., Crowther, M., Garcia, D., Chai-Adisaksopha, C., Messe, S. and Cuker, A. (2019) Inherited Thrombophilia and the Risk of Arterial Ischemic Stroke: A Systematic Review and Meta-Analysis. *Journal of the American Heart Association*, 8(19), e012877.

14. Huang, J., Zhang, X., **Tong, J.**, Du, J., Duan, R., Yang, L., Moore, J.H., Tao, C. and Chen, Y., (2019). Comparing drug safety of hepatitis C therapies using post-market data. *BMC medical informatics and decision making*, 19(4), p.147.

15. Luo, C., Zhang, X., Yang, L., Du, J., **Tong, J.**, Huang, J., Tao, C. and Chen, Y., (2019), August. Rare adverse effects analysis of new generation hepatitis C medications using EMR and FAERS data.*Pharmacoepidemiology and Drug Safety* (Vol. 28, pp. 580-581). 111 RIVER ST, HOBOKEN 07030-5774, NJ USA: WILEY.

16. **Tong, J.**, Huang, J., Du, J., Cai, Y., Tao, C. and Chen, Y. (2018) The Use of Likelihood Ratio Test to Identify Rare Adverse Events with Year-varying Reporting Rates for FLU4 Vaccine in VAERS, *AMIA Annual Symposium Proceedings*. American Medical Informatics Association. (Vol. 2018, p. 1544)

17. Huang, J., Zhang, X., **Tong, J.**, Du, J., Duan, R., Yang, L., Moore, J.H., Chen, Y. and Tao, C. (2018). Comparing adverse effects of Hepatitis C drugs using FAERS data. *In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1653-1656). IEEE.

Under review/Under revision/Preprints

18. **Tong, J.**, Luo, C., Islam, M.N., Sheils, N., Buresh, J., Edmondson, M., Merkel, P.A., Lautenbach, E., Duan, R. and Chen, Y. (2021) An efficient distributed algorithm with application to COVID-19 data from heterogeneous clinical sites. *medRxiv*. (Under revision by NPJ Digital Medicine)

19. **Tong, J.**, Hubbard, R.A., Saine, M.E., Xu, H., Zuo, X., Kimmel, S.E., Cuker, A. and Chen, Y., (2021). Advancing timely and reliable evidence synthesis in the era of COVID-19: A novel method for including preprints in systematic reviews. (Under review by Research Synthesis Methods).

20. **Tong, J.**, Yang, N., Islam, N., Qu, A. and Chen, Y. (2021) On the proportional likelihood ratio model for sparse data. (Under review by Biometrics).

21. **Tong, J.**, Duan, R., Luo, C., Moore, J., Shaw, P., Zhu, J. and Chen, Y. (2021) Quantifying and correcting bias due to outcome dependent self-reported weights in longitudinal study of weight loss interventions, (Under review by JAMIA).

22. **Tong, J.**, Islam, N., Buresh, J., Chen, Y. and Sheils, N.(2021) Which one to report? Odds ratio or Relative risk, (Under review by American Journal of Epidemiology).

23. Liang, W., **Tong, J.**, Cappola, A. , Carroll, R., Chu, H., and Chen., Y. (2021) On Robust Inference for Reference Ranges, (Under review by Biometrika)

24. Chen, Y., **Tong, J.**, Hubbard, R.A., Saine, M.E., Xu, H., Zuo, X., Kimmel, S.E., Cuker, A. (2021). A New Approach for Timely Systematic Reviews of Rapidly Evolving Research: Identifying Evidence in Which We Can Place Our Confidence. *medRxiv.*

25. Duan, R,. **Tong, J.**, Asch, D., Sutton, A., Chu, H., Schmid, C., Chen, Y. (2021) Origami plot: a novel tool for visualizing multivariate data, (Under review by BMJ)

26. Duan, R., **Tong, J.**, Lin, L., Levine, L.D., Sammel, M.D., Stoddard, J., Li, T., Schmid, C.H., Chu, H. and Chen, Y. (2021). PALM: Patient-centered Treatment Ranking via Large-scale Multivariate Network Meta-analysis. *medRxiv.* (Under review by The Annals of Applied Statistics)

27. Li, W., **Tong, J.**, Anjum, M., Mohammed, N., Chen, Y. and Jiang, X., (2021). Federated Learning Algorithms for Generalized Mixed-effects Model (GLMM) on Horizontally Partitioned Data from Distributed Sources. *arXiv preprint arXiv:2109.14046.*

