

# CURRICULUM VITAE

## SHAOPENG GU

Email: [shaopeng.gu@osumc.edu](mailto:shaopeng.gu@osumc.edu)

[shaopeng.gu@usd.edu](mailto:shaopeng.gu@usd.edu)

Phone: 605-691-1562

Street Address: 2255 Kenny Rd, Columbus, OH 43020

---

### PROFESSIONAL EXPERIENCE

OHIO STATE UNIVERSITY COMPREHENSIVE CANCER CENTER

Columbus, OH

**Biostatistics Senior Analyst**

12/2023 - Present

- Develop computational methods and perform analysis for NIH *All of Us* Research Program Database
- Develop deep learning and machine learning methods to predict the outcome of diseases from real-world Electronic Health Records (EHR) and genetic data
- Perform computational analysis for single-cell, RNA, DNA, CHIP-Seq, and many other sequencing data

UNIVERSITY OF SOUTH DAKOTA

Sioux Falls, SD

**Instructor (Volunteer), Department of Internal Medicine**

02/2022 - Present

- Research Interest: Exploring the genetic risk factors for drug-related disease (Pharmacogenomics)

SANFORD HEALTH

Sioux Falls, SD

**Computational Bioinformatics Analyst**

08/2019 – 12/2023

- Develop deep learning and machine learning methods to predict the outcome of diseases from real-world Electronic Health Records (EHR) and genetic data
- Develop computational methods for calculating Polygenic Risk Score (PRS) in AWS platform using patients' genetic data and return clinical report
- Develop computational methods for genotyping Next Generation Sequencing (NGS) and Third Generation (Oxford Nanopore) sequencing data

### EDUCATION

SOUTH DAKOTA STATE UNIVERSITY

Brookings, SD

Master of Science: Mathematics – Statistics Specialization

2019

Thesis topic: *Applying Machine Learning Algorithms for the Analysis of Biological Sequences and Medical Records.*

SOUTH DAKOTA STATE UNIVERSITY

Brookings, SD

Bachelor of Science: Electrical Engineering

2018

### RESEARCH INTERESTS

Applying Artificial Intelligence (AI) in Medicine  
Bioinformatics  
Pharmacogenomics  
Precision Medicine

