# RUAN, Peifeng

*Phone: +1(202) 394-0508 • Email: peifeng.ruan@utsouthwestern.edu*

## Working Experiences

**Assistant Professor at UT Southwestern Medical Center.** Sept 2022 - present, Dallas, TX

- Develop risk models for atherosclerotic cardiovascular disease and heart failure using proteomics data.
- Develop gene signatures for hepatocellular carcinoma patients with HBV infection.
- Develop novel statistical models for multi-omics single cell analysis.

**Postdoctoral Associate at Yale University.** June 2020 – July 2022, New Haven, CT
**Joint program with Boehringer Ingelheim**

- Supervisor: Hongyu Zhao
- Develop novel multi-omics clustering methods with applications to Idiopathic Pulmonary Fibrosis (IPF) and cancer.
- Conduct joint model analyses for longitudinal proteomics data and time-to-event data to investigate the association between the proteomics biomarkers and IPF patients' progression outcomes.

**Statistical Analyst Intern at Genentech**. May 2018 - Aug 2018, San Francisco, CA

- Developed machine learning pipelines to identify biomarkers that define subtypes of cancer patients who would respond better to treatments, using the methods such as random forests, boosting tree models and network assisted algorithms.

**Biostatistician at GWU Biostatistics Center**. Mar 2017 - Jan 2018, Rockville, MD

- Constructed tree algorithms for identifying the diabetes patient subgroup which responds to metformin better.
- Conducted GWAS analyses with the interaction of metabolism factors for diabetes patients.

**Visiting Scholar at Columbia University Medical Center** Nov 2014 - May 2015, New York, NY

- Constructed a network assisted algorithm to prioritize cancer associated methylated genes.

**Data Analyst at NYU Langone Medical Center**. Jan 2013 – Jun 2013, New York, NY

- Conducted data analyses for a long-term project investigating the efficacy of interventions for children with attention deficit hyperactivity disorder (ADHD).

## Education

**Ph.D. in Biostatistics.** The George Washington University Jan 2016 – May 2020

- Thesis Advisor: Hua Liang
- Doctoral Thesis: Advanced Statistical Models in Cancer Research with Omics Data

**Ph.D. in Computer Science.** Fudan University Sept 2010– Dec 2015

- Thesis Advisor: Shuigen Zhou
- Doctoral Thesis: Improving Computational Methods for Association Studies in Genome-Wide Scale

**M.A. in Statistics.** Columbia University Aug 2011– Dec 2012
**B.S. in Applied Mathematics.** Fudan University Sept 2006–Jun 2010

## Publications

### Peer Reviewed Papers

- **Peifeng Ruan** and Shuang Wang. 2021. DiSNEP: a Disease-Specific gene Network Enhancement to improve Prioritizing candidate disease genes. *Briefings in Bioinformatics*, 22(4), p.bbaa241.
- **Peifeng Ruan**, Shuang Wang and Hua Liang. 2020. mirPLS: a partial linear structure identifier method for cancer subtyping using microRNAs. *Bioinformatics*, 36(19), pp.4902-4909.

- **Peifeng Ruan**, Ya Wang, Ronglai Shen, and Shuang Wang. 2019. Using association signal annotations to boost similarity network fusion. *Bioinformatics*, 35(19), pp.3718-3726.
- Hailin Huang, Jizi Shangguan, **Peifeng Ruan** and Hua Liang. 2019. Bi-level feature selection in high dimensional AFT models with applications to a genomic study. *Statistical Applications in Genetics and Molecular Biology*, 18(5).
- Ya Wang, Min Qian, **Peifeng Ruan**, Andrew E. Teschendorff, and Shuang Wang. 2019. Detection of epigenetic field defects using a weighted epigenetic distance-based method. *Nucleic Acids Research*, 47(1), pp.e6-e6.
- **Peifeng Ruan**, Jing Shen, Regina M. Santella, Shuigen Zhou, and Shuang Wang. 2016. NEpiC: a Network-Assisted Algorithm for Epigenetic Studies Using Mean and Variance Combined Signals. *Nucleic Acids Research*, 44(16), e134-e134.

