Wei Zhang

+1-786-630-2400 | wei.zhang60@med.miami.edu | noblegasss.github.io
nova-weizhang | noblegasss | Google Scholar | ResearchGate
Miami, FL - 33136

Research Interests

Multi-omics Integration, Machine Learning, Deep Learning, Random Forests, Variable Selection, Meta-analysis, Biomarker Detection, Subtype Clustering, Statistical Genomics, Epigenetics, Neurodegenerative Disease and Cancers

Research Experience

University of Miami *Miami, FL*
Postdoctoral Associate Sep 2024 - Present

- Develop and implement advanced computational and machine learning methods for the analysis of large-scale omics data
- Conduct computational analyses to identify biomarkers and therapeutic targets using multi-omics data integration
- Develop and apply machine learning models for predictive analytics in biomedical research
- Maintain and optimize computational clusters and cloud computing environments to support large-scale data analysis

Graduate Research Assistant May 2022 - Aug 2024

- Collaborated with a diverse team to research and analyze genomic data for association studies, biomarker discoveries, and disease predictions in late-onset Alzheimer's Disease, triple-negative breast cancer, and colorectal cancer
- Published multiple research papers contributing to the field of biomarker detection and disease prediction
- Supported in drafting and editing grant proposals, ensuring clarity and alignment with project objectives

Education

University of Miami *Miami, FL*
Ph.D. in Biostatistics | Advisor: Chen, X. Steven, Ph.D. Aug 2024

- Dissertation: Integrative Multi-Omics Analysis Using Multivariate Random Forest

The George Washington University *Washington, DC*
M.S. in Statistics May 2019

State University of New York at Binghamton *Binghamton, NY*
B.S. in Economics Analysis & Actuarial Math May 2017

Projects

- **Multivariate Random Forest Framework for Multi-omics Data Integration** *Ongoing*
  Tools: R, Random Forest, method development
  - Developed a random forestbased method integrating multi-omics datasets to improve predictive accuracy compared to single-omics analyses significantly.
  - Implemented simulation and real data validation compared with benchmarking methods, achieving enhanced biomarker detection accuracy.
  - Created comprehensive documentation and methodology pipelines, ensuring replicability of translational research applications.
  - Ongoing work in developing multi-omics integration for disease subtyping and data imputation.
  - Github Repositories:
    - An Integrative Multi-Omics Random Forest Framework for Robust Biomarker Discovery
- **Tumor and CellLine Transcriptomes Alignment with Deep Neural Networks** *Ongoing*
  Tools: Python, Pytorch, Variational Autoencoder (VAE)
  - Developed a deep learning framework integrating Variational Autoencoders and domain-adversarial training to align tumor and cell-line transcriptional profiles.
  - Enabled generalization to new transcriptomic datasets by projecting them into a shared latent space and mitigating systematic domain-specific biases.
- **Epigenetic Biomarkers for Alzheimers and Cognitive Health** *2023 - 2025*
  Tools: R, statistical modeling, longitudinal cohort analysis
  - Identified and validated blood-based DNA methylation signatures predictive of incident dementia in longitudinal cohorts.
  - Evaluated reliability of the Illumina MethylationEPIC v1.0 platform for robust epigenomic measurements.
  - Discovered distinct cerebrospinal fluid biomarker-associated methylation profiles from Alzheimers patients and cognitively normal subjects.

- Collaborated on statistical methodology development and comprehensive data analyses to advance cognitive resilience research.
- Github Repositories:
  - Blood DNA Methylation Signature for Incident Dementia: Evidence from Longitudinal Cohorts
  - DNA Methylation Signature of a Lifestyle-based Resilience Index for Cognitive Health
  - Critical Evaluation of the Reliability of DNA Methylation Probes on the Illumina MethylationEPIC v1.0 BeadChip Microarrays
  - Distinct CSF biomarker-associated DNA methylation in Alzheimer's disease and cognitively normal subjects

- **Prediction Models for Cancer Biomarkers** *2023 - 2024*

  Tools: R, WGCNA, meta-analysis
  - Developed transcriptome-based prediction models for chemotherapy response using matched colorectal tumor-organoid gene expression data.
  - Implemented network-based biomarker selection methods to enhance prediction accuracy in oncology studies.
  - Performed a meta-analysis on triple-negative breast cancer datasets to identify robust gene signatures linked to neoadjuvant chemotherapy outcomes.
  - Contributed to interdisciplinary research initiatives, informing personalized treatment strategies in cancer care.
  - Github Repositories:
    - Enhancing Chemotherapy Response Prediction via Matched Colorectal Tumor-Organoid Gene Expression Analysis and Network-Based Biomarker Selection
    - Transcriptome Meta-Analysis of Triple-Negative Breast Cancer Response to Neoadjuvant Chemotherapy

Publications and Preprints

J=Journal, S=In Submission, T=Thesis

[J.1] **Zhang W**, Young JI, Gomez L, Schmidt MA, Lukacsovich D, Kunkle B, Chen XS, Martin ER, Wang L. (2025). **Blood DNA methylation signature for incident dementia: Evidence from longitudinal cohorts**. *Alzheimer's & Dementia*, 21:e14496.