28. Zhang, M., **Tong, J.**, Ma, W., Luo, C., Liu, H., Jiang, Y., Qin, L., Wang, X., Zhang, J., Peng, F., Chen, Y., Li, W., Jiang, Y.(2021) Predictors of pulmonary adenocarcinoma with leptomeningeal metastases: a 2021 target-therapy assisted molGPA model – a 2021 target-therapy assisted molGPA model for LM (Under review by Annals of Oncology)

29. Wang, Y., Ye, Z., **Tong, J.**, Tang. C., Chen., Y. (2021) Analysis of Time to Event Data Under Two-Layer Censoring. (Under review by Biometrika)

30. Duan, R., Piao, J., Marks-Anglin, A., **Tong, J.**, Lin, L., Chu, H., Ning, J., Chen, Y. (2021) Testing for publication bias in meta-analysis under Copas selection model. *arXiv preprint* arXiv:2007.00836. (Under review by Journal of the Royal Statistical Society: Series A)

31. Luo, C., Islam, M.N., Sheils, N.E., Reps, J.M., Buresh, J., Duan, R., **Tong, J.**, Edmondson, M., Schuemie, M.J. and Chen, Y. (2021) Lossless Distributed Linear Mixed Model with Application to Integration of Heterogeneous Healthcare Data. *medRxiv.* (under revision by Nature Communications)

Oral Presentations

1. **Tong, J.**, Hubbard, R., Saine, E., Xu, H., Zuo, X., Kimmel, S., Umscheid, S., Cuker, A. and Chen, Y. (2021) Advancing timely and reliable evidence synthesis for COVID-19 research by including preprints in systematic reviews **The 34th New England Statistics Symposium (NESS)**, **invited session**, virtual October 2021.

2. **Tong, J.**, Hubbard, R., Saine, E., Xu, H., Zuo, X., Kimmel, S., Umscheid, S., Cuker, A. and Chen, Y. (2021) Advancing timely and reliable evidence synthesis for COVID-19 research by including preprints in systematic reviews **2021 Joint Statistical Meetings (JSM)**, virtual August 2021.

3. **Tong, J.**, Luo, C., Islam, M.N., Sheils, N., Buresh, J., Edmondson, M., Merkel, P.A., Lautenbach, E., Duan, R. and Chen, Y. (2021) Distributed Binary Regression of Electronic Health Records Data Across Heterogeneous Clinical Sites **Eastern North American Region (ENAR)**, virtual March 2021.

4. **Tong, J.**, Chen, Z., Duan, R., Lo-Ciganic, W., Lyu, T., Tao, C., Merkel, P., Kranzler, H., Bian, J., Chen, Y.. (2020) Identifying Clinical Risk Factors of Opioid Use Disorder using a Distributed Algorithm to Combine Real-World Data from a Large Clinical Data Research Network, **AMIA 2020 Annual Symposium**, virtual Nov 2020.

5. Duan, R., Chen Z., **Tong, J.**, Luo, C., Lyu, T., Tao, C., Maraganore, D., Bian, J. and Chen, Y.. (2020) Leverage real-world longitudinal data in large clinical research networks for Alzheimer's disease and related dementia, **AMIA 2020 Annual Symposium**, virtual Nov 2020.

6. **Tong, J.**, Ren, B., Liu, Y., Moore, J., Xu, H. and Chen, Y. (2020) Xmeta-COVID19: A Comprehensive Web-based Toolbox for Meta-analysis on COVID19 Research, **2020 OHDSI**, virtual Oct 2020.

7. **Tong, J.**, Luo, C., Duan, R., and Chen, Y. (2020) Distributed Learning From EHR across Multiple Heterogenous Clinical Sites, **2020 Joint Statistical Meetings (JSM)**, virtual July 2020.

8. **Tong, J.**, Huang, J., Chubak, J., Wang, X., Hubbard, R., and Chen, Y. (2020) An Augmented Estimation Procedure for EHR-based Association Studies, **Eastern North American Region (ENAR)**, virtual March 2020.

9. **Tong, J.**, Duan, R., Li, R., Scheuemie, M. J., Moore, J. H., and Chen, Y. (2020). Robust-ODAL: Learning from heterogeneous health systems without sharing patient-level data. **Pacific Symposium on Biocomputing**, Hawaii Jan 2020

10. **Tong, J.**, Li, R., Zhou, D., Duan, R., Moore, J., and Chen, Y. (2019) Improving the reproducibility of EHR-based association studies for pleiotropic effects by accounting for phenotyping errors, **Eastern North American Region (ENAR)**, Philadelphia Oct 2019.

11. **Tong, J.**, Huang, J., Du, J., Cai, Y., Tao, C. and Chen, Y. (2018) The Use of Likelihood Ratio Test to Identify Rare Adverse Events with Year-varying Reporting Rates for FLU4 Vaccine in VAERS, **AMIA 2018 Annual Symposium**, San Francisco Nov 2018.

