Star Liu

U.S. Citizenship   👤 Personal Website   🗐 Latest CV   🎓 Google Scholar   @ sliu197@jhmi.edu   📱 +470 538 5974

🏛 EDUCATION

2021-2023	Master of Science in Biomedical Informatics and Data Science Research
	Johns Hopkins University School of Medicine, Baltimore, Maryland, United States
2017-2021	Bachelor of Science in Quantitative Sciences, Biology Concentration
	Emory University, Atlanta, Georgia, United States

💻 COMPETENCIES

Programming	Python, R, SQL, HTML, JavaScript, CSS, Java
Frameworks	Flask, Dash, React JS
Databases	Microsoft SQL Server, CosmosDB, MongoDB
Vocabularies	ICD-9, ICD-10, RxNorm, LOINC
Data Sources	EHR, Claims, Healthcare Cost and Utilization Project, Medical Expenditure Panel Survey
Other Softwares	GitHub, Overleaf, Microsoft Excel, Slurm, FileZilla, Tableau, QGIS, Google Analytics

📄 PUBLICATIONS AND PREPRINTS

1. **Liu S**, Ding X, Belouali A, Bai H, Raja K, Kharrazi H. Assessing the Racial and Socioeconomic Disparities in Postpartum Depression Using Population-Level Hospital Discharge Data : Longitudinal Retrospective Study. *JMIR Pediatrics Parenting*. 2022;5(4):e38879. Published 2022 Oct 17. 🔗 doi:10.2196/38879

2. Belouali A, Bai H, Raja K, **Liu S**, Ding X, Kharrazi H. Impact of Social Determinants of Health on Improving the LACE Index for 30-Day Unplanned Readmission Prediction. *JAMIA Open*. 2022;5(2):ooac046. Published 2022 Jun 10. 🔗 doi:10.1093/jamiaopen/ooac046

📄 PUBLICATIONS UNDER REVIEW

1. Sood P, **Liu S**, Kharrazi H. Evaluating the Effect of Type-2 Diabetes Phenotype Definitions in Identifying Populations of Interest for Clinical Studies and Research. *BMJ Open Diabetes and Care*. [Under Review] 🔗 Link to Manuscript

2. Hassan S, **Liu S**, Patel SA, Emmert-Fees K, Suvada K, Tandon N, Sosale A, Anjana RM, Mohan V, Chwastiak L, and Ali MK. Association of collaborative care intervention components with depression and metabolic outcomes in the INDEPENDENT study : A mixed methods study. *J GEN INTERN MED*. [Under Review] 🔗 Link to Manuscript

3. Benameur K, Gross PJ, **Liu S**, Koneru S. The Mediterranean Soul Food Diet Intervention after Stroke, a Feasibility Study in the Stroke Belt. *Neurology*. [Under Review] 🔗 Link to Manuscript

📄 PUBLICATIONS IN PREPARATION

1. **Liu S**, Wei SX, Lehmann HP. Applicability Area : A Novel Utility-based Approach for Evaluating Predictive Models Beyond Discrimination. *AMIA*. [In Preparation] 🔗 Link to Manuscript

2. **Liu S**, Nagy P. A Holistic Framework for Understanding the Ecosystem of An Open Scientific Community Using Public Artifacts. *JAMIA Open*. [In Preparation] 🔗 Link to Latest Draft

3. Sood P, **Liu S**, Kharrazi H. Assessing the Effect of Data Quality and Data Type on Phenotyping Type-2 Diabetes using EHR and Claims Data. *JAMIA Open*. [In Preparation]

4. Sood P, **Liu S**, Kharrazi H. Impact of Type-2 Diabetes Phenotype Definitions in Predicting Healthcare Utilization Using EHR vs. Claims. *JAMIA Open*. [In Preparation]

5. Sood P, **Liu S**, Kharrazi H. Assessing the racial bias in Predicting Healthcare Utilization Stemmed from Variations in Diabetes Phenotype Definitions. *JAMIA Open*. [In Preparation]

*See latest CV for updates (found in the header).