[J.2] **Zhang W**, Wu C, Huang H, Bleu P, Zambare W, Alvarez J, Wang L, Paty PB, Romesser PB, Smith JJ, Chen XS. (2025). **Enhancing chemotherapy response prediction via matched colorectal tumor-organoid gene expression analysis and network-based biomarker selection**. *Translational Oncology*, 52:102238.

[S.1] **Zhang W**, Huang H, Wang L, Lehmann BD, Chen XS. (2025). **An Integrative Multi-Omics Random Forest Framework for Robust Biomarker Discovery**. Manuscript submitted for publication in *GigaScience*. Preprint available at bioRxiv.

[S.2] Chen XS, Lukacsovich D, Zambare W, Wu C, Huang H, **Zhang W**, Kim MJ, et al. (2025). **Integrating Tumor and Organoid DNA Methylation Profiles Reveals Robust Predictors of Chemotherapy Response in Rectal Cancer**. Preprint available at medRxiv.

[S.3] **Zhang W**, Lukacsovich D, Young JI, Gomez L, Schmidt MA, Martin ER, Kunkle BW, Chen X, OShea DM, Galvin JE, Wang L. (2024). **DNA Methylation Signature of a Lifestyle-based Resilience Index for Cognitive Health**. *Alzheimer's Research & Therapy*. In press.

[T.1] **Zhang W**. (2024). **Integrative Multi-Omics Analysis using Multivariate Random Forest**. PhD Thesis, University of Miami.

[J.3] **Zhang W**, Young JI, Gomez L, Schmidt MA, Lukacsovich D, Varma A, Chen XS, Kunkle B, Martin ER, Wang L. (2024). **Critical evaluation of the reliability of DNA methylation probes on the Illumina MethylationEPIC v1.0 BeadChip microarrays**. *Epigenetics*, 19(1):2333660.

[J.4] Lukacsovich D, O'Shea D, Huang H, **Zhang W**, Young JI, Chen XS, et al. (2024). **MIAMI-AD: An integrative knowledgebase facilitating exploration of DNA methylation across sex, aging, and Alzheimers disease**. *Database*, 2024, baae061.

[J.5] **Zhang W**, Young JI, Gomez L, Schmidt MA, Lukacsovich D, Varma A, Chen XS, Martin ER, Wang L. (2023). **Distinct CSF biomarker-associated DNA methylation in Alzheimer's disease and cognitively normal subjects**. *Alzheimer's Research & Therapy*, 15:78.

[J.6] **Zhang W**, Li E, Wang L, Lehmann BD, Chen XS. (2023). **Transcriptome meta-analysis of triple-negative breast cancer response to neoadjuvant chemotherapy**. *Cancers*, 15(8):2194.

[J.7] Silva TC, **Zhang W**, Young JI, Gomez L, Schmidt MA, Varma A, Chen XS, Wang L. (2022). **Distinct sex-specific DNA methylation differences in Alzheimers disease**. *Alzheimer's Research & Therapy*, 14(1), 121.

Presentations

O=Contributed Talk, P=Poster

[O.1] **An Integrative Multi-Omics Random Forest Framework for Robust Biomarker Discovery**, *STATGEN: Conference on Statistics in Genomics and Genetics*. May 2025. Minneapolis, MN, USA.

[P.1] **An X chromosome-wide DNA methylation study of Alzheimers disease**, *Alzheimer's Association International Conference (AAIC)*. Jul 2024. Virtual Poster.

[O.2] **Unlocking the potential of multi-omics data integration using multivariate random forest approach**, *International Biometric Society Eastern North American Region (ENAR) Annual Meeting*. Mar 2024. Baltimore, MD, USA.

[P.2] **Distinct CSF biomarker-associated DNA methylation in Alzheimer's disease and cognitively normal subjects**, *Alzheimer's Association International Conference (AAIC)*. Jul 2023. Virtual Poster.

[P.3] **Iterative Multivariate Random Forest for Feature Selection in Integrating Multi-Omics Datasets**, *Annual American Statistical Association (ASA) Florida Chapter Meeting*. Mar 2023. Gainesville, FL, USA.

Teaching Experience

EPH705 Advanced Statistical Methods *2022 - 2024*
Professor: Wang, Lily | University of Miami

STAT6201 Applied Linear Models *2018*
Professor: Barut, Emre | The George Washington University

Professional Development

Duke Electronic Health Records Study Design Workshop *Dec 2024*
Duke University, NC

Code Rigor and Reproducibility with R Boot Camp *July 2023*
Columbia University, NY

Honors and Awards

Award of Academic Merit *Aug 2024*
University of Miami

Student Competition Award *March 2023*
ASA Florida Chapter Meeting [🌐]
◦ Best student poster

Travel Award *Mar 2023*
University of Miami

Professional Services & Activities

Manuscript Peer Review
Manuscript Reviewer
◦ Nature Communication; Scientific Reports; Discover Applied Sciences; Biology Direct; Discover Oncology; Medicine in Omics

Membership
Member
◦ International Biometric Society (ENAR)
◦ American Statistical Association (ASA)
◦ International Society to Advance Alzheimer's Research and Treatment (ISTAART)

Skills

- **Proficient in R/Rstudio and Python** for package building, data analysis, and visualization
- **Comprehensive skills in SAS** for various statistical applications
- **Familiar with Linux** in system and command

Additional Information

Languages: English (Proficiency level), Mandarin (Native Speaker), Cantonese (Proficiency level)
Interests: Tennis