

Dhrubajyoti Ghosh

Assistant Professor, Department of Data Science and Analytics, Kennesaw State University
dghosh3@kennesaw.edu | Website | Google Scholar

Research Interests: Longitudinal data analysis; causal inference; high-dimensional and network-based methods; time series; statistical methods for neurodegenerative disease and biomedical imaging.

Academic Appointments

2025–present	Assistant Professor , Department of Data Science and Analytics, Kennesaw State University
2022–2025	Postdoctoral Research Associate , Department of Biostatistics and Bioinformatics, Duke University. Advisor: Sheng Luo
2018–2022	Research Assistant , Laboratory for Analytic Sciences. Advisor: Soumendra Lahiri
2019–2022	Teaching Assistant , Department of Mathematics and Statistics, Washington University in St. Louis
2017–2019	Teaching Assistant , Department of Statistics, North Carolina State University

Education

2017–2022	Ph.D., Statistics , Washington University in St. Louis Thesis: Contribution to Data Science: Time Series, Uncertainty Quantification and Applications Advisors: Soumendra Lahiri and Tucker McElroy <i>Transferred from North Carolina State University with advisor in Fall 2019.</i>
2015–2017	M.Stat., Statistics , Indian Statistical Institute, Kolkata, India
2012–2015	B.Stat., Statistics , Indian Statistical Institute, Kolkata, India

Publications

[†] Student or trainee supervised/co-supervised by D. Ghosh.

Peer-Reviewed and Accepted Publications

1. [†]Ibenye, C., [†]Jimenez, A.-V., [†]Jaeger, S. & **Ghosh, D.** When Machines Lie Differently: Detecting AI vs Human Fake News. *arXiv preprint arXiv:2607.21967*. (2026). Accepted at ICMLA Conference.
2. **Ghosh, D.** A Rank-Based Information Fusion Framework for Comparing Clustered Multivariate Socioeconomic Outcomes. In *2026 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* 1–5 (2026).
3. **Ghosh, D.**, Boettcher, W., Johnston, R. & Lahiri, S. Machine Learning Based Bot Detection on X With Temporal and Semantic Feature Integration. *IEEE Transactions on Computational Social Systems*. (2026).
4. **Ghosh, D.**, McElroy, T. & Lahiri, S. Polyspectral Mean Estimation of General Nonlinear Processes. *Bernoulli*. (2026). Accepted for publication in Bernoulli, in press.

5. **Ghosh, D.** & Pal, S. Efficacy Analysis in Clinical Trials: A Comprehensive Review of Statistical and Machine Learning Approaches. *The New England Journal of Statistics in Data Science*, 1–23. ISSN: 2693-7166. (2026).
6. **Ghosh, D.**, Tushar, F., Dahal, L., Vancoillie, L., Lafata, K. J., Samei, E., Lo, J. Y. & Luo, S. Demographic distribution matching between real-world and virtual phantom population. *Medical physics* **53**, e70364. (2026).
7. McElroy, T., **Ghosh, D.** & Lahiri, S. Polyspectral factorization. *Annals of the Institute of Statistical Mathematics*, 1–25. (2026).
8. Peddi[†], P. K. & **Ghosh, D.** Integrating Causal Inference with Graph Neural Networks for Alzheimer’s Disease Analysis. In *Proceedings of the 15th International Conference on Data Science, Technology and Applications - Volume 1: DATA* 450–457 (SciTePress, 2026). ISBN: 978-989-758-854-9.
9. Dahal, L., Ghoghnejad, M., Vancoillie, L., **Ghosh, D.**, Bhandari, Y., Kim, D., Ho, F. C., Tushar, F. I., Luo, S., Lafata, K. J., *et al.* XCAT 3.0: A comprehensive library of personalized digital twins derived from CT scans. *Medical image analysis* **103**, 103636. (2025).
10. **Ghosh, D.** Assessing air quality extremes: a comparative extreme value analysis of metropolitan cities across India and the world. *Environmental Monitoring and Assessment* **197**, 276. (2025).
11. **Ghosh, D.** Polyspectral mean based time series clustering of Indian stock market. *Discover Data* **3**, 10. (2025).
12. **Ghosh, D.**, Boettcher, W. A., Johnston, R. & Lahiri, S. THANOS: A Predictive Model of Electoral Campaigns Using Twitter Data and Opinion Polls. *Data Science in Science* **4**, 2484180. (2025).
13. **Ghosh, D.** & Luo, S. A non-parametric U-statistic testing approach for multi-arm clinical trials with multivariate longitudinal data. *Journal of Multivariate Analysis*, 105447. (2025).
14. **Ghosh, D.**, Pal, S., Lutz, M., Luo, S. & Initiative, A. D. N. Ensemble survival analysis for preclinical cognitive decline prediction in Alzheimer’s disease using longitudinal biomarkers. *Journal of Alzheimer’s Disease*, 13872877251365621. (2025).
15. **Ghosh, D.**, Xu, X., Luo, S. & Database, C. I. P. Power and sample size calculation for multivariate longitudinal trials using the longitudinal rank sum test. *Statistics in Medicine* **44**, e70261. (2025).
16. Mayan, I., Roth, H., **Ghosh, D.**, Whitson, H. E. & Johnson, K. G. Genetic and biomarker disclosure process in a memory and aging study. *Journal of Alzheimer’s Disease* **0**. PMID: 39924880, 13872877251315957. (2025).
17. Pal, S., **Ghosh, D.** & Yang, S. Penalized FCI for Causal Structure Learning in a Sparse DAG for Biomarker Discovery in Parkinson’s Disease. *arXiv preprint arXiv:2507.00173*. (2025). Accepted at Annals of Applied Statistics, in press.
18. Tushar, F. I., Vancoillie, L., McCabe, C., Kavuri, A., Dahal, L., Harrawood, B., Fryling, M., Zarei, M., Sotoudeh-Paima, S., Ho, F. C., *et al.* Virtual lung screening trial (VLST): An in silico study inspired by the national lung screening trial for lung cancer detection. *Medical Image Analysis* **103**, 103576. (2025).
19. Xu, X., **Ghosh, D.**, Luo, S. & Database, C. I. P. A novel longitudinal rank-sum test for multiple primary endpoints in clinical trials: Applications to neurodegenerative disorders. *Statistics in Biopharmaceutical Research*, 1–17. (2025).
20. McElroy, T. S., **Ghosh, D.** & Lahiri, S. Quadratic Prediction of Time Series via Auto-Cumulants. *Sankhya A* **86**, 431–463. (2024).
21. Sotoudeh-Paima, S., Segars, W. P., **Ghosh, D.**, Luo, S., Samei, E. & Abadi, E. A systematic assessment and optimization of photon-counting CT for lung density quantifications. *Medical Physics* **51**, 2893–2904. (2024).