🎤 PRESENTATIONS, POSTERS, AND CASE STUDIES

1. Sood P, **Liu S**, Kharrazi H. Measuring the Effect of EHR Data Quality in Identifying Type-2 Diabetes Population Across Common Phenotype Definitions of Diabetes. *AMIA Symposium*. 2022. [Poster] 🔗 Link to Poster

2. **Liu S**, Golozar A, McLeggon JA, Black A, Nagy P. The OHDSI Community Dashboard : Tracking the Health and Impact of the Open Science Observational Health Data Sciences and Informatics Community. *Observational Health Data Sciences and Informatics (OHDSI) Symposium*. 2022. [Software Demonstration] 🔗 Link to Abstract   🔗 Link to Software

3. **Liu S**, Wall E, Patel SA, Park Y. COVID-19 Health Equity Dashboard - Addressing Vulnerable Populations. *Visualization for Communication Workshop*. 2020. doi:10.31219/osf.io/2frha. [Case Study] ⬏ Link to Case Study

☰ RESEARCH EXPERIENCE

November 2021 Present -

Research Assistant | PI : Hadi Kharrazi | Bloomberg School of Public Health, BATIMORE, MD
Study the effect of data quality on using phenotype definitions to identify cohorts with diabetes
> Derive descriptive statistics from the EHR and claims consisting of 200,000+ patients using SQL
> Simulate induced data quality issues on ICD, RxNorm, and LOINC Codes for both claims and EHR data
> Construct logistic regression models to assess disparities in data quality by race in claims and EHR
> Assist the first author in writing and preparing 4 manuscripts for submission

Data Quality | Phenotyping | EHR | Claims | R | SQL | Logistic Regression | Manuscript

November 2021 Present -

Graduate Researcher | Advisor : Harold Lehmann | Johns Hopkins School of Medicine, BALTIMORE, MD
Assessing the impact of asymmetric cost in machine learning – clinical utilities analysis
> Program fully factorial simulations with 6 parameters and 6 machine learning models using Python
> Establish a novel utility-based approach to evaluating machine learning model performance
> Apply the novel utility-based approach to 3 medical datasets and compare against traditional metrics
> Lead in writing the manuscript for journal submission

Decision Analysis | Python | Simulation | Cost-effectiveness | Manuscript

October 2021 Present -

Graduate Researcher | Advisor : Paul Nagy | Johns Hopkins School of Medicine, BALTIMORE, MD
OHDSI Community Dashboard : Assessing the health and impact of the OHDSI ecosystem
(`http://ohdsi.azurewebsites.net/`)
> Develop an automated pipeline to collect data from PubMed and map texts to standard terminologies
> Build the front-end dashboard using Flask and Dash supported by Azure back-end
> Establish an open-source framework for evaluating open-source scientific communities
> Lead in writing the manuscript for journal submission

OHDSI | APIs | Flask | Dash | HTML | Azure CosmosDB | GitHub | Manuscript

January 2021 Present -

Research Assistant | PI : Mohammed Ali | Rollins School of Public Health, ATLANTA, GA
The associations of collaborative care intervention components with depressive symptoms and metabolic outcomes in the INDEPENDENT study : A mixed methods study
> Analyze longitudinal data on 196 patients and conducted multiple logistic regressions using R
> Write the statistical methods section and produce the appropriate figures for the manuscript
> Assist the first author in writing and preparing the manuscript for journal submission

Longitudinal clinical trials data | R | Mixed methods | Regression Analysis | Manuscript

October 2021 September 2022

Graduate Researcher | Advisor : Hadi Kharrazi | Johns Hopkins School of Medicine, BALTIMORE, MD
Assessing the Racial and Socioeconomic Disparities in Postpartum Depression Using Population-level Hospital Discharge Data : A Longitudinal Retrospective Study [Published]
> Conducted literature reviews on postpartum depression, diagnosis, and racial disparities
> Performed logistic regression, multinomial regression, survival analysis, and sensitivity analysis using R and Healthcare Cost and Utilization Project datasets for Maryland (2016-2019)
> Led the manuscript writing, journal submission, revisions, and correspondence