**PUBLICATIONS**Full list: <https://scholar.google.com/citations?user=2i987IQAAAAJ&hl=en>

1. **Gu, S.\*<sup>S</sup>**, Rajendiran, G., Forest, K., Tran, T. C., Denny, J. C., Larson, E. A., & Wilke, R. A. (2023). Drug-Induced Liver Injury with Commonly Used Antibiotics in the All of Us Research Program. *Clinical Pharmacology and Therapeutics/Clinical Pharmacology & Therapeutics*, 114(2), 404–412. <https://doi.org/10.1002/cpt.2930>
2. Kwak, S. H., Hernandez-Cancela, R. B., DiCorpo, D. A., Condon, D. E., Merino, J., Wu, P., Brody, J. A., Yao, J., Guo, X., Ahmadizar, F., Meyer, M., Sincan, M., Mercader, J. M., Lee, S., Haessler, J., Vy, H. M. T., Lin, Z., Armstrong, N. D., **Gu, S.**, . . . Meigs, J. B. (2024). Time-to-Event Genome-Wide Association Study for incident cardiovascular disease in people with Type 2 Diabetes. *Diabetes Care*. <https://doi.org/10.2337/dc23-2274>
3. Fohle, E., Nadkarni, N., Naseer, U., Stys, P., Weaver, M., Stys, K., Rotter, J., Merrill, L., Yu, E., Guo, X., Larson, E., Hajek, C., Bartaria, S., **Gu, S.**, Rajpurohit, N., Petrasko, M. S., Stys, A. T., & Stys, T. P. (2024). ATRIAL FIBRILLATION POLYGENIC RISK SCORE IS CORRELATED WITH THE INCIDENCE OF NONISCHEMIC CARDIOMYOPATHY. *Journal of the American College of Cardiology*, 83(13), 124. [https://doi.org/10.1016/s0735-1097\(24\)02114-4](https://doi.org/10.1016/s0735-1097(24)02114-4)
4. Oleszak, F., Maryniak, A., Coy, K., Weaver, M., Pham, M., Stys, K., Anders, J., Ung, A., Hajek, C., Bartaria, S., **Gu, S.**, Rotter, J. I., Guo, X., Tan, J., Larson, E., Stys, A. T., Petrasko, M. S., Rajpurohit, N., & Stys, T. P. (2024). ELEVATED CORONARY ARTERY POLYGENIC RISK SCORE IS ASSOCIATED WITH SEVERE AORTIC STENOSIS. *Journal of the American College of Cardiology*, 83(13), 2035. [https://doi.org/10.1016/s0735-1097\(24\)04025-7](https://doi.org/10.1016/s0735-1097(24)04025-7)
5. Maryniak, A., Oleszak, F., Coy, K., Weaver, M., Anders, J., Yu, E., Stys, K., Hajek, C., Bartaria, S., **Gu, S.**, Rotter, J., Guo, X., Tan, J., Larson, E., Rajpurohit, N., Petrasko, M. S., Stys, A. T., & Stys, T. P. (2024). ATRIAL FIBRILLATION POLYGENIC RISK SCORE CORRELATES WITH INCIDENCE OF ISCHEMIC CEREBROVASCULAR EVENTS. *Journal of the American College of Cardiology*, 83(13), 1776. [https://doi.org/10.1016/s0735-1097\(24\)03766-5](https://doi.org/10.1016/s0735-1097(24)03766-5)
6. **Gu, S.\***, Lovrien, S., Qamar, M. Z., & Wilke, R. A. (2023). Rural land management and kidney Health. *PubMed*, 76(3), 102–110. <https://pubmed.ncbi.nlm.nih.gov/36898195>
7. Ma, A., Wang, X., Li, J., Wang, C., Xiao, T., Liu, Y., Cheng, H., Wang, J., Li, Y., Chang, Y., Li, J., Wang, D., Jiang, Y., Su, L., Xin, G., **Gu, S.**, Li, Z., Liu, B., Xu, D., & Ma, Q. (2023). Single-cell biological network inference using a heterogeneous graph transformer. *Nature Communications*, 14(1). <https://doi.org/10.1038/s41467-023-36559-0>
8. Fohle, E., Nadkarni, N., Oleszak, F., Hajek, C., Stys, A., Petrasko, M. S., Rajpurohit, N., Larson, E., Bartaria, S., **Gu, S.**, Weaver, M., Rotter, J. I., Guo, X., & Stys, T. (2023). Abstract 13144: High Polygenic Risk Score is a Predictor for Major Adverse Cardiac Event. *Circulation*, 148(Suppl\_1). [https://doi.org/10.1161/circ.148.suppl\\_1.13144](https://doi.org/10.1161/circ.148.suppl_1.13144)
9. Maryniak, A., Nadkarni, N., Singh, A., Hajek, C., Larson, E., Bartaria, S., **Gu, S.**, Weaver, M., Rotter, J. I., Guo, X., Rajpurohit, N., Petrasko, M. S., Stys, A., & Stys, T. P. (2023). Abstract 12423: CAD polygenic risk score, a novel screening tool for predicting likelihood of coronary artery disease, was not superior to coronary calcium score in predicting MACE. *Circulation*, 148(Suppl\_1). [https://doi.org/10.1161/circ.148.suppl\\_1.12423](https://doi.org/10.1161/circ.148.suppl_1.12423)

