Yichen Jia

Email: YIJ22@pitt.edu
Immigration status: U.S. Lawful Permanent Resident

Education

University of Pittsburgh, Pittsburgh, PA   Anticipated 2022
Ph.D. in Biostatistics
Dissertation advisor: Jong H. Jeong, Ph.D.

University of Washington, Seattle, WA   Aug 2017
M.S. in Biostatistics
Thesis advisor: Timothy A. Thornton, Ph.D.

University of Iowa, Iowa City, IA   May 2015
Graduate with Distinction
B.S. in Mathematics
B.A. in Biochemistry

Work Experience

Sanofi Pasteur, Swiftwater, PA; Cambridge, MA; Remote

Biostatistician L2   Aug 2021 – present
- Focus on omics data related projects

Biostatistician I (contractor)   Sep 2020 – Aug 2021
- Prepared SAP for clinical trials with transcriptomics data
- Analyzed transcriptomics data (microarray, RNA-seq)

Biostatistics Summer Intern   Jun 2020 – Aug 2020
- Developed an R package for data analysis and power analysis using exact method for vaccine efficacy trial
- Classification of morphological phenotype using deep learning methods

University of Pittsburgh Medical Center, Department of Psychiatry, Pittsburgh, PA

Graduate Student Researcher   Aug 2017- Aug 2021
  PI: Mary Ganguli, M.D., MPH; Supervisor: (Joyce) Chung-Chou H. Chang, Ph.D.
- Provided statistical analysis and consultation for studying risk/protective factors of mild cognitive impairment/dementia and related entities

University of Pittsburgh, Pittsburgh, PA

Teaching Assistant   Jan 2019 – Apr 2019
- BIOST 2086, Applied Mixed Model Analysis

University of Washington, Department of Biostatistics, Seattle, WA

Graduate Research Assistant   Sep 2016 – Jun 2017
  Supervisor: Timothy A. Thornton, Ph.D.

- Analyzed high-dimensional genomic dataset (human liver whole genome RNA-seq data) using machine learning tools

University of Iowa, Department of Obstetrics and Gynecology, Iowa City, IA

Undergraduate Research Assistant   Dec 2011-May 2015
      P.I.: Kimberly K. Leslie, M.D.; Supervisor: Shujie Yang, Ph.D.
- Examined the mechanism of progesterone receptor (PR) down-regulation during endometrial cancer progression

Honors and Awards

ASA Student of the Year (Honorable mention), Pittsburgh Chapter   2021

ASA Lifetime Data Science (LiDS) Section Student Paper Award at JSM   2021

    Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg

ICSA Applied Statistics Symposium Poster Award   2020

    Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg

ENAR Distinguished Student Paper Award   2020

    Cause-specific Quantile Regression on Inactivity Time

University of Washington School of Public Health Awards of Excellence   2017

    Outstanding Master's Student

University of Iowa Student Employee of the Year Recognition   2014

Publications

Biostatistics – Methodology

- **Jia, Y.** & Jeong, J. H. (2021). Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg. *Computational Statistics and Data Analysis,* 107323. link

- **Jia, Y.** & Jeong, J. H. (2021). Cause-specific Quantile Regression on Inactivity Time. *Statistics in Medicine, 40*(7), 1811-1824.

Biostatistics – Collaboration

- Lee. S., **Jia, Y.**, Snitz, B. E., Chang, C. C. H. & Ganguli, M. (2021+). Should Social Cognition Be Assessed in Older Adults? A Population-Based Study. *Submitted.*

- Runk A., **Jia, Y.**, Liu, A., Chang, C. C. H., Ganguli, M & Snitz, B. E. (2021+). Associations between visual acuity and cognitive decline in older adulthood: A 9-year longitudinal study. *Under revision.*

- Bhojak T., **Jia, Y.**, Jacobson, E., Snitz, B. E., Chang, C. C. H. & Ganguli, M. (2021). Driving Habits of older adults: A Population-Based Study. *Alzheimer Disease & Associated Disorders, in press.*

- Ganguli, M., Hughes, T. F., **Jia, Y.**, Lingler, J., Jacobson, E., & Chang, C. C. H. (2020). Aging and Functional Health Literacy: A Population-Based Study. *The American Journal of*

Geriatric Psychiatry, in press. Epub 2020 Dec 11.

- Lee. S., Jacobson, E., **Jia, Y.**, Snitz, B. E., Chang, C. C. H. & Ganguli, M. (2020). Reading the Mind in the Eyes: A Population-Based Study of Social Cognition in Older Adults. *The American Journal of Geriatric Psychiatry, in press.* Epub 2020 Dec 01.

- **Jia, Y.**, Chang, C. C. H., Hughes, T. F., Wang, S., & Ganguli, M. (2020). Predictors of Dementia in the Oldest Old: A Novel Machine Learning Approach. *Alzheimer Disease & Associated Disorders, 34*(4), 325-332.

- Cohen, A. D., **Jia, Y.**, Smagula, S. F., Chang, C. C. H., Snitz, B. E., Jacobson, E., & Ganguli, M. (2020). Cognitive functions predict trajectories of sleepiness over ten year: a population-based study. *The Journals of Gerontology: Series A*, glaa120.

