# ZHONGYI (JAMES) GUO

zhongyig123@gmail.com LinkedIn: zhongyi-guo Personal Website (+1) 607-262-3415 Bay Area, CA

## Summary of Experience

As a data scientist with 5+ years of experience, I am proficient in turning complex, high-dimensional data into clear, actionable insights through reporting and visualization.

## Work Experience

**DATA SCIENTIST I, W74** **03/2026 - PRESENT**

- Provided **statistical and deep learning expertise** to inform clients' business decision.
- Improved performance by 25% using advanced models (DNN, LSTM, Transformer).

**RESEARCH ASSOCIATE, THE GLADSTONE INSTITUTES** **07/2025 - PRESENT**

- **Pioneered an R package** implementing a permutation-based pipeline to reduce false positives in biomedical research, integrating eight massive-scale scRNA-seq datasets.
- Performed QC and filtered low-quality scATAC-seq data using **advanced R** frameworks.
- Applied **AlphaGenome, a transformer-based model,** to study disease progression.
- Communicated results using **story-telling and data visualization** to diverse audiences.

## Project Experience

**GRADUATE STUDENT RESEARCHER, UNIVERSITY OF CALIFORNIA, SAN FRANCISCO** **10/2023 - 06/2025**

- Applied **t-test, PCA, PLS-DA, random forest, SVM, and regression models** to high-dimensional metabolomics data after cleaning; presented results to diverse audiences.

**GRADUATE STUDENT RESEARCHER, STANFORD UNIVERSITY** **04/2024 - 06/2025**

- **Preprocessed, integrated, and clustered** multi-modal data using data science skills.
- Performed **case-control statistical analyses** (Wilcoxon rank-sum test, Wald test, and MAST) and find-tuned **CNN-based deep learning models.**

## Education

**Stanford University** – MS in Epidemiology & Clinical Research **(GPA: 3.86/4.00)** 2025
Core Courses: Deep Learning, Causal Inference, Biostatistics, Bioinformatics, Data Visualization

**Cornell University** – BS in Statistics and Biology **(GPA: 3.57/4.30)** 2023
Core Courses: Data Science, Machine Learning, Statistical Computing, Linear Algebra, Theory of Statistics, Macroeconomics, OOP and Data Structure, UNIX scripting, Computing using Python, Biostatistics

**Case Western Reserve University** – Biochemistry Major **(GPA: 4.0/4.0)** 2020
Core Courses: Calculus I & II & III, Microeconomics, pre-med

## Skills

- Languages: Python, R, SQL, SAS (Base certified), Javascript (D3.js), LaTeX, HTML, CSS
- Tools: Git, GitHub, Conda, RStudio, ipynb, MySQL, AWS EC2, HPC (Slurm), Claude
- Aesthetics: Adobe (Illustrator, Photoshop, After Effect), Procreate
- Soft: cross-functional collaboration, communication, storytelling, vibe coding

## Leadership

**EPIDEMIOLOGY M.S. STUDENT AMBASSADOR, STANFORD UNIVERSITY** **05/2024 - PRESENT**

- Responded to inquiries from applicants, explaining curriculum and resources.

**SCIENTIFIC REVIEW EDITOR (EDUCATION BRANCH), COMMUNITY HEALTHED** **03/2022 - 05/2023**

- Validated accuracy of cited information and converted scientific jargon to plain language to ensure accessibility in weekly newsletters.

**ORIENTATION LEADER, CORNELL UNIVERSITY** **08/2021 AND 01/2022**

- Conducted icebreakers to connect students; guided campus tours to welcome students.

## Teaching Experience

**CORNELL UNIVERSITY** **01/2021 - 05/2023**

- **Beta Tester** and **Teaching Assistant,** Introduction to Data Science
- **Grader,** Probability Models & Inference
- **Teaching Assistant,** Laboratory in Genetics and Genomics

**JNC STUDY ABROAD PLATFORM** **07/2022 - 08/2022**

- **Teaching Assistant,** Introductory Biology and Physics I

## Publication

[1] (Submitted to Nature Neuroscience, 2025) Shin, J.‡, Brady, E.‡, Chen, C., Lauderdale, K., Agrawal, A., Zhang, Y., Jiang, X., Nambiar, P., Herbert, J., Mallen, D., Ly, K., **Guo, Z.,** Sant, C., Thomas, R., Miller, S., Cobos, I., Palop, J.. APOE4 and Aβ synergize to drive neuronal network dysfunction and lysosomal-ER proteostasis dysregulation in the preclinical stages of Alzheimer's disease.

Manuscripts in preparation:

[1] (In Progress; Aiming for BMC Medicine) **Guo, Z.‡,** Chen, D.‡, Stopsack, K. H., Soule, P., Ajit, D., Ramamoorthy, P., Hoffmann, T. J., Chan, J. M., Mucci L. A., Graff, R. E.. Metabolomic Disparities Between Black and White Men with Metastatic Hormone-Sensitive Prostate Cancer: A Pilot Study.

[2] (In Progress; Aiming for Cell) Qu, P., Wang, T., Jessa, S., **Guo, Z.,** Guo, H., Purmann, C., Monte, E., Jiang, L., Yang, X., Zhou, B., Kundu, S., Kundaje, A., Wong, W., Hallmayer, J. F., Urban, A. E., Snyder, M. P.. Multi-modal functional genomics analysis of bipolar disorder and schizophrenia. (Title is tentative.)

[3] (In Progress; Aiming for Nature Neuroscience) Sant, C., **Guo, Z.,** Corces, M. R.. Preventing false discoveries in Alzheimer's disease single-cell sequencing data using permutation testing. (Title is tentative.)

‡ indicates co-first authorship.