# YUSI FANG

914-314-9339 | yuf31@pitt.edu | 230 N Craig St. APT707-Pittsburgh-PA-15213 | Personal Website

## Education

**Graduate School in Public Health, University of Pittsburgh** Pittsburgh, PA
*Ph.D. Candidate in Biostatistics* *Expected: Apr.2023*
Cumulative GPA: 4.0/4.0

Research Interest: High-dimensional Statistics, Meta-analysis, Statistical Genomics and Genetics, Machine Learning Methods

Relevant coursework: Bayesian Data Science, Applied Mixed Model Analysis, SAS, High-Dimensional Statistics, Introduction to Genomics Analysis, High-Dimensional Data with Omics Application, Advanced R Computing, Nonparametric Theory, Asymptotic Methods, Survival analysis

**School of Mathematical Sciences, Xiamen University** Xiamen, China
*B.S. in Mathematics and Applied Mathematics* *Aug. 2013 – May 2017*
Major GPA: 3.94 / 4.00 (top 2/71 in the program)

## Experience

**Research Assistant** Jul. 2017 – Present
*Department of Biostatistics, School of Public health, University of Pittsburgh* *Pittsburgh, PA*
Advisors: George Tseng and Zhao Ren

- Develop statistical methodology with an emphasis on novel p-value combination methods
- Collaborate with biologists for clinical and genomics data analysis
- Conduct daily maintenance and management of laboratory servers

**Biostatistics Graduate Intern** Jun. 2021 – Aus. 2021
*Biogen Inc.* *Cambridge, MA*

- Evaluated performance of 2-in-1 flexible design and its variants with comprehensive simulations
- Derived theoretical justification for 2-in-1 flexible design and its variants

**Undergraduate Research Assistant** Jul. 2016 – May 2017
*School of Mathematical Sciences, Xiamen University* *Xiamen, China*
Advisor: Wei Liang

- Developed algorithm for kernel-based semi-supervised Bayesian quantile regression with application to cell lineage data for the detection of abnormal asynchrony of division between sister cells

## Selected Research Projects

**On p-value combination of independent and frequent signals** Aus. 2021 – Mar. 2022
*Department of Biostatistics, University of Pittsburgh*

- Evaluated asymptotic efficiencies of modified Fisher's methods using the criterion of exact slope
- Performed comprehensive simulations to evaluate the finite-sample performance of modified Fisher's methods
- Developed a ensemble method (FE) to ensemble modified Fisher's method with fast computation and theoretical guarantee on asymptotic efficiency
- Developed an extension of the FE method for enhanced statistical power on detecting signals with concordant effect size directions
- Performed extensive simulations and a application to AGEMAP dataset to justify the universally competitive performance of the proposed methods

**Flexible 2-in-1 design and its potential application in the real drug development** Jun. 2021 – Aus. 2021
*Biogen Inc.*

- Derived theoretical justifications on the type I error control procedures of the flexible 2-in-1 design and its variants
- Verified the assumptions for flexible 2-in-1 design and its variants' type I error control procedures under a simulation setup that mimics the lupus drug development program
- Evaluated the numeric performance of the flexible 2-in-1 design and its variants under the alternatives mimicking the lupus drug development program

**Heavy-tailed distribution for robust combination of dependent p-values** Oct. 2019 – Sep. 2020
*Department of Biostatistics, University of Pittsburgh*

- Proposed a family of robust tests for combining dependent p-values by the transformations by regularly varying distributions
- Provided theoretical guarantee on the robustness of the test family for the type I error control
- Showed the family of tests are asymptotically optimal in terms of detection boundary for detecting weak and sparse signals
- Performed comprehensive simulations to verify our theoretical insights and provide a recommendation to the real practice
- Conducted a neuroticism GWAS application to illustrate the theoretical findings and the advantages of our proposed methods

**Outcome-guided disease subtyping for high-dimensional omics data** Jun. 2019 – Jul. 2020
*Department of Biostatistics, University of Pittsburgh*

- Developed a unified latent generative model to perform outcome-guided disease subtyping for continuous clinical outcomes with feature selection for omics data
- Modified the model for survival outcomes by embedding the accelerated failure time model
- Performed simulations to evaluate the performance of the modified model for survival outcomes

**Combining p-values for detecting heterogeneous, weak and sparse signals** Jun. 2020– Present
*Department of Biostatistics, University of Pittsburgh*

