# JooChul Lee, Ph.D.

*206A Extension Hall, Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849*

jzl0375@auburn.edu | sites.google.com/site/joochul07

## Research Interests

- Active Learning and Testing, and Semi-supervised Learning
- Electronic Health Records Data Analysis
- Sport Data Analytics

## Employment

**Auburn University, Department of Mathematics and Statistics** — *AL, USA*

Assistant Professor — *July 2024 - present*

**University of Pennsylvania,**
**Department of Biostatistics, Epidemiology, and Informatics** — *PA, USA*

Postdoctoral Researcher — *Jun 2021 - May 2024*

## Education

**University of Connecticut** — *Storrs, CT, USA*

Doctor of Philosophy in Statistics — *2021*

Advisors : Dr. Elizabeth Schifano, Dr. Haiying Wang
- Dissertation Title: Computational Topics in the Statistical Analysis of Big Data

**Korea University** — *Seoul, South Korea*

Master of Statistics in Statistics — *2016*

**Dongguk University** — *Seoul, South Korea*

Bachelor of Science in Statistics — *2013*

## Experiences

**University of Connecticut, School of Nursing** — *CT, USA*

Research Assistant — *Summer 2018 - Spring 2021*

Advisor: Dr. Xiaomei Cong
- Analyzed longitudinal clinical data and microbiome data

**University of Connecticut, Statistical Consulting Services** — *CT, USA*

Administrative Specialist and Graduate Assistant Consultant — *Fall 2018 - Spring 2021*

Director: Dr. Timothy E. Moore
- Conducted 9 consulting projects and 4 workshops

**Korea University, Department of Statistics** — *Seoul, South Korea*

Research Assistant — *Mar 2014 - Feb 2015*
- Participated in a project, 'Statistical Testing Technology for Modeling in Pyroprocessing Safeguards '

## Grants

1. 2025 NVIDIA Academic Grant, DGX Spark Hardware
2. (Pending) NSF, CAREER: Statistically Principled Active Learning for High-Dimensional Inference and Intelligent Data Labeling, PI
3. (Pending) NIH-R03, Intelligent Active Modeling Framework for Multimodal Electronic Health Records, PI
4. (Pending) NSF, Building Undergraduate Data Literacy Through Authentic Video-Derived Data Analysis, PI
5. (Pending) NIFA, Strategic Risk Management in Southeastern Agriculture: Exploring the Interplay Between Irrigation, Crop Insurance, and Cover Crops, co-PI (PI: Sunjae Won)
6. (Pending) National Endowment for the Humanities, Transnational Academic Identity and the Cultural Dynamics of Knowledge in US Higher Education, co-PI (PI: Yuhyun Park)

## Publications

1. **JooChul Lee**, Guorong Dai, Daniel J Rader, Jinbo Chen (2025). 'Model Calibration and Evaluation via Optimal Sampling in Electronic Health Record Data', *Journal of the Royal Statistical Society: Series A.*.
2. SeJin Hong, Soyeon Park, Namsu Kim, Minsuh Chung, Youlee Jung, **JooChul Lee**, Kyounghae Kim (2025). Feasibility of a Gamified Mobile-Based Self-Management Intervention for Individuals With Nonspecific Chronic Lower Back Pain, *Nursing Research*.
3. **JooChul Lee**, Weidong Ma, Ziyang Wang (2024). Unbiased Active Semi-supervised Binary Classification Models, *The 33rd International Joint Conference on Artificial Intelligence (IJCAI)*.
4. **JooChul Lee**, Likhitha Kolla, Jinbo Chen (2024). Towards Optimal Model Evaluation: Enhancing Active Testing with Actively Improved Estimators, *Scientific Reports*.
5. Jinkyung Park, **JooChul Lee**, Kyounghae Kim (2023). Pathways through which participatory treatment decision making affects health-related quality of life among people with chronic conditions, *Journal of Health Communication*.
6. **JooChul Lee**, Elizabeth Schifano, Haiying Wang (2022). Sampling-based Gaussian Mixture Regression for Big Data, *Journal of Data Science*.
7. Jie Chen, Yiming Zhang, Zahra Amirkhanzadeh Barandouzi, **Joochul Lee**, Tingting Zhao, Wanli Xu, Ming-Hui Chen, Bin Feng, Angela Starkweather, Xiaomei Cong (2022). The effect of self-management online modules plus nurse-led support on pain and quality of life among young adults with irritable bowel syndrome: a randomized controlled trial. *International Journal of Nursing Studies*.
8. Tingting Zhao, Yiming Zhang, **Joochul Lee**, Angela Starkweather, Erin Young, Xiaomei Cong (2022). 'The Associations of Single Nucleotide Polymorphisms with Risk and Symptoms of Irritable Bowel Syndrome', *Journal of Personalized Medicine*.
9. Zahra A. Barandouzi, **Joochul Lee**, Maria del Carmen Rosas, Jie Chen, Wendy A. Henderson, Angela R. Starkweather, Xiaomei S. Cong (2022), Associations of Neurotransmitters and the Gut Microbiome with Emotional Distress in Irritable Bowel Syndrome, *Scientific Reports*.
10. Jie Chen, Zahra Amirkhanzadeh Barandouzi, **JooChul Lee**, Wanli Xu, Bin Feng, Angela Starkweather, Xiaomei Cong (2021). 'Psychosocial and Sensory Factors Contribute to Self-Reported Pain and Quality of Life in Young Adults with Irritable Bowel Syndrome', *Pain Management Nursing*.
11. **JooChul Lee**, Elizabeth Schifano, Haiying Wang (2021). 'Fast Optimal Subsampling Probability Approximation for Generalized Linear Models'. *Econometrics and Statistics*.
12. Zahra A. Barandouzi, **Joochul Lee**, Kendra Maas, Angela R. Starkweather, Xiaomei S. Cong (2021), Altered gut microbiota in irritable bowel syndrome and its association with nutrients intake. *Journal of Personalized Medicine*.
13. Kyounghae Kim, Yiming Zhang, **Joochul Lee**, Thomas P. Meehan, Kim Gans (2020). 'Age at HPV vaccine initiation among adolescents and young adults from 22 pediatric practices in the northeastern United States'. *Journal of Nursing Scholarship*.
14. **JooChul Lee**, Haiying Wang, Elizabeth Schifano (2020). 'Online Updating Method to Correct for Measurement Error in Big Data Streams'. *Computational Statistics and Data Analysis*.