Poster Presentations

12. **Tong, J.**, Duan, R., Liu, Y., Hong, C., Chi, K., and Chen, Y. (2018) A Comprehensive Toolbox for Advanced Meta-analysis, **2018 DBEI & CCEB Research Day**, University of Pennsylvania, Philadelphia April 2018.

Research/Job Experiences

1. ***Data Scientist and Research Graduate Intern***
   OptumLabs Unitedhealth Group *Jun 2021 - Aug 2021*
   Advisors: Dr. Natalie Sheils, Dr. Nazmul Islam *Minnetonka, MN*

   - Research on two commonly used relative measures in biomedical and epidemiologic research studies: odds ratio and relative risk.

- Investigated the difference between two measures through simulation study with a wide spectrum of settings and use cases with real-world claims data from UnitedHealth Group database.

2. ***Research Assistant***
   The Perelman School of Medicine, University of Pennsylvania *Oct 2017 - Present*
   Advisor: Dr. Yong Chen *Philadelphia, PA*

   - Statistical methodology development: worked with collaborators and developed an augmented method to handle misclassified phenotype in EHR data to improve reproducibility in EHR-based findings; applied the augmented method to multiple datasets.
   - Applied projects: applied a pairwise likelihood model to Keep it Off study, a three-arm randomized controlled trial study to analyze weight loss data that could be potentially missing not at random; applied semiparametric proportional likelihood ratio model to analyze the length of hospitalization with data from Penn ICU.
   - Collaborative projects: collaborated concurrently on multiple projects, including post-marketing pharmacovigilance study (e.g., vaccine safety) with medical informaticians from UTHealth and meta-analysis with physicians from Penn Medicine.
   - Web and module development: built and added interactive online meta-analysis module to *XMETA* (`https://www.xmeta.org/`), which is aimed to assist the doctors without prior programming experience to conduct meta-analyses.

3. ***Independent Undergraduate Research*** *Jun 2017 - present*
   Department of Mathematics, UC San Diego *La Jolla, CA*

   - Research on "Forecasting Stock Prices by Fuzzy Support Vector Machines".
   - Used time-series model analysis in R to study the trending of stock prices; selected technical indicators with machine learning methods for stock prices prediction.

4. ***Research Assistant***
   Department of Mathematics, UC San Diego *Sep 2016 - Jun 2017*
   Advisor: Dr. Todd Kmep *La Jolla, CA*

   - Designed "Maximally Correlated Band Random Matrices"; performed simulations using different distributed entries and investigated the pattern of the eigenvalues.
   - Proposed and proved eigenvalue pattern with Wick's Theorem and the method of matrix moment.
   - Wrote thesis paper "Random Matrices with Blocks of Intermediate Scale Strongly Correlated Band Matrices", which was awarded High Distinction Honors.

5. ***Summer Research Assistant***
   Department of Mathematics, UC San Diego *Jun 2016 - Sep 2016*
   Advisor: Dr. Todd Kemp *La Jolla, CA*

   - Worked on random matrices with blocks whose entries are highly correlated; performed simulations and investigated the pattern of random matrices' eigenvalues.
   - Tested and computed scalings of entries; performed normalization of random matrices; gave proof of the scalings which contribute to the convergence; led the weekly discussion and developed random matrices with new structure and distribution of correlated entries.

6. ***Research Assistant***
   Department of Political Science, UC San Diego — *Sep 2015 - Apr 2016*
   Advisor: Dr. Erik Gartzke — *La Jolla, CA*

   - Prepared and cleaned data of U.S. territories and possessions and foreign oversea area of the 1960s to the 1980s for research about Democratic Peace.

Teaching Experience

1. ***Teaching Assistant*** — *Nov 2021 - Present*
   BSTA 754: Advanced Survival Analysis — *Philadelphia, PA*

2. ***Teaching Assistant*** — *Sep 2021 - Nov 2021*
   BSTA 656: Longitudinal Data Analysis — *Philadelphia, PA*

3. ***Teaching Assistant*** — *Nov 2020 - Dec 2020*
   BSTA 660: Design of Observational Studies — *Philadelphia, PA*

4. ***Teaching Assistant*** — *Sep 2020 - Nov 2020*
   BSTA 661: Design Of Interventional Studies — *Philadelphia, PA*

Professional Activities & Services

1. ***Reviewer of Pacific Symposium on Biocomputing (PSB) 2022***

2. ***Reviewer of American Medical Informatics Association (AMIA) Annual Symposium, 2019, 2020, 2021***

3. ***4th Annual Summer Institute in Statistics for Big Data (SISBID)***

Computing Skills

- R, Matlab, Python, Java, SQL