### Book Chapter

- Ya Wang, Min Qian, **Peifeng Ruan**, Andrew E. Teschendorff, and Shuang Wang 2020. Detection of Epigenetic Field Defects Using a Weighted Epigenetic Distance-Based Method. In *Stem Cell Transcriptional Networks* (pp. 109-131). Humana, New York, NY.

## Honors and Awards

- The George Washington University Statistics Departmental Prize (2020)
- The George Washington University Graduate Fellowship Award (2019 – 2020)
- ASA Biopharmaceutical Workshop Student Award (2019)

## Teaching Experiences

**Teaching Assistant**. The George Washington University

| | |
|---|---|
| • STAT1053 - Introduction to Statistics for Social Sciences (Recitation). | Spring 2020 |
| • STAT4157 - Introduction to Mathematical Statistics. | Spring 2020 |
| • STAT5216(10) - Applied Multivariate Analysis, section I. | Fall 2019 |
| • STAT5216(11) - Applied Multivariate Analysis, section II. | Fall 2019 |

## Software

**DiSNEP:**
- R package for an improved analytical tool for prioritizing genes associated with diseases using gene network information.
- Reference: **Peifeng Ruan** and Shuang Wang. 2021. DiSNEP: a Disease-Specific gene Network Enhancement to improve Prioritizing candidate disease genes. *Briefings in Bioinformatics*, 22(4), p.bbaa241.

**mirPLS**:
- R package for a partial linear structure identifier algorithm identifying non-linearly outcome-associated miRNAs for cancer subtyping.
- Reference: **Peifeng Ruan**, Shuang Wang and Hua Liang. 2020. mirPLS: a partial linear structure identifier method for cancer subtyping using microRNAs. *Bioinformatics*, 36(19), pp.4902-4909.

**abSNF**:
- R package for an algorithm using association signal annotations to boost similarity network fusion for diseases subtyping.
- Reference: **Peifeng Ruan**, Ya Wang, Ronglai Shen, and Shuang Wang. 2019. Using association signal annotations to boost similarity network fusion. *Bioinformatics*, 35(19), pp.3718-3726.

**NEpiC**:
- R package for a biological network assisted greedy algorithm for identifying epigenetic biomarkers.
- Reference: **Peifeng Ruan**, Jing Shen, Regina M. Santella, Shuigen Zhou, and Shuang Wang. 2016.

NEpiC: a Network-Assisted Algorithm for Epigenetic Studies Using Mean and Variance Combined Signals. *Nucleic Acids Research*, 44(16), e134-e134.

## Ad hoc reviewer

*Bioinformatics, BioMedInformatics, Briefings in Bioinformatics, BMC Genomics, Healthcare, Human Genomics, IEEE Journal of Biomedical and Health Informatics, Life, PLOS Computational Biology*

## Presentations and Posters

- “Dat*a-type Data-Type Weighted Multi-Omics Spectral Clustering for Disease Subtyping*”
  Contributed Presentation at Joint Statistical Meetings Washington, D.C., 2022
- “Dat*a-type Data-Type Weighted Multi-Omics Spectral Clustering for Disease Subtyping*”
  Invited Presentation at ICSA 2022 Applied Statistics Symposium Gainesville, FL, 2022
- “*Disease-Specific Enhancement for Biological Networks Improves Prioritizing Candidate Disease Genes*”,
  Poster at Annual Meeting of American Society of Human Genetics Huston, TX, Oct 2019
- “*Using Association Signal Annotations to Boost Similarity Network Fusion*”
  Poster at ASA Biopharmaceutical Workshop Washington, D.C., Sept 2019
- “*Incorporating Prior Biology Information in Building Prognostic Biomarkers*”
  Invited Presentation at Biostatistics Division, Genentech San Francisco, CA, Aug 2018
- “*Weighted Similarity Network Fusion through Integrating Genomic Functional Annotation*”
  Poster at Annual Meeting of American Society of Human Genetics Orlando, FL, Oct 2017
- *“Network-Assisted Algorithm for Epigenetic Studies”*
  Invited Presentation at Biostatistics Center, GWU Rockville, MD, Aug 2017
- “*Network-assisted Method of Genome-Wide DNA Methylation Association Studies in Cancers*”
  Poster at Annual Meeting of American Society of Human Genetics Baltimore, MD, Oct 2015