

Wei Zong

PH.D. STUDENT IN BIostatISTICS

5030 Centre Ave, Apt 860, Pittsburgh, PA, 15213

☎ (+1)412-478-2437 | ✉ wez97@pitt.edu

Education

University of Pittsburgh

PH.D. IN BIostatISTICS

Pittsburgh, PA, USA

Aug. 2018 - Apr. 2023 (Expected)

- Advisor: Prof. George C. Tseng, ScD.
- Methodological research interest: Statistical Learning, High-Dimensional Data, Multi-omics Data Analysis, Bayesian Statistics.
- Collaborative research interest: Multi-omics Circadian analysis, Schizophrenia, Bipolar, Aging.
- GPA: 4.0/4.0; Qualifying Exam Rank: 1st;

University of Cambridge

MAST IN MATHEMATICAL STATISTICS

Cambridge, UK

Oct. 2016 - Jul. 2017

- GPA 4.0/4.0; Merit;
- Relevant coursework: Modern Statistical Methods (e.g., Regularized regression, SVM), Bayesian Modelling and Computation, Applied Statistics (e.g., GLM, Nonparametric regression, Mixed effects models).

Heriot-Watt University

BSC IN ACTUARIAL SCIENCE

Edinburgh, UK

Sept. 2014 - Jun. 2016

- GPA 4.0/4.0; First class honours; Rank: 1/93;
- Actuarial Exemptions: CT1-CT8;

Southwestern University of Finance and Economics

BEC IN BILINGUAL ACTUARIAL SCIENCE

Sichuan, China

Sept. 2012 - Jun. 2014

- GPA 4.0/4.0; Rank: 1/67;

Research Interest

My research interest lies in the intersection between statistical methodology and its applications to genomics and bioinformatics. I am particularly interested in statistical modeling of high-dimensional data, Bayesian modeling, cross-species congruence evaluation and circadian analysis. In addition to methodology development, I have also collaborated with researchers in the field of psychiatry for data analysis and motivation of new methodology.

Professional Experience

Department of Biostatistics, University of Pittsburgh

METHODOLOGY RESEARCH FOR DOCTORAL THESIS

Pittsburgh, PA, USA

Aug. 2018 - present

- Multivariate Guided Disease Subtyping: develop a regularized generative model to identify meaningful molecular clusters that are relevant to multiple model-selected variables collectively.
- Model-based Multiple Clustering: develop a model-based algorithm to simultaneously discover multiple clusterings from a high-dimensional dataset.
- Congruence Analysis of Model Organisms (CAMO): define novel metrics from the 3×3 contingency table constructed from Bayesian estimates to objectively quantify concordance and discordance level across multiple (cross-species/tissues) transcriptomic response analyses.

Department of Psychiatry, University of Pittsburgh Medical Center (UPMC)

GRADUATE STUDENT RESEARCHER; MENTOR: PROF. COLLEEN A. MCCLUNG

Pittsburgh, PA, USA

Aug. 2018 - Present

- Provide statistical consulting for multiple investigators.
- Analyze high-throughput multi-omics datasets, including RNA-Seq, CUT&RUN and methylation data, starting from preprocessing raw data using bioinformatics tools to downstream analysis such as differentially expressed (DE) analysis, cosinor model fitting, pathway enrichment analysis, dimension reduction, clustering analysis and visualization.
- Involved in multiple projects investigating the association between various psychiatric disorders (e.g., bipolar disorder, major depression and drug addiction) and the molecular circadian clock.

Department of Biostatistics, University of Pittsburgh

TEACHING FELLOW

Pittsburgh, PA, USA

Jan. 2022 - May. 2022

- Teach B10ST 2094-Advanced R Computing as a co-lecturer.
- Re-designed the original R computing course to include advanced programming skills (parallel computing and Rcpp) and genomics analysis using R.

Bayer U.S. LLC

STATISTICIAN INTERN; SUPERVISOR: DANIEL HAVERSTOCK

Whippany, NJ, USA

May. 2020 - Aug. 2020

- Designed and implemented Bayesian power calculation for GO/NO-GO/INDETERMINATE decisions of continuity to binary and continuous outcome proof-of-concept (PoC) trials.
- Developed a user-friendly R-shiny app for future internal use.

Grant Application Experience

NIH UL1TR001857

TOTAL FUNDING: \$10,000

- Robust order-based machine learning framework on omics data for medical decisions.
- Applied as a co-PI for the CTSI Quantitative Methodologies Pilot Program.

NIH K01MH128763

TOTAL FUNDING: \$174,722

- Title: Role of DNA methylation in regulating striatal molecular rhythm.
- Helped power calculation using preliminary data to determine the adequate sample size for circadian analysis. Analyzed preliminary methylation data to illustrate the feasibility of rhythm detection on methylation data.