- Developed a novel modified Fisher's method to aggregate p-values based on weakly geometric system
- Showed the proposed method's asymptotic optimality in terms of separating rate under a high-dimensional setup
- Performed comprehensive simulations to evaluate the statistical power under the high-dimensional setup
- Developed fast-computing algorithm for the proposed methods based on importance sampling and cross-entropy method

**Data analysis of women over 70 years old with clinically node negative breast cancer** May. 2020– Sep. 2020
*Magee-Womens Research Institute and Foundation, University of Pittsburgh Medical Center*

- Fitted Cox-proportional hazards models for overall survival and disease-free survival analyses
- Conducted propensity score matching over selected baseline covariates and performed down-stream analysis

## Teaching experience

**BIOST 2081 Mathematical Methods for Statistics Teaching Fellow** 2021 Fall
*Department of Biostatistics, University of Pittsburgh*

**BIOST 2094 Advanced R Computing Teaching Fellow** 2021 Spring
*Department of Biostatistics, University of Pittsburgh*

**BIOST 2094 Advanced R Computing Teaching Fellow** 2020 Spring
*Department of Biostatistics, University of Pittsburgh*

## Publications

- **Fang, Yusi** & Chang, Chung, & Tseng, George. (2022+). On p-value combination of independent and frequent signals. *submitted to Biometrics.* [arxiv]
- **Fang, Yusi** & Chang, Chung, & Park, Yongseok, & Tseng, George. (2022+). Heavy-tailed distribution for combining dependent p-values with asymptotic robustness. *submitted to Statistica Sinica.* [arxiv]
- Liu, Peng & **Fang, Yusi** & Ren, Zhao & Tang, Lu & Tseng, George. (2022+). Outcome-Guided Disease Subtyping for High-Dimensional Omics Data. *submitted to Biometrics.* [arxiv]
- Yujia Li, & **Fang, Yusi**,& Peng Liu, & Tseng, George. (2022+). Association study between gene expression and multiple phenotypes in omics applications of complex diseases. *submitted to Genes.* [arxiv]
- Orr, Brian & Mahdi, Haider & **Fang, Yusi** & Strange, Mary & Uygun, Ibrahim & Rana, Mainpal & Zhang, Lixin & Mora, Adria & Pusateri, Alexandra & Elishaev, Esther & Kang, Chaeryon & Tseng, George & Gooding, William & Edwards, Robert & Kalinski, Pawel & Vlad, Anda. (2022). Phase I Trial Combining Chemokine-Targeting with Loco-Regional Chemoimmunotherapy for Recurrent, Platinum-Sensitive Ovarian Cancer Shows Induction of CXCR3 Ligands and Markers of Type 1 Immunity. Clinical Cancer Research. clincanres.3659.2021. 10.1158/1078-0432.CCR-21-3659.
- Taylor, Sarah & **Fang, Yusi** & Wield,Alyssa & Bhargava, Rohit & Lang, Susan & Tseng, George & Soong, Rinda & Edwards, Robert & Oesterreich, Steffi & Coffman, Lan. (2022). Endocrine biomarkers in low grade serous ovarian cancers and serous ovarian tumors of low malignant potential. *submitted to Gynecologic Oncology.*
- Carleton, Neil & Zou, Jian & **Fang, Yusi** & Koscumb, Stephen & Shah, Osama & Chen, Fangyuan & Beriwal, Sushil & Diego, Emilia & Brufsky, Adam & Oesterreich, Steffi & Shapiro, Steven & Ferris, Robert & Emens, Leisha & Tseng, George & Marroquin, Oscar & Lee, Adrian & McAuliffe, Priscilla. (2021). Outcomes After Sentinel Lymph Node Biopsy and Radiotherapy in Older Women With Early-Stage, Estrogen Receptor-Positive Breast Cancer. JAMA network open. 4. e216322. 10.1001/jamanetworkopen.2021.6322.
- Li, Zheqi & Wu, Yang & Bahreini, Seyed Amir & Priedigkeit, Nolan & Ding, Kai & Sartorius, Carol & Miller, Lori & Rosenzweig, Margaret & Wagle, Nikhil & Richer, Jennifer & Muller, William & Buluwela, Laki & Ali, Simak & **Fang, Yusi** & Zhu, Li & Tseng, George & Gertz, Jason & Atkinson, Jennifer & Lee, Adrian & Oesterreich, Steffi. (2021). ESR1 mutant breast cancers show elevated basal cytokeratins and immune activation. *to appear in Nature Communications.*
- Liu, Peng & Liu, Silvia & **Fang, Yusi** & Xue, Xiangning & Zou, Jian & Tseng, George & Konnikova, Liza. (2020). Recent Advances in Computer-Assisted Algorithms for Cell Subtype Identification of Cytometry Data. Frontiers in Cell and Developmental Biology. 8. 234. 10.3389/fcell.2020.00234.
- Lin, Chien-Wei & Chang, Lun-Ching & Ma, Tianzhou & Oh, Hyunjung & French, Beverly & Puralewski, Rachel & Mathews, Fasil & **Fang, Yusi** & Lewis, David & Kennedy, James & Mueller (Müller), Daniel J. & Marshe, Victoria & Jaffe, Andrew & Chen, Qiang & Ursini, Gianluca & Weinberger, Daniel & Newman, Anne & Lenze, Eric & Nikolova, Yuliya & Sibille, Etienne. (2020). Older molecular brain age in severe mental illness. Molecular Psychiatry. 1-11. 10.1038/s41380-020-0834-1.
- Grabosch, Shannon & Bulatovic, Mirna & Zeng, Feitianzhi & Ma, Tianzhou & Zhang, Lixin & Ross, Malcolm & Brozick, Joan & **Fang, Yusi** & Tseng, George & Kim, Eun & Gambotto, Andrea & Elishaev, Esther & Edwards, Robert & Vlad, Anda. (2019). Cisplatin-induced immune modulation in ovarian cancer mouse models with distinct inflammation profiles. Oncogene. 38. 10.1038/s41388-018-0581-9.
- Liang, Wei & Yuxiao Yang & **Yusi Fang**& Zhongying Zhao & Jie Hu. "Bayesian Detection of Abnormal Asynchrony of Division Between Sister Cells in Mutant Caenorhabditis elegans Embryos." Journal of Computational Biology 26, no. 5 (2019): 495-505.

