Sarah Jiang

sarah.jiang@unc.edu • linkedin.com/in/sarahhjiang • sarahhjiang.me

RESEARCH INTERESTS

First-year PhD student and NSF Graduate Research Fellow seeking to develop **machine learning and causal inference** methods for healthcare that **integrate multimodal data** (biosignals from wearable devices, EHR, social determinants of health) with particular focus on **health equity and ethical AI** system design. Interested in utilizing self-supervised learning & representation learning approaches to bridge individual care with population-level health insights.

EDUCATION

UNC CHAPEL HILL – GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH Chapel Hill, NC
PhD Biostatistics 2025 – Present

- Supported by NSF Graduate Research Fellowship
- Li Lab – applying mathematical, statistical, and machine learning theory to real world health projects
- Relevant Courses: Probability and Statistical Inference I, Intermediate Statistical Methods, Fundamentals of Epidemiology

DUKE UNIVERSITY – PRATT SCHOOL OF ENGINEERING Durham, NC
B.S.E. Biomedical Engineering, B.A. Computer Science 2021 – 2025

- Concentrations: Data Science & Biomedical Imaging
- Activities: Research Fellow in the BIG IDEAs Lab, Head Undergraduate Teaching Assistant CS 216: Everything Data, Undergraduate Teaching Assistant Engineering First-Year Design, BME 254L (Medtech Prototyping) & 554L (Embedded Systems) Grader, Duke eNable (prosthetic hand development), Global Health Focus Program, Trinity Ambassador, Biomedical Engineering Design Fellow, Asian Student Association (Vice-President, Outreach), Rhythm & Blue A Cappella (2x President, Music Director)
- Relevant Courses: BME Data Science, Deep Learning for Protein Engineering, Cardiac Ultrasound Systems, Biomedical Signals & Systems, Imaging Systems, Medical Instrumentation, Embedded Medical Devices & Software, Medtech Device Prototyping, Ethics of Bioengineering, Data Structures & Algorithms, Database Systems, Computer Architecture, Design & Analysis of Algorithms, Quantitative Physiology & Biostatistical Applications, Modeling Cellular & Molecular Systems, Evidence for Policy in a Pandemic, Linear Algebra, Multivariable Calculus, Ordinary and Partial Differential Equations

HONORS/AWARDS

GRADUATE RESEARCH FELLOWSHIP — NATIONAL SCIENCE FOUNDATION, SPRING 2025

THEO C. PILKINGTON MEMORIAL AWARD — DUKE UNIVERSITY, SPRING 2025

Established in 1993 by the Whitaker Foundation in memory of Theo C. Pilkington, this award recognizes one student in the graduating class demonstrating outstanding perseverance and accomplishment in the study of biomedical engineering. Through determination and effort, the recipient of the award has succeeded academically and has exhibited characteristics that ensure personal and professional success in the future. Selected at the end of the student's junior year, presented at graduation.

INDEPENDENT STUDY RESEARCH GRANT — DUKE UNIVERSITY, FALL 2024

CONFERENCE TRAVEL GRANT — DUKE UNIVERSITY, FALL 2024

DEAN'S LIST WITH DISTINCTION — DUKE UNIVERSITY, FALL 2021

SKILLS

- Programming & Web: Python, Java, C/C++, SQL, MATLAB, HTML/CSS, LaTeX, Flask, Django, FastAPI
- AI/ML: PyTorch, TensorFlow, scikit-learn, foundation modeling, causal inference, self-supervised & representation learning
- DevOps/Infra: Azure, AWS, PostgreSQL
- Hardware & Design: SOLIDWORKS, Fusion 360, KiCad, PCB design
- Languages: Mandarin (native), English (native), Spanish (Intermediate)

PUBLICATIONS & PRESENTATIONS

M. M. H. Shandhi, H. Jeong, **S. Jiang**, P. Ashar, S. Kavirayuni, A.V. Kotla, M. Fudim, H. Pontzer, W.E. Kraus, J. Dunn, *"Assessment of Cardiorespiratory Fitness and Functional Capacity: From Clinics to Real-World Settings,"* under review at **Circulation**, Aug. 2025

S. Jiang, P. Ashar, M. M. H. Shandhi, and J. Dunn, *"Demographic reporting in biosignal datasets: a comprehensive analysis of the PhysioNet open access database,"* **The Lancet Digital Health**, vol. 6, no. 11, pp. e871–e878, Oct. 2024, doi: 10.1016/S2589-7500(24)00170-5.

L. Lederer, M. Liu, B. Chen, **S. Jiang**, S. Kim, D. MacKenzie, E. Ho, G. Guerreri, A. Roghanizad, J. Dunn, *Determinants of Opioid Use Disorder Relapse from the Biopsychosocial Perspective: A Systematic Review* [Poster Presentation]. In Jan. 2025 CERSI Summit, San Francisco, CA.

S. Jiang, J. Dunn, *Analyzing Demographic Data Gaps in the PhysioNet Open Access Database: Toward Mitigating AI Bias in Biosignal Algorithms* [Poster Presentation]. In Oct. 2024 Biomedical Engineering Society (BMES), Baltimore, MD.

L. Lederer, M. Liu, B. Chen, **S. Jiang**, S. Kim, D. MacKenzie, E. Ho, G. Guerreri, A. Roghanizad, J. Dunn, "*Determinants of Opioid Use Disorder Relapse from the Biopsychosocial Perspective: A Systematic Review,"* under review at the FDA. Aug. 2024.

P. Yang, **S. Jiang,** J. Wang, P. Chang, S. Sakai, A. Chompre, D. Bajaj, A. Kaakati, B. Chen, L. Lederer, K. Singh, P. Cho, A. Roghanizad, J. Dunn, *Improving Infection Detection With Wearable Device Data* [Poster Presentation]. In May 2023 Bass Connections Research Symposium, Durham, NC.