22. Tushar, F. I., Vancoillie, L., McCabe, C., Kavuri, A., Dahal, L., Harrawood, B., Fryling, M., Zarei, M., Sotoudeh-Paima, S., Ho, F. C., *et al.* Virtual NLST: towards replicating national lung screening trial. In *Medical Imaging 2024: Physics of Medical Imaging* **12925**, 442–447 (2024).

Manuscripts Under Review / Revision

1. Pal, S. & **Ghosh, D.** Network-aware IV Regression for Causal Node Discovery and Estimation. *arXiv preprint arXiv:2604.24969*. (2026). Major Revision at Journal of the American Statistical Association.
2. Ghosal, R., **Ghosh, D.**, Yue, Y. & Roy, A. Penalized Cox Regression with Graph-Structured Distribution-Valued Predictors (2026). Submitted to Journal of Computational and Graphical Statistics.
3. **Ghosh, D.** Spatiotemporal Characterization of Overdose Mortality in Georgia, USA Using Spectral and Nonlinear Interaction Analysis, 2003-2021. *arXiv preprint arXiv:2604.00353*. (2026). Third Round Revision at Spatial and Spatio-temporal Epidemiology.
4. **Ghosh, D.**, Song, Z., Xiao, L. & Luo, S. A comprehensive goodness-of-fit test for covariance structures in multivariate longitudinal data. (2026). Submitted to Biometrics.
5. [†]Jaeger, S., [†]Ibeneye, C., [†]Vera-Jimenez, A. & **Ghosh, D.** Human vs. Machine Deception: Distinguishing AI-Generated and Human-Written Fake News Using Ensemble Learning. *arXiv preprint arXiv:2604.09960*. (2026). Submitted to American Journal of Undergraduate Research.
6. Karami, A., Imran, M., **Ghosh, D.** & Carlisle, J. Predicting Areas at the Risk of Lung Cancer Based on Modifiable Risk Factors (2026). Revised manuscript under review at JCO Clinical Cancer Informatics.