- Smagula, S. F., **Jia, Y.,** Chang, C. C. H., Cohen, A., & Ganguli, M. (2019). Trajectories of daytime sleepiness and their associations with dementia incidence. *Journal of Sleep Research, 29*(6): e12952.

- Stoehr, G. P., Jacobson, E., **Jia, Y.**, Snitz, B. E., & Ganguli, M. (2019). Trends in the use of medication to treat or prevent dementia: a population-based study. *Alzheimer Disease & Associated Disorders, 34*(2):148-155.

- Shaaban, C. E., **Jia, Y.**, Chang, C. C. H., & Ganguli, M. (2019). Independent and joint effects of vascular and cardiometabolic risk factor pairs for risk of all-cause dementia: a prospective population-based study. *International psychogeriatrics, 31*(10), 1421-1432.

- Ganguli, M., **Jia, Y.**, Hughes, T. F., Snitz, B. E., Chang, C. C. H., Berman, S. B., ... & Kamboh, M. I. (2019). Mild Cognitive Impairment that Does Not Progress to Dementia: A Population-Based Study. *Journal of the American Geriatrics Society, 67*(2), 232-238.

Biochemistry and Cell Biology

- Kavlashvili, T., **Jia, Y.**, Dai, D., Meng, X., Thiel, K. W., Leslie, K. K., & Yang, S. (2016). Inverse relationship between progesterone receptor and Myc in endometrial cancer. *PLoS One, 11*(2), e0148912.

- Meng, X., Yang, S., Zhang, Y., Wang, X., Goodfellow, R. X., **Jia, Y.,** ... & Leslie, K. K. (2015). Genetic deficiency of Mtdh gene in mice causes male infertility via impaired spermatogenesis and alterations in the expression of small non-coding RNAs. *Journal of Biological Chemistry, 290*(19), 11853-11864.

- Yang, S., **Jia, Y.**, Liu, X., Winters, C., Wang, X., Zhang, Y., ... & Xu, Y. (2014). Systematic dissection of the mechanisms underlying progesterone receptor downregulation in endometrial cancer. *Oncotarget, 5*(20), 9783.

- Yang, S., Xiao, X., **Jia, Y.**, Liu, X., Zhang, Y., Wang, X., ... & K Leslie, K. (2014). Epigenetic modification restores functional PR expression in endometrial cancer cells. *Current pharmaceutical design*, 20(11), 1874-1880.

- Leslie KK, Thiel KW, Goodheart MJ, De Geest K, **Jia Y**, Yang S. Endometrial cancer. Obstet Gynecol Clin North Am. 2012 Jun;39(2):255-68. doi: 10.1016/j.ogc.2012.04.001.

Presentations

- (Talk) "Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg", *JSM (virtual),* Aug 2021

- (Poster) "Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg", *ASA Pittsburgh Chapter Spring Banquet (virtual),* Apr 2021.

- (Poster) "Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg", *Dean's Day, Graduate School of Public Health, University of Pittsburgh (virtual),* Apr 2021.

- (Poster) "Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg", *ENAR (virtual),* Mar 2021.

- (Poster) "Deep Learning for Quantile Regression under Right Censoring: DeepQuantreg", *Biostatistics Student Research Day, University of Pittsburgh (virtual)*, Mar 2021.

- (Poster) "Deep Learning for Quantile Regression: DeepQuantreg", *ICSA (virtual),* Dec 2020.

- (Talk) "Cause-specific Quantile Regression on Inactivity Time", *ENAR (virtual),* Mar 2020.

- (Talk) "Cause-specific Quantile Regression on Inactivity Time", *Biostatistics Student Research Day, University of Pittsburgh*, Pittsburgh, PA, Feb 2020.

Statistical Packages

DeepQuantreg: A Python package (with GPU acceleration) for deep (Deep) censored quantile (Quant) regression (reg). Available on GitHub. link

QRegIT: An R package (with functions written in Rcpp) for quantile (Q) regression (Reg) on inactivity (I) time (T). Available on GitHub. link

ExactVE: An R package for exact method on vaccine efficacy trial with a non-zero lower bound. (internal use at Sanofi)

TEDVisualizationTool: An R package (utilizing *ggplot2*) for data visualization in vaccine study. (Joint work with Xiaoyu Lu, internal use at Sanofi)

Computational Skills

R, Python (PyTorch, TensorFlow, Keras, Scikit-learn, NumPy, Pandas, Matplotlib, etc.), UNIX shell scripting, Stata, SAS, LaTex, cloud-based Jupyter notebook environment (e.g., Google Colab)

Professional Affiliations

Member of American Statistical Association (ASA)
Member of Eastern North American Region International Biometric Society (ENAR)
Member of International Chinese Statistical Association (ICSA)

Services

Business manager, Pitt ASA Student Chapter	Aug 2020 – Jul 2021
Student volunteer, LiDS conference	May 2019
Student volunteer, ASA Pittsburgh Chapter Poster session and Banquet	Apr 2019