

Keren Xu

San Francisco, CA

✉ keren.xu@tempus.com | 🏠 xukeren.github.io | 📺 XUKEREN | 🌐 [kerenxu](https://kerenxu.com) | 🐦 [kerenxuepi](https://twitter.com/kerenxuepi)

Education

University of Southern California, Keck School of Medicine

Los Angeles, CA

Doctor of Philosophy in Epidemiology, GPA: 3.92/4.00

2018-2022

- Dissertation: Genetic Epidemiological Approaches in the Study of Risk Factors for Hematologic Malignancies

Columbia University, Mailman School of Public Health

New York, NY

Master of Public Health in Epidemiology, Certificate in Advanced Epidemiology, GPA: 3.81/4.00

2015-2017

- Thesis: Modeling the Impacts of Numbers of Affected Relatives and Perceived Chance of Having Epilepsy-Related Mutation

East China University of Science and Technology, School of Pharmacy

Shanghai, China

Bachelor of Science in Pharmaceutical Sciences, Minor in English, GPA: 3.73/4.00

2011-2015

- Thesis: Pluripotent Stem Cells Culture and ICOSL Gene Knockout

Employment and Research

Tempus Labs

Redwood City, CA

Bioinformatics Scientist

June 2022-Present

- Design and conduct analysis to improve variant calling, classification and analysis systems
- Work in interdisciplinary groups of scientists, engineers, and product developers to translate research into clinically actionable insights to clients

University of Southern California Center for Genetic Epidemiology

Los Angeles, CA

Predoctoral Fellow

Aug. 2018-May 2022

- Built bioinformatics pipelines for germline/somatic variants (short variants/structural variations/copy number alterations) discovery and mutational signature analysis of over 10-terabyte human whole-genome/exome sequencing data on high-performance computing cluster
- Conducted epigenome-wide association studies, methylation quantitative trait loci analyses, causal mediation analyses, and meta-analyses to identify genetic risks that lead to leukemia by altering DNA methylation
- Conducted genome-wide association studies and construct polygenetic risk scores to assess the associations of blood cell traits with childhood ALL patient characteristics
- Detected accelerated aging in Down syndrome newborns using epigenetic age clocks derived from elastic net regression
- Assessed the impact of prenatal smoking on somatic gene deletion burden in childhood ALL patients using the polyepigenetic DNA methylation smoking score constructed by logistic lasso regression
- Held weekly lectures and office hours as a teaching assistant for 5 graduate-level courses in Biostatistics, Data analysis, and Epidemiology

Dartmouth-Hitchcock Medical Center

Lebanon, NH

Statistician

July 2017-June 2018

- Analyzed patient survey data in multiple projects using R programming
- In charge of cluster randomized trial design, survey design, IRB submission, and patient recruitment
- Collaborated with medicine fellow in writing study proposal which received \$15,000 Cardiovascular Medicine Fellowship Research Award
- Organized stakeholder group meeting including 4 clinicians and 8 patients to collect data for grant application

Columbia University Medical Center

New York, NY

Graduate Research Assistant

May 2016-July 2017

- Analyzed questionnaire data using structural equation modeling, multilevel modeling, and multinomial logistic regression in SAS and R
- Created and maintained three-year survey database in SPSS to efficiently track project outcomes

Relevant Skills

Data Analysis: Prediction, Inference, Machine Learning, Visualization, Data Cleaning, Modelling

Programming: R, Python, Linux/Unix, Bash, Git, Singularity, Docker, SAS, STATA, SPSS, AWS, GCP

Bioinformatics: GATK, PLINK, Conda, Bioconductor, WDL, EWAS, GWAS, PRS

Languages: Mandarin (Native), English (Full professional working proficiency)

Publications and Presentations

11 peer-reviewed journal publications (4 first author), 3 conference presentations, 3 workshops teaching R programming, and 10 accepted conference abstracts. Full list available on Google Scholar: <https://scholar.google.com/citations?user=pNidMJAAAAAJ&hl=en>