## Abstract accepted by conference

- McAuliffe, Priscilla & Carleton, Neil & Zou, Jian & **Fang, Yusi** & Koscumb, Stephen & Shah, Osama & Chen, Fangyuan & Beriwal, Sushil & Diego, Emilia & Brufsky, Adam & Oesterreich, Steffi & Shapiro, Steve & Ferris, Robert & Emens, Leisha & Tseng, George & Marroquin, Oscar & Lee, Adrian. (2021). Abstract PS1-10: Outcomes after sentinel lymph node biopsy and radiation therapy in women over 70 years old with ER+, HER2-, clinically node negative breast cancer. PS1-10. 10.1158/1538-7445.SABCS20-PS1-10.

- Nasrazadani, Azadeh & Li, Yujia & **Fang, Yusi** & Shah, Osama & Atkinson, Jennifer & Lee, Joanna & McAuliffe, Priscilla & Lee, Adrian & Tseng, George & Lucas, Peter & Oesterreich, Steffi & Wolmark, Norman. (2021). Abstract PS7-15: Mixed invasive ductal lobular carcinomas (mDLC) are clinically more similar to invasive lobular carcinoma (ILC) than to invasive ductal carcinoma (IDC). PS7-15. 10.1158/1538-7445.SABCS20-PS7-15.
- Taylor, Sarah & Wield, Alyssa & **Fang, Yusi** & Bhargava, Rohit & Lang, Susan & Tseng, George & Coffman, Lan & Oesterreich, Steffi. (2020). Endocrine biomarkers in low-grade serous ovarian cancers (LGSC) and serous ovarian tumors of low malignant potential (LMP). Journal of Clinical Oncology. 38. e18045-e18045. 10.1200/JCO.2020.38.15_suppl.e18045.

## Presentations

- (Poster) "Heavy-tailed distribution for combining dependent p-values with asymptotic robustness.", Pitt Biostatistics Student Research Day, Pittsburgh, PA, February 2020.
- (Virtual Talk) " On p-value combination of independent and frequent signals", Pitt Biostatistics Student Research Day, Pittsburgh, PA, February 2022.
- (Talk) " Robust aggregation of p-values with unknown dependency structure for SNP-set test", ENAR, Houston, TX, March 2022

## Technical Skills

**Languages**: English and Chinese
**Statistical Packages and Software**: RShiny, Rcpp, STATA, SPSS, SAS
**Programming Languages**: R, Python, Linux bash shell script, LaTeX