Working Papers

1. [†]Shahid, I., Luo, S. & **Ghosh, D.** Longitudinal Rank-Sum Testing for incomplete clinical trials (2026). To be submitted.
2. [†]Vera-Jimenez, A., [†]Jaeger, S., [†]Ibenye, C. & **Ghosh, D.** Cross-Prompt Generalization in Detecting AI-Generated Fake News Using Interpretable Linguistic Features. (2026). To be submitted.
3. **Ghosh, D.** & Lahiri, S. Subsampling and Convolved Bootstrap Inference for Polyspectral Means in Nonlinear Time Series (2026). To be submitted.
4. **Ghosh, D.** Nonparametric Global Tests for Longitudinal Treatment Effects Using Integrated Mean Curves (2026). To be submitted.
5. **Ghosh, D.** & Chakraborty, N. Geographic and Socioeconomic Inequality in High School Graduation Success Across the United States (2026). To be submitted.
6. **Ghosh, D.**, Dahal, L., Vancoillie, L., Nejad, M. G., Samei, E., Luo, S. & Lo, J. SADDHU: Segmentation Anomaly Detection using Distribution based Holistic Unsupervised approach (2026). To be submitted to Medical Physics.
7. **Ghosh, D.** & Luo, S. Residual LRST: Covariate-Adjusted Rank Inference for Longitudinal Data (2026). To be submitted.
8. **Ghosh, D.**, Luo, S., Arbeev, K., Akushevich, I. & Yashin, A. A Holistic Assessment of Psychological and Cognitive Health in Aging Using Long Life Family Study Data (2026). To be submitted.
9. **Ghosh, D.** & Pal, S. Representation Learning for Causal Inference in Multimodal Repeated Cross-Sectional Surveys with Block-missingness (2026).

10. Olszewska, D. A., **Ghosh, D.**, Dam, T., Simuni, T., Luo, S. & Marras, C. Performance of the Neuronal α -Synuclein Disease Integrated Staging System Framework for Parkinson's Disease in the Critical Path for Parkinson's Integrated Parkinson's Disease Datasets (2026). To be submitted.

Research Funding

External Funding - Pending

- NIH R15** **Principal Investigator.** 1R15AG104246, *Network-aware Mendelian randomization for causal brain imaging in AD/PD*. Revision submitted; pending IRG review.
- NIH R21** **Multiple Principal Investigator.** 1R21AG108860, *Longitudinal brain-intensity distributions predicting time to AD* (with Rahul Ghosal, University of South Carolina). Pending IRG review.

Institutional Research Support

- 2025 Kennesaw State University Startup Fund (\$250,000; PI)
- 2026, 2027 First-Year Scholars Program Funding, Kennesaw State University (\$3,000 each)
- 2027 Sophomore Scholars Program Funding, Kennesaw State University (\$2,000)

Research Software

- pfc**i: CRAN R package (with Samhita Pal and Shu Yang) for penalized FCI causal structure learning in sparse DAGs.
- ivgl**s: CRAN R package (with Samhita Pal) for network-aware instrumental variable regression.
- Polyspectra**: R package for polyspectra, polyspectral means, and asymptotic variance estimation.
- QuadraticPredictionR**: R package for quadratic forecasting of time series.

Student Mentoring

Doctoral Students

- 2026–present Imaan Shahid, Ph.D.
- 2026–present Param V. Kesireddy, Ph.D.
- 2025–present Faruk Muritala, Ph.D. (co-advisor: Austin Brown)

Master's Students

- 2025–2026 Imaan Shahid, M.S. (Data Science and Analytics)
- 2025–2026 Pranay Kumar Peddi, M.S. (Computer Science)
- 2026–2026 Hemanth Reddy Mandla, M.S. (Data Science and Analytics)
- 2025–2025 Shelby Bradley, M.S. (Data Science and Analytics)

Master's Thesis Committee Service

- 2026 Md Istiak Morsalin, M.S. (Advisor: Amir Karami)

2026 Louis Muhammad, M.S. (Advisor: Md Shazibul Islam Shamim)
 2026 Abdullah Al Noman, M.S. (Advisor: Mohammad Abdul Halim)

Undergraduate Researchers

2025–present Aya-Vera Jimenez; Samuel Jaeger; Calvin Ibeneye

Presentations

Invited Talks

1. **Ghosh, D.**, McElroy, T. & Lahiri, S. *Polyspectral mean estimation of general nonlinear processes* International Indian Statistical Association (IISA) Conference, Invited Session. Dec. 2026. Scheduled.
2. **Ghosh, D.** & Pal, S. *Network-aware IV Regression for Causal Node Discovery and Estimation* Joint Statistical Meetings (JSM), Invited Session. Aug. 2026.
3. **Ghosh, D.**, Boettcher, W., Johnston, R. & Lahiri, S. *THANOS: A Predictive Model of Electoral Campaigns Using Twitter Data and Opinion Polls* Joint Statistical Meetings (JSM), Invited Session. Aug. 2025.
4. **Ghosh, D.** & Luo, S. *A Novel Longitudinal Rank-Sum Test for Multiple Primary Endpoints in Clinical Trials: Applications to Neurodegenerative Disorders* International Indian Statistical Association (IISA) Conference. Dec. 2024.
5. **Ghosh, D.** & Luo, S. *A Comprehensive Goodness-of-Fit Test for Covariance Structures in Multivariate Longitudinal Data* International Indian Statistical Association (IISA) Conference. May 2023.
6. **Ghosh, D.**, McElroy, T. & Lahiri, S. *Polyspectral Mean Estimation of General Nonlinear Processes* Joint Statistical Meetings (JSM). Aug. 2021.
7. **Ghosh, D.**, Boettcher, W., Johnston, R. & Lahiri, S. *Analyzing Influence of Social Media Through Twitter* Laboratory for Analytic Sciences (LAS) Symposium. Oct. 2019.