A. Chompre, **S. Jiang**, P. Yang, *Improving infection detection efficiency with wearables* [Poster Presentation]. In Aug. 2022 Data+ Symposium, Durham, NC.

RESEARCH EXPERIENCE & PROJECTS

MOVEMENT FOUNDATION MODELS UCLA
Hail Lab May 2025 – Present

- Using contrastive learning and autoregressive embeddings on high-frequency accelerometer data to pretrain general-purpose health foundation models
- Working under Dr. Yuzhe Yang

PROTEIN LANGUAGE MODELING Duke University
BME 590 (Graduate) Jan 2025 – April 2025

- Designed a classifier-guided protein generation pipeline using natural language prompts (CLIPPro + DPLM)
- Benchmarked general vs. domain-specific protein LMs (GPTNeo, ProtGPT2, ProGen2), integrating custom tokenization and structural evaluation to assess tradeoffs in generative quality
- Enhanced IDR and signal peptide prediction through ESM-2 embeddings and BiLSTM/CNN models, leveraging LoRA and partial fine-tuning to improve performance

HEART DISEASE OUTCOME PREDICTION Duke University
BIG IDEAs Lab Aug 2024 – Present

- Developing Causal Inference Techniques and predictive algorithms for heart disease outcomes using wearable (Fitbit), electronic health record (EHR), healthcare access and utilization, and social determinant data
- Identifying biomarkers for heart function/disease indicators via activity and HR/HRV data
- Utilizing All of Us survey data (Social Determinants of Health, Basics, Lifestyle, Healthcare Access and Utilization)
- All of Us data merged with ScHARe for ground-truth healthcare access data

PHYSIONET DATABASE SYSTEMATIC REVIEW Duke University
BIG IDEAs Lab May 2023 – June 2024

- **First author of manuscript published in the Lancet Digital Health** regarding the demographic imbalances in biosignal research and open access data
- Conducted database review of 176 datasets and studies, identifying relationships in demographic data reporting and collection in biosignal studies

WEARABLE INFECTION DETECTION ENGINE Duke University
BIG IDEAs Lab Aug 2023 – Jan. 2025

- Analyzed COVID-19 testing, demographics data from the greater Durham community, and wearable data to develop a data-driven timeline for disease onset & extracted missingness features via machine learning and signal processing algorithms
- Merged wearable data with electronic health records, symptom data, and test results
- Studied wearable data missingness patterns and incorporated this as a separate feature
- Pivoted from COVID prediction engine to developing anomaly detection pipeline for symptom onset, ongoing model development

RECYCLEHEALTH RESEARCH DASHBOARD Duke University
BIG IDEAs Lab May 2024 – Nov. 2024

- **Deployed and maintained web app** on servers; developed unit tests for user sign-in, account data storage, device donations and requests, and user flow aspects for Recycle Health (wearable non-profit)
- Developed geographic visualizations for where devices were donated from and being sent to post-processing
- Built end-to-end Flask web app coded in Python, HTML, & CSS hosted via Render with backend postgreSQL database to store wearable device and donor data

CARDIORESPIRATORY FITNESS & FUNCTIONAL CAPACITY REVIEW Duke University
BIG IDEAs Lab Feb 2022 – May 2025

- **Co-author of manuscript of scoping review paper** studying the role of wearable devices in improving cardiorespiratory fitness and functional capacity outcomes

- Comparing the gold standard clinical measurement tools with commercial wearable devices to evaluate metrics such as SpO_2, heart rate, heart rate variability, activity, and sleep

OPIOID USE DISORDER SYSTEMATIC REVIEW Duke University
BIG IDEAs Lab May 2024 – Aug 2024

- **Co-author of manuscript for systematic review** of Opioid-Use Disorder relapse and craving to develop a biosignal prediction engine for OUD relapse
- Conducted extensive literature review regarding social determinants of health that affect opioid use relapse and craving

FULL-STACK WEB APP DEVELOPMENT Duke University
BIG IDEAs Lab May 2022 – May 2023

- **Built a cloud-based (Azure) lab-wide web app** to navigate between surveys, wearable data, and user login and account information using Python and HTML
- Developed unit tests to assess code functionality for user accounts, data collection, and database connections
- Self-taught Django, FastAPI and integrated OAuth 1.0 and 2.0 libraries to allow for research use of high-resolution wearable data and developed cron jobs to automatically pull data upon initial authorization into SQL databases

TEACHING EXPERIENCE

DUKE UNIVERSITY COMPUTER SCIENCE Durham, NC
Head Undergraduate Teaching Assistant Dec 2022 – May 2025

- Developed curriculum final data science application project, created updated demo materials from previous 2020 demo milestones and led teams of 10 UTAs each semester to guide and provide feedback for 80+ student teams
- Created weekly and monthly check-in points for student teams and open forms for students to report concerns throughout the semester
- Held weekly office hours, graded projects and exams, assisted with in-class coding exercises, attended weekly teaching team meetings

DEWEY SMART Remote
Academic Tutor & Admissions Counselor May 2022 – May 2024

- Tutored high school geometry, chemistry, biology, English, and SAT reading, writing, and math for four students
- College admissions and scholarship counseling for five high school seniors

FIRST-YEAR DESIGN Durham, NC
Undergraduate Teaching Assistant Aug 2022 – Dec 2022

- Guided 4 student teams through the engineering design process to present deliverables to local clients via semester-long design projects
- Supervised use of Innovation CoLab spaces for engineering students to safely prototype community design solutions