LANQIU "Kate" YAO

545 1st Ave, Apt 10Q, New York, NY 10016.
(919)6384885 · lanqiu.yao@nyulangone.org
https://www.linkedin.com/in/lanqiuyao

EDUCATION

NEW YORK UNIVERSITY, Grossman School of Medicine, Department of Population Health, Division of Biostatistics, New York, NY
Ph.D. of Biostatistics, May 2022
Relevant coursework includes: Statistical Inference, Advanced Topics in Survival Analysis, Causal inference, Epidemiology Method, Algorithms and Data Structures, Advanced Regression Modeling, Rigor & Reproducibility

DUKE UNIVERSITY, School of Medicine, Department of Biostatistics and Bioinformatics, Durham, NC
Master of Biostatistics, May 2018
Relevant coursework includes: Introduction to Statistical Theory and Methods I/II, Applied Biostatistics Methods I/II, Introduction to the Practice of Biostatistics I/II, Introduction to Statistical Programming I/II, Measure Theory, Generalized Linear Model, Genetic Statistics

PEKING UNIVERSITY, Health Science Center, Beijing, China
Bachelor of Medicine, Clinical Medicine, July 2016
Bachelor of Science, Statistics, July 2016
Relevant coursework includes: Principles of Mathematical Analysis, Probability Theory, Real Analysis, Advanced Algebra, Modern Algebra, Statistics, Survey Sampling

RESEARCH EXPERIENCE

NEW YORK UNIVERSITY, DEPARTMENT OF POPULATION HEALTH, New York, NY, USA   2019-Present
"Discovering linear biosignatures for treatment response based on maximizing Kullback-Leibler Divergence in linear mixed-effect models"
- Reviewed and summarized literatures for functional data analysis
- Derived an algorithm for data clustering with consideration of baseline covariates
- Conducted simulations to check the consistency of the method
- Preparing a paper for publication.

NEW YORK UNIVERSITY, DEPARTMENT OF POPULATION HEALTH, New York, NY, USA   2018-2019
"A new Algorithm for Convex Biclustering and Its Extension to Compositional Data"
- Helped to build a new Biclustering algorithm based on the alternating direction method of multipliers (ADMM)
- Algorithm optimization
- Conducted simulations to compare the clustering effects with other biclustering methods
- Analyzed microbiome data with the new biclustering algorithm

NEW YORK UNIVERSITY, DEPARTMENT OF POPULATION HEALTH, New York, NY, USA   2018-2019
"Alternative non-parametric estimation of the Kaplan-Meier estimator for dependent censoring"
- Studied non-parametric estimations method to analyze survival data without independence assumption
- Generalized the independence assumption and derived a new valid semi-parametric method under the generalized assumption.
- Conducted simulations to compare the semi-parametric estimator with Kaplan-Meier estimator.

DUKE UNIVERSITY DEPARTMENT OF BIOSTATISTICS, Durham, NC, USA   2017-2018
Studied on Methods for Handling Missing Data in Cluster Randomized Trials (CRTs)
- Summarized a literature review for existed methods aiming to deal with missing values in cluster randomized trials (CRTs)
- Conducted simulations to compare the missingness handling methods using R
- Applied missingness handling methods to analyze the data from Health and Literacy Intervention project (HALI)
- Published a paper for the project.

DUKE UNIVERSITY DEPARTMENT OF SOCIOLOGY, Durham, NC, USA   2016-2017
Conducted a Meta-analysis to present the correlation between church strictness and church growth
Data Science Intern
- Conducted a system review for relevant papers and extracted data
- Conducted meta-analysis with consideration of a hierarchical data structure and build a model with the correlations.
- Preparing a paper for the project.

PEKING UNIVERSITY INTERNATIONAL GENETICALLY ENGINEERED MACHINE COMPETITION TEAM, Beijing, China   2013-2014
Developed An Integrated Solution to Harmful Algal Bloom by Genetically Engineered Escherichia coli
Lab Manager
- Applied transgene technology to *E. coli* cells to secrete lysozyme and tested the efficiency to eliminate cyanobacteria
- Applied transgene technology to *E. coli* cells to express microcystinase for the degradation of toxins
- Surface displayed of a lectin on the E. coli cell membrane to achieve binding towards the cyanobacteria to improve the efficacy
- Won a gold medal in the find competition

SINGLE-MOLECULE & NANOBIOLOGY LABORATORY, Beijing, China   2012-2013
Research Assistant
- Performed diversified lab assistant work, including lab safety (EHS), instruments, maintenance, etc
- Assisted other lab mates to design and conduct experiments

EXPERIENCE

NEW YORK UNIVERSITY, Grossman School of Medicine, New York, NY
Teaching Assistant: Rigor & Reproducibility, Fall 2020

DUKE CLINICAL RESEARCH INSTITUTION, Durham, NC, USA   2016-2018
Prediction Model for Necrotizing Enterocolitis
Data Science Intern
- Conducted data exploration for the raw data collected from Duke hospitals.
- Applied the regular expressions to extract information from big data
- Built logistic regression models to predict the disease-Necrotizing Enterocolitis

BEIJING JISHUITAN HOSPITAL, Beijing, China   2014-2016
Intern (similar as Resident)
- Rotated in 15 clinical departments including Cardiology, Surgery, Gynecology, Obstetrics, Pediatrics
- Assisted doctors' work including conducting physical examination for patients and writing medical records
- Aided in surgeries including Appendectomy and TIPS(Transjugular intrahepatic portosystemic shunt)

HONORS and AWARDS

- Eisai (China) Pharmaceutical Scholarship. 2014
- China National Scholarship, 2013
- Beijing excellent student award, 2013
- Peking University excellent learning award, 2012, 2014

SCIENTIFIC MEMBERSHIPS

- Eastern North American Region International Biometric Society, 2021-Present
- American Statistical Association, 2018-Present

TECHNICAL SKILLS

- R, Python, SAS, C, C++, Matlab, SPSS, SPSS modeler, EViews
- Microsoft Office (Excel, Word, PowerPoint), Adobe Photoshop, Adobe After Effects
- Molecular cloning, Western blotting, PCR, ELISA, Protein Purification, Protein Engineering, Fluorescence Microscopy, Flow Cytometry

PUBLICATION

- Turner EL, **Yao L,** Li F, Prague M. Properties and pitfalls of weighting as an alternative to multilevel multiple imputation in cluster randomized trials with missing binary outcomes under covariate-dependent missingness. *Statistical methods in medical research.* 2019 Jul 11:0962280219859915

- Wang, B., **Yao, L.**, Hu, J. and Li, H., 2020. A New Algorithm for Convex Biclustering and Its Extension to the Compositional Data. arXiv preprint arXiv:2011.12182.

- Zhang Z, Meng L, Ni C, **Yao L**, Zhang F, Jin Y, Mu X, Zhu S, Lu X, Liu S, Yu C. Engineering Escherichia coli to bind to cyanobacteria. *Journal of bioscience and bioengineering.* 2017 Mar 1;123(3):347-52.

PRESENTATIONS and POSTERS

- A Single Index Model for Longitudinal Outcomes to Optimize Individual Treatment Decision Rules
  Joint Statistical Meetings, American Statistical Association, 2020
- Discovering Linear Biosignatures for Treatment Response Based on Maximizing Kullback-Leibler Divergence in Linear mixed-effect Models
  Precision Health Interest Group Meetings, NYU, Grossman School of Medicine, 2019
- Contributed Posters: Discovering Linear Biosignatures for Treatment Response: A Convexity-Based Clustering Approach
  Tom Ten Have Symposium on Statistics in Mental Health, Yale University, 2019