10. Naseer, U., Fohle, E., Hajek, C., Petrasko, M. S., Stys, A., Rajpurohit, N., Larson, E., Weaver, M., Bartaria, S., **Gu, S.**, & Stys, T. (2023). Abstract 15267: Differential effect of calcium score on major adverse cardiac events in men and women. *Circulation*, *148*(Suppl\_1). [https://doi.org/10.1161/circ.148.suppl\\_1.15267](https://doi.org/10.1161/circ.148.suppl_1.15267)
11. Shaukat, M. H. S., Massmann, A., Van Heukelom, J., Knouse, M., Stys, P., Glanzer, A., Rotter, J., **Gu, S.**, Christensen, K., Guo, X., Hickingbotham, M., Zoltick, E., Larson, E., Hajek, C., Petrasko, M. S., Rajpurohit, N., Stys, A. T., & Stys, T. P. (2023). DOES TIMING OF CYP2C19 TESTING AFFECT MAJOR ADVERSE CARDIOVASCULAR EVENT AND BLEEDING RISK IN PATIENTS INITIATED ON DUAL ANTI-PLATELET THERAPY? REAL-WORLD EXPERIENCE FROM a POPULATION GENOMIC SCREENING PROGRAM. *Journal of the American College of Cardiology*, *81*(8), 1219. [https://doi.org/10.1016/s0735-1097\(23\)01663-7](https://doi.org/10.1016/s0735-1097(23)01663-7)
12. Shaukat, M. H. S., Van Heukelom, J., Massmann, A., Knouse, M., Glanzer, A., Stys, P., Rotter, J., Zoltick, E., Hickingbotham, M., **Gu, S.**, Christensen, K., Guo, X., Hajek, C., Larson, E., Rajpurohit, N., Petrasko, M. S., Stys, A. T., & Stys, T. P. (2023). EFFECT OF CYP2C19 TEST TIMING ON GENOTYPE-GUIDED P2Y12 INHIBITON IN ACUTE CORONARY SYNDROME AND PCI: INSIGHT FROM a POPULATION GENOMIC SCREENING PROGRAM. *Journal of the American College of Cardiology*, *81*(8), 1208. [https://doi.org/10.1016/s0735-1097\(23\)01652-2](https://doi.org/10.1016/s0735-1097(23)01652-2)
13. Nadkarni, N., Fohle, E., Singh, A., Hajek, C., Stys, A. T., Petrasko, M. S., Rajpurohit, N., Larson, E., Bartaria, S., **Gu, S.**, Weaver, M., Rotter, J., Guo, X., & Stys, T. P. (2023). ADDITION OF CAD POLYGENIC RISK SCORE TO CORONARY ARTERY CALCIUM SCORE ENHANCES PREDICTIVE VALUE OF MAJOR ADVERSE CARDIOVASCULAR EVENTS. *Journal of the American College of Cardiology*, *81*(8), 1118. [https://doi.org/10.1016/s0735-1097\(23\)01562-0](https://doi.org/10.1016/s0735-1097(23)01562-0)
14. Cherukuri, P. F., Soe, M. M., Condon, D. E., Bartaria, S., Meis, K., **Gu, S.**, Frost, F. G., Fricke, L. M., Lubieniecki, K. P., Lubieniecka, J. M., Pyatt, R. E., Hajek, C., Boerkoel, C. F., & Carmichael, L. (2022). Establishing analytical validity of BeadChip array genotype data by comparison to whole-genome sequence and standard benchmark datasets. *BMC Medical Genomics*, *15*(1). <https://doi.org/10.1186/s12920-022-01199-8>
15. Zhao, J., **Gu, S.**, & McDermaid, A. (2019). Predicting outcomes of chronic kidney disease from EMR data based on Random Forest Regression. *Mathematical Biosciences*, *310*, 24–30. <https://doi.org/10.1016/j.mbs.2019.02.001>
16. Wilke, R. A., Qamar, M., Lupu, R. A., **Gu, S.**, & Zhao, J. (2019). Chronic kidney disease in agricultural communities. *The American Journal of Medicine*, *132*(10), e727–e732. <https://doi.org/10.1016/j.amjmed.2019.03.036>
17. McDermaid, A., Chen, X., Zhang, Y., Wang, C., **Gu, S.**, Xie, J., & Ma, Q. (2018). A new Machine Learning-Based framework for mapping uncertainty analysis in RNA-SEQ read alignment and gene expression estimation. *Frontiers in Genetics*, *9*. <https://doi.org/10.3389/fgene.2018.00313>
18. Adhikari, N., Dubey, A., Gaml, E. A., Vaagensmith, B., Reza, K. M., Mabrouk, S. a. A., **Gu, S.**, Zai, J., Qian, X., & Qiao, Q. (2016). Crystallization of a perovskite film for higher performance solar cells by controlling water concentration in methyl ammonium iodide precursor solution. *Nanoscale*, *8*(5), 2693–2703. <https://doi.org/10.1039/c5nr06687e>

19. Dubey, A., Adhikari, N., Venkatesan, S., **Gu, S.**, Khatiwada, D., Wang, Q., Mohammad, L., Kumar, M., & Qiao, Q. (2016). Solution processed pristine PDPP3T polymer as hole transport layer for efficient perovskite solar cells with slower degradation. *Solar Energy Materials & Solar Cells/Solar Energy Materials and Solar Cells*, *145*, 193–199. <https://doi.org/10.1016/j.solmat.2015.10.008>
20. Dubey, A., Adhikari, N., Venkatesan, S., **Gu, S.**, Khatiwada, D., Wang, Q., Mohammad, L., Kumar, M., & Qiao, Q. (2016b). Shelf life stability comparison in air for solution processed pristine PDPP3T polymer and doped spiro-OMeTAD as hole transport layer for perovskite solar cell. *Data in Brief*, *7*, 139–142. <https://doi.org/10.1016/j.dib.2016.02.021>

### **REVIEWER**

2023 – Present	Computational Biology and Chemistry
2023 – Present	Heliyon

### **CONFERENCE POSTERS**

1. **Gu, S.**, Forest, K., Tran, T. C., Denny, J. C., Larson, E. A., & Wilke, R. A. RISK DETERMINANTS FOR DRUG-INDUCED LIVER INJURY WITH COMMON ANTIBIOTICS IN THE ALL OF US RESEARCH PROGRAM. ASCPT 2024. March 2024. Colorado Spring, CO.
2. **Gu, S.**, Forest, K., Tran, T. C., Denny, J. C., Larson, E. A., & Wilke, R. A. DRUG-INDUCED LIVER INJURY WITH COMMON ANTIBIOTICS IN THE ALL OF US RESEARCH PROGRAM. Pharma R&D 2024. February 2024. Boston, MA.