



Sima Najafzadehkhoei

Department of Population Health Sciences, University of Utah

- **Place of Birth:** Tehran, Iran
- **Date of Birth:** November 4, 1997
- **Current Status:** Second-year Ph.D. student in PHS Biostatistics at the UoU
- **Address:** University of Utah, Salt Lake City, UT, USA
- **Phone:** +1 (801) 301 2749
- **Email:** sima.njf@utah.edu



About Myself

I am a biostatistics Ph.D. student focused on building reliable, data-driven tools for infectious disease modeling and public health decision-making. My current work centers on (i) **machine-learning-based calibration for agent-based epidemic models**—developing a three-layer bidirectional LSTM calibrator that improves accuracy and runtime relative to ABC and is being productized as `epiworldRcalibrate`—and (ii) **comparative evaluation of effective reproduction number (R_t) estimators** using agent-based network models. I collaborate closely with **Dr. Bernardo Modenesi**, **Dr. Yue Zhang**, and **Dr. George Vega Yon**, contribute to open-source tools (e.g., `epiworldR`), and manage complex EHR/VA datasets with reproducible workflows.

Beyond methods development, I **develop R packages** with an emphasis on clear APIs, documentation, and reproducibility; current efforts include `epiworldRcalibrate` for practical calibration of epidemic ABMs and `imaginarycss` for Cognitive Social Structures (software paper in preparation, Sept. 2025). I share results with both technical and applied audiences (CDC 2024; ENAR 2025; JSM 2025 selection), aiming to translate rigorous statistics and machine learning into actionable insights for public health.

Education

- Expected 2027 **Ph.D. in PHS Biostatistics**, *University of Utah*, Salt Lake City, UT, USA
 - GPA: 4.0/4.0
 - **Relevant Coursework:** Multilevel Data Analysis, Statistical Practice, Advanced R Programming, EHR Data Analysis, Advanced Statistical Inference, Survival Analysis
- July 2020 **B.A. in Statistics and Applications**, *University of Tehran*, Tehran, Iran
 - GPA: 3.7/4.0
 - **Key Coursework:** Mathematical Statistics, Linear Algebra, Stochastic Processes, Data Mining, Time Series Analysis, Computational Statistics

Research Interests

Application of advanced statistical methodologies in biostatistics, encompassing survival analysis, causal inference, and longitudinal studies, integrated with machine learning techniques for comprehensive healthcare data analysis.

Research Experience

- Ongoing **Research Assistant to Dr. Yue Zhang**, *University of Utah*, Salt Lake City, UT, USA
- **Project:** Comparative Analysis of R_t Estimation Methods Using Agent-Based Network Models.
 - Collaborating on CDC-related initiatives and SDBC projects.
 - Managing and analyzing VA datasets for epidemiological research.
- Ongoing **Research Assistant to Dr. George Vega Yon**, *University of Utah*, Salt Lake City, UT, USA
- **Project:** Automated Calibration of Agent-Based Models Utilizing Deep Neural Networks.
 - Developing machine learning algorithms to enhance the precision and efficiency of epidemiological models.
 - **CRAN Package Development:** Integrating automatic calibration tools into the `epiworldR` R package for broader accessibility.

Publications, Software, and Presentations

Preprints

- 2025 **Machine Generalize Learning in Agent-Based Models: Going Beyond Surrogate Models for Calibration in ABMs**, *arXiv preprint*
Authors: S. Najafzadehkhoei, G. Vega Yon, B. Modenesi, D. S. Meyer
- Introduces a supervised machine learning calibrator for epidemic agent-based models using a three-layer bidirectional LSTM.
 - Demonstrates improved accuracy and computational efficiency over Approximate Bayesian Computation (ABC).
 - Includes an open-source implementation in R via `epiworldR`.
- July 9, 2025 **A novel approach for classifying Monoamine Neurotransmitters by applying Machine Learning on UV plasmonic-engineered Auto Fluorescence Time Decay Series (AFTDS)**, *arXiv preprint arXiv:2507.07227*
Authors: M. Mohammadi, S. Najafzadehkhoei, G. Vega Yon, Y. Wang
- Integrates aluminum concave nanocubes (AICNCs) with ML to enhance native fluorescence for probe-free, label-free detection of DA, NE, and DOPAC.
 - Reports fluorescence enhancement up to $12\times$ (DA), $9\times$ (NE), and $7\times$ (DOPAC) over silicon substrates.
 - Uses LSTM models on time-dependent fluorescence data; outperforms KNN and Random Forest; achieves $>89\%$ classification accuracy.