Conference Presentations

1. **Ghosh, D.** *A Rank-Based Information Fusion Framework for Comparing Clustered Multivariate Socioeconomic Outcomes* IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). 2026.
2. **Ghosh, D.**, Xu, X. & Luo, S. *The Longitudinal Rank-Sum Test (LRST)* ENAR (Eastern North American Region). Mar. 2026.
3. Peddi, P. K. & **Ghosh, D.** *Integrating Causal Inference with Graph Neural Networks for Alzheimer's Disease Analysis* International Conference on Data Science, Technology and Applications (DATA). 2026.
4. **Ghosh, D.**, Xu, X. & Luo, S. *The Longitudinal Rank-Sum Test (LRST)* 19th International Joint Conference on Computational and Methodological Statistics. Dec. 2025.
5. **Ghosh, D.** & Lahiri, S. *Rate of Convergence and Optimal Choice of m in the m out of n Bootstrap for Sample Extremes* International Indian Statistical Association (IISA) Conference, Student Paper Competition. June 2019.

Posters

1. **Ghosh, D.** & Luo, S. *The Longitudinal Rank-Sum Test (LRST) Framework: A Unified Rank-Based Approach for Testing Treatment Effects in Multivariate Longitudinal Studies* Southeast Regional Clinical & Translational Science Conference. Mar. 2026.
2. Ibenye, C., Jimenez, A.-V., Jaeger, S. & **Ghosh, D.** *When Machines Lie Differently: Detecting AI vs Human Fake News* IEEE International Conference on Machine Learning and Applications (ICMLA). Oct. 2026. Scheduled.
3. **Ghosh, D.** & Lahiri, S. *Analyzing Influence of Social Media Through Twitter* Symposium on Data Science and Statistics (SDSS), Cambridge, MA. June 2019.
4. **Ghosh, D.** *Application of Projection to RESCAL Algorithm* Southern Regional Council on Statistics (SRCOS). July 2018.

Teaching

Kennesaw State University

2026	Causal Inference (STAT 7990), Instructor and course developer
2026	Time Series Forecasting (STAT 8332), Instructor and course developer
2025	Introduction to Categorical Data Analysis (STAT 8310), Instructor
2025	Introduction to Probability and Statistics (STAT 2332), Instructor

Washington University in St. Louis

2022	Theory of Statistics II (Math 5062), Assistant Instructor
2020–2021	Theory of Statistics (Math 5061), Teaching Assistant
2020	Topics in Statistics: Machine Learning (Math 579), Teaching Assistant
2019	Elementary and Intermediate Statistics and Data Analysis (Math 3200), Teaching Assistant

North Carolina State University

2017–2018	Probability and Statistics for Engineers (ST 370), Teaching Assistant
-----------	---

Professional Service

Conference Organization

2026	Organizer, invited session, Joint Statistical Meetings (JSM)
2025	Chair and Organizer, invited session, CMSTAT
2025	Chair, invited session, Joint Statistical Meetings (JSM)

Journal Refereeing

Biostatistics; Statistics in Medicine; New England Journal of Statistics and Data Science; Journal of Alzheimer's Disease; Journal of Statistical Planning and Inference; Journal of the Korean Statistical Society; Archives of Proteomics and Bioinformatics.

University Service

2026–present	Chair, Colloquium Committee, Kennesaw State University
2026–present	Chair, Open Rank Professor Selection Committee, Kennesaw State University

Awards and Honors

2021	Harshbarger Student Travel Award, Southern Regional Council on Statistics
2019	IISA Student Paper Competition Travel Award
2018	Harshbarger Student Travel Award, Southern Regional Council on Statistics
2015	Selected for summer research internship with Heinz Koeppel, TU Darmstadt, Germany
2012	KVPY Fellowship, India (Rank 18)
2012	Indian Institute of Technology Joint Entrance Examination (IIT-JEE), Rank 1358 (top 0.005%)

Selected Media Coverage

2022	The Ampersand, Are bots winning the war to control social media?
------	--