Manuscript | Maternal Health Disparity | HCUP | R | Regression Analysis

May 2020 June 2021

Research Assistant | PI : Shivani Patel | Rollins School of Public Health, ATLANTA, GA
COVID-19 Health Equity Dashboard (`https://covid19.emory.edu/`).
> Transformed data using Python and led front-end development using NPM, React JS, and HTML
> Implemented and managed backend MongoDB database solutions using Python and JavaScript
> Led the writing as the first author of a case study accepted by the IEEE VisComm 2020

COVID-19 | Interactive Dashboard | JavaScript | Python | MongoDB | Case Study Presentation

January 2020 May 2020

Undergraduate Researcher | Advisor : Seunghwa Rho | Emory University, ATLANTA, GA
Atlanta Fire and Rescue Department : Predicting Fire Incident Risk ⬏ Link to Report   ⬏ Link to PPT
> Pre-processed and aggregated +8 million observations of geospatial data on historical fires
> Optimized a random forest model to predict fires with 40% sensitivity and 90% specificity
> Created an R Shiny interactive dashboard to visualize fire predictions and model performance
> Presented findings to the Atlanta Mayor's Office, City Planning, and the Fire Department

Predictive Modeling | Random Forest | Interactive Dashboard | Geospatial and Longitudinal Data

September 2019 June 2020	Volunteer Research Assistant | Rollins School of Public Health, ATLANTA, GA

> - Wrangled Medical Expenditure Panel Survey and Healthcare Cost and Utilization Project data using R
> - Conducted literature reviews on diabetes, preventable hospitalization, and care utilization
> - Conducted consistent 30, 90, and 120-day phone interviews with a cohort of 24 patients with stroke

Diabetes | Survey data | Panel data | R | Literature Review | Phone Interviews

💼 PROFESSIONAL EXPERIENCE

May 2019 August 2019	Analyst | Supervisor : Michael Trader | RightPatient, DUNWOODY, GA

> - Sourced 2 million images from public safety agencies for the R&D project on opioid addiction
> - Developed 4 client leads for the ACO patient-centered care software and relayed to the demo team
> - Entrusted with managing 358 client accounts on Zoho CRM for our patient wellness initiative

Healthcare Technology Startup | R&D | Customer Relations | Sales

👥 TEACHING & MENTORSHIP

Oct 2022 Dec 2022	Graduate Teaching Assistant | Johns Hopkins School of Medicine, BALTIMORE, MD

Clinical Data Analysis with Python

> - Assist the faculties in teaching Python and host weekly office hours

Python | Office Hours

July 2022 July 2022	Bootcamp Teaching Assistant | PI : Hadi Kharrazi | Bowie State University, BOWIE, MD

Public Health Informatics and Technology 4-Week Bootcamp

> - Assisted the program faculties in teaching R and Tableau to a class of 15 students
> - Hosted daily office hours and coached 4 capstone project groups

Bootcamp | R | Visualization | Capstone projects

May 2018 May 2021	Academic Fellow Captain, Academic Fellow | Emory University, ATLANTA, GA

Annual International Student Welcome and Mentoring Program

> - Coordinated two orientations with 14 colleagues, oversaw 26 fellows, and mentored 24 mentees

Peer Mentorship | Program coordination

✒️ PEER REVIEWS

2021 - Present	Reviewer | JAMIA Open

❇️ CERTIFICATES

April 2021 - No Expiration Date	Information Privacy Security | JHMI CITI Program | Credential ID : 45798326
April 2021 - No Expiration Date	Health Privacy Issues for Researchers | JSPH CITI Program | Credential ID : 45562106
October 2021 - October 2026	Basic Human Subjects Research | JSPH CITI Program | Credential ID : 45562104
October 2021 - October 2024	Social and Behavioral Research Best Practices for Clinical Research | JSPH CITI Program | Credential ID : 45562105
October 2021 - October 2023	Human Subjects Research - Biomedical Research | JHMI CITI Program | Credential ID : 45733041

🌐 LANGUAGES

English	Native
Chinese	Native
Spanish	Beginner