Manuscripts in Preparation

- September 15, 2025 **imaginarycss: An R package for Cognitive Social Structures**, *Manuscript in preparation*
Authors: S. Najafzadehkhoei, G. Vega Yon, K. Tanaka (Affiliations: University of Utah; Aarhus University)
- Provides a CSS-native data model, perceptual error taxonomy, tie-level accuracy decomposition, and accuracy-preserving null models.
 - Enables reproducible analysis of discrepancies between perceived and ground-truth networks across sociology, psychology, and organizational science.

Software Contributions

In `epiworldRcalibrate`: **An R package for Machine Generalize Learning in Agent-Based Models**, CRAN/JOSS submission in preparation

Authors: **S. Najafzadehkhoie**, G. Vega Yon, B. Modenesi, D. S. Meyer

- Practical tools for calibration of agent-based epidemic models, extending methods from the 2025 arXiv preprint.

In `imaginarycss`: **An R package for Cognitive Social Structures**, Software package (R)

Authors: **S. Najafzadehkhoie**, G. Vega Yon, K. Tanaka

- CSS-native data structure, perceptual error classification, tie-level accuracy, and accuracy-based null models; complements `statnet` and `igraph`.

Conference Presentations

Oral Presentations

- Najafzadehkhoie, S. (2025). *Automatic Calibration of Agent-Based Models using Deep Neural Networks*. ENAR Spring Meeting, New Orleans, LA. March 26, 2025.
- Najafzadehkhoie, S. (2025). *Automatic Calibration of Agent-Based Models using Deep Neural Networks*. Joint Statistical Meetings (JSM), Nashville, TN. Accepted for oral presentation, August 2025 (unable to attend). *Competitive selection at one of the largest international statistical conferences*.
- Najafzadehkhoie, S. (2024). *Comparing the Performance of Different R_t Estimations: Insights from an Agent-Based Network Model Study*. Centers for Disease Control and Prevention (CDC), Atlanta, GA. July 11–12, 2024.

Poster Presentations

- Najafzadehkhoie, S. (2025). *Automatic Calibration of Agent-Based Models using Deep Neural Networks*. DELPHI Data Science Initiative. January 2025.
- Najafzadehkhoie, S. (2025). *Automatic Calibration of Agent-Based Models using Deep Neural Networks*. Joint Statistical Meetings (JSM), Nashville, TN. Accepted for poster presentation, August 2025 (unable to attend).

Achievements

- 2025 **Oral Presentation** at ENAR Spring Meeting, New Orleans, LA (March 26).
- 2025 **JSM 2025 Oral & Poster Selections** (unable to attend); competitive international conference.
- 2025 **Poster Presentation** at DELPHI Data Science Initiative (January).
- 2024 **Oral Presentation** at the Centers for Disease Control and Prevention (CDC), Atlanta, GA (July 11–12).
- 2023 **Awarded a Fully Funded Ph.D. Position in Biostatistics** at the University of Utah.
- 2020 **Ranked First** in the Bachelor's Degree Program at the University of Tehran.

Skills & Interests

Technical Skills

Programming Languages R, Python, Java, \LaTeX , Stata

Machine Learning & Data TensorFlow, scikit-learn, Pandas, NumPy

Reproducibility R package development, Git/GitHub, \LaTeX

Languages

English Fluent, TOEFL Score: 97 (Reading: 19, Listening: 27, Writing: 26, Speaking: 25)

Persian Native

Interests

Hiking, Skiing, Programming