NIH R01LM014142

ESTIMATED TOTAL COST: \$1,355,480; IMPACT SCORE: 25; PERCENTILE: 8

- Title: Disease subtyping guided by clinical phenotype for precision medicine.
- Helped generating preliminary results using the outcome guided disease subtyping model. Later extended to the multivariate clinical variable guided disease subtyping model as the second part of my thesis.

Mentoring Experience

Peer Mentor

CURRENT STATUS OF MENTEE (XIANGNING XUE): 4TH YEAR PHD STUDENT

Pittsburgh, PA, USA

Sept. 2019 - present

- Teach omics data preprocessing and downstream analysis in collaborative projects.
- Mentor the development of DiffCircaPipeline, a methodological project under revision of Bioinformatics.

Peer Mentor

CURRENT STATUS OF MENTEE (RUOFEI YIN): 2TH YEAR MS STUDENT

Pittsburgh, PA, USA

Jun. 2022 - present

- Mentor the development of a methodological project investigating the strange association of library size and normalized expression level in many publicly available RNA-Seq datasets.

Publication * indicates co-authorship

METHODOLOGY

Wei Zong, George C. Tseng. (2022+) Model-based multiple clustering algorithm for high-dimensional omics data. *In Preparation*.

Wei Zong, Lu Tang, George C. Tseng. (2022+) Multivariate guided disease subtyping for high-dimensional omics data. *Ready to Submit*.

Wei Zong, Tanbin Rahman, Li Zhu, Xiangrui Zeng, Yingjin Zhang, Jian Zou, Song Liu, Zhao Ren, Jingyi Jessica Li, Steffi Osterreich, Tianzhou Ma, George C. Tseng. (2022+) Transcriptomic congruence analysis for evaluating model organisms. *Tentatively Accepted in PNAS*.
<https://www.biorxiv.org/content/10.1101/2021.11.21.469371v1>

Wei Zong, Marianne L. Seney, Kyle D. Ketchesin, Michael T. Gorczyca, Andrew C. Liu, Karyn A. Esser, George C. Tseng, Colleen A. McClung, Zhiguang Huo. (2022+) Experimental design and power calculation in omics circadian rhythmicity detection. *Submitted to Bioinformatics*

Xiangning Xue, **Wei Zong**, Zhiguang Huo, Kyle D. Ketchesin, Madeline R. Scott, Kaitlyn A. Petersen, Ryan W. Logan, Marianne L. Seney, Colleen McClung and George Tseng. (2022+) DiffCircaPipeline: A framework for multifaceted characterization of differential rhythmicity. *Under Revision in Bioinformatics*

Xiangrui Zeng*, **Wei Zong***, Chien-Wei Lin, Zhou Fang, Tianzhou Ma, David A Lewis, John F Enwright, George C Tseng. (2020) Comparative Pathway Integrator: a framework of meta-analytic integration of multiple transcriptomic studies for consensual and differential pathway analysis. *Genes*. <https://www.mdpi.com/2073-4425/11/6/696>

APPLICATION

Kyle D Ketchesin*, **Wei Zong***, Mariah A Hildebrand, Madeline R Scott, Marianne L Seney, Kelly M Cahill, Vaishnavi G Shankar, Jill R Glausier, David A Lewis, George C Tseng, Colleen A McClung. (2022) Diurnal alterations in gene expression across striatal subregions in psychosis. *Biological Psychiatry*. <https://www.sciencedirect.com/science/article/abs/pii/S0006322322015220>

Darius D Becker-Krail, Kyle D Ketchesin, Jennifer N Burns, **Wei Zong**, Mariah A Hildebrand, Lauren M DePoy, Chelsea A Vadnie, George C Tseng, Ryan W Logan, Yanhua H Huang, Colleen A McClung. (2022) Astrocyte molecular clock function in the nucleus accumbens is important for reward-related behavior. *Biological Psychiatry*. <https://www.sciencedirect.com/science/article/abs/pii/S0006322322000944>

Kyle D Ketchesin*, **Wei Zong***, Mariah A Hildebrand, Marianne L Seney, Kelly M Cahill, Madeline R Scott, Vaishnavi G Shankar, Jill R Glausier, David A Lewis, George C Tseng, Colleen A McClung. (2021) Diurnal rhythms across the human dorsal and ventral striatum. *Proceedings of the National Academy of Sciences*. <https://www.pnas.org/doi/10.1073/pnas.2016150118>

Xiangning Xue, **Wei Zong**, George Tseng, Ryan Logan, Colleen McClung, Marianne Seney. (2021) Sex and disease differences in circadian rhythms of gene expression in the human brain. *Biological Psychiatry*. <https://www.sciencedirect.com/science/article/abs/pii/S0006322322000944>