## Manuscripts

1. Tzu-Ying Chen, **JooChul Lee**. Evaluating NFL Quarterback Execution Using In-Flight Passing Windows and Execution Stability. 2026, *Under review*.
2. **JooChul Lee**, Huan He. Patient-Conditioned Confidence Calibration for Chest X-ray Object Detection 2026, *Under review*.
3. **JooChul Lee**, Ziyang Wang, Huan He. Minimum Active Sample Conditional Objective Estimator for Efficient Active Learning. 2025, *Under review*.

## Teaching

### Instructor

| | |
|---|---|
| Experimental Statistics I | *Fall 2024, 2025, Spring 2026* |
| Statistics for Engineers and Scientists | *Spring 2025* |
| Introduction to Mathematical Statistics II | *Spring 2021* |
| Introduction to Mathematical Statistics I | *Fall 2020* |

## Presentations

### Invited Talks

1. 'A Pipeline for Sports Analytics: From Video to Performance Metrics', American Statistical Association (ASA) San Antonio Chapter Virtual Seminar, 2025
2. 'HiAL: High-dimensional Inference under Active Learning', WNAR/IMS Annual Meeting 2025 (Organizer).
3. 'HiAL: High-dimensional Inference under Active Learning', ENAR Spring Meeting 2025
4. 'Towards optimal sample-efficient model evaluation', New England Statistics Symposium 2024 (Organizer).
5. 'Towards optimal sample-efficient model evaluation', Department of Mathematical Sciences, University of Memphis, Feb. 2024.
6. 'Towards optimal sample-efficient model evaluation', Mathematical Sciences Department, Worcester Polytechnic Institute, Jan. 2024.
7. 'Towards optimal sample-efficient model evaluation', Department of Mathematics and Statistics, University of Vermont, Jan. 2024.
8. 'Towards optimal sample-efficient model evaluation', Department of Mathematics and Statistics, Auburn University, Jan. 2024.
9. 'Towards optimal sample-efficient model evaluation', Mathematics and Statistics Department, Missouri University of Science and Technology, Dec. 2023.
10. 'Model Calibration and Evaluation via Optimal Sampling in Electronic Health Record Data', ICSA Applied Statistics Symposium 2023.
11. 'Model Calibration and Evaluation via Optimal Sampling in Electronic Health Record Data', International Conference on Design of Experiments 2023.
12. 'Sampling-based Gaussian Mixture Regression for Big Data', New England Statistics Symposium 2021 (Organizer).

## Services

### Referee Service

- Journals - Computational Statistics and Data Analysis, IEEE Transactions on Knowledge and Data Engineering, Journal of Applied Statistics, Journal of Machine Learning Research, Scientific Reports, Statistics and Computing, Statistics and Its Interface, Statistica Sinica, Test, AISTAT
- Fundings - UK Research and Innovation

### Conference Session Chair/Organizer

- 'Intelligent Risk Prediction via Causal and Efficient Modeling in Electronic Health Records', the 33rd International Biometric Conference 2026 (Organizer).

## Honors/Awards

- Recognized as a meritorious finalist in Open Philanthropy's AI Safety research grant program (2025), reaching the randomized allocation stage.
- Spring 2020 Service Award, Department of Statistics, University of Connecticut
- 2020 Student & Early Career Funding Award, 2020 Symposium on Data Science & Statistics
- Spring 2020 Doctoral Student Travel Fellowship, Department of Statistics, University of Connecticut
- Spring 2020 Doctoral Dissertation Fellowship, Department of Statistics, University of Connecticut
- 2019 KAUPA Gold scholarship, Korean American University Professors Association
    - Award for Potential in Academic Excellence as a Scholar
- 2019 Travel award, International Conference on Statistical Distributions and Applications 2019
- 2019, 2020 Pre-doctoral Fellowship, Department of Statistics, University of Connecticut
- Graduated first, Department of Statistics, Dongguk University
- Magna cum laude, Dongguk University
- Scholarships for honor students, Dongguk University
- Scholarships for 2010, 2012 at Gwang Ju Citizen Scholarship Foundation, South Korea

## Skills

**Programming** R, Python, SAS, SPSS

## Additional Information

- President, Korean Graduate Student Association, University of Connecticut, 2018
- President of Soccor Club, Dongguk University, 2011