Lauren M DePoy, Darius D Becker-Krail, **Wei Zong**, Kaitlyn Petersen, Neha M Shah, Jessica H Brandon, Alyssa M Miguelino, George C Tseng, Ryan W Logan, Colleen A McClung. (2021) Circadian-dependent and sex-dependent increases in intravenous cocaine self-administration in Npas2 mutant mice. *Journal of Neuroscience*. <https://www.jneurosci.org/content/41/5/1046.abstract>

Despoina Aslanoglou, Suzanne Bertera, Marta Sánchez-Soto, R Benjamin Free, Jeongkyung Lee, **Wei Zong**, Xiangning Xue, Shristi Shrestha, Marcela Brissova, Ryan W Logan, Claes B Wollheim, Massimo Trucco, Vijay K Yechoor, David R Sibley, Rita Bottino, Zachary Freyberg. (2021) Dopamine regulates pancreatic glucagon and insulin secretion via adrenergic and dopaminergic receptors. *Translational psychiatry*. <https://www.nature.com/articles/s41398-020-01171-z>

Marianne L Seney, Sam-Moon Kim, Jill R Glausier, Mariah A Hildebrand, Xiangning Xue, **Wei Zong**, Jiebiao Wang, Micah A Shelton, BaDoi N Phan, Chaitanya Srinivasan, Andreas R Pfenning, George C Tseng, David A Lewis, Zachary Freyberg, Ryan W Logan. (2021) Transcriptional alterations in dorsolateral prefrontal cortex and nucleus accumbens implicate neuroinflammation and synaptic remodeling in opioid use disorder. *Biological Psychiatry*. <https://www.sciencedirect.com/science/article/pii/S000632232101369X>

Ryan W Logan, Angela R Ozburn, Rachel N Arey, Kyle D Ketchesin, Alicia Winquist, Andrew Crain, Brian TD Tobe, Darius Becker-Krail, Matthew B Jarpe, Xiangning Xue, **Wei Zong**, Zhiguang Huo, Puja K Parekh, Xiyu Zhu, Ethan Fitzgerald, Hui Zhang, Jeffrey Oliver-Smith, Lauren M DePoy, Mariah A Hildebrand, Evan Y Snyder, George C Tseng, Colleen A McClung. (2021) Valproate reverses mania-like behaviors in mice via preferential targeting of HDAC2. *Molecular psychiatry*. <https://www.nature.com/articles/s41380-020-00958-2>

Marianne L Seney, Kelly Cahill, John F Enwright, Ryan W Logan, Zhiguang Huo, **Wei Zong**, George Tseng, Colleen A McClung. (2019) Diurnal rhythms in gene expression in the prefrontal cortex in schizophrenia. *Nature Communications*. <https://www.nature.com/articles/s41467-019-11335-1>

Awards

| | | |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| 2022 | ASA Biopharmaceutical Student Travel Grant , Evaluated by a committee composed of members from industry, the FDA, and academia. | DC, USA |
| 2022 | Outstanding Graduate Student Researcher , Awarded by the Department of Biostatistics (UPitt) for excellent collaborative work. | Pittsburgh, USA |
| 2021 | Dean's Day Biostatistics Doctoral Award , Awarded by the School of Public Health (UPitt) for poster presentation on Dean's day. | Pittsburgh, USA |
| 2016 | The Roger Gray Memorial Prize of Statistics , Awarded by the Department of Actuarial Mathematics and Statistics (HW) to the best Senior year student. | Edinburgh, UK |
| 2015 | University Prize and Scottish Widows Prize , Awarded by the Department of Actuarial Mathematics and Statistics (HW) to the best Junior year student. | Edinburgh, UK |
| 2014 | Academic Excellence Scholarship-First Prize , Awarded for excellent academic performance (SWUFE). | Sichuan, China |
| 2013 | Academic Excellence Scholarship-First Prize , Awarded for excellent academic performance (SWUFE). | Sichuan, China |

Presentations

| | |
|-----------------------------------------------------------------------------------------------------|--------------------|
| Poster Presentation at ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop | Washington DC, USA |
| "MULTIVARIATE CLINICAL VARIABLE GUIDED DISEASE SUBTYPING" | Sept. 2022 |

Poster Presentation at Symposium on Data Science & Statistics

"CAMO: A MOLECULAR CONGRUENCE ANALYSIS FRAMEWORK FOR EVALUATING MODEL ORGANISMS"

Pittsburgh, USA

Jun. 2022

Poster Presentation at ASA Pittsburgh Chapter Spring Banquet

"CAMO: A MOLECULAR CONGRUENCE ANALYSIS FRAMEWORK FOR EVALUATING MODEL ORGANISMS"

Pittsburgh, USA

Apr. 2022

Poster Presentation at Dean's Day Competition

"DIURNAL RHYTHMS ACROSS THE HUMAN DORSAL AND VENTRAL STRIATUM"

Pittsburgh, USA

Apr. 2021