Yu Gu

Room 231, Run Run Shaw Building
The University of Hong Kong
Pokfulam, Hong Kong

Email: yugu@hku.hk
Phone: (852) 3917-8322
Website: https://yugu-stat.github.io/

Research Interests

Survival analysis; Non- and semi-parametric inference; Statistical machine learning & AI; Functional data analysis; Alzheimer's disease; Infectious disease; Cancer.

Employment

The University of Hong Kong — Pokfulam, Hong Kong
Assistant Professor, Department of Statistics and Actuarial Science — 2023–

Education

University of North Carolina at Chapel Hill — Chapel Hill, NC, USA
Ph.D. in Biostatistics, Gillings School of Global Public Health — 2018–2023

- Thesis: Statistical Methods for Analyzing Interval-Censored Multi-State Data
- Advisors: Drs. Danyu Lin and Donglin Zeng

University of Science and Technology of China — Hefei, Anhui, China
B.S. in Statistics, School of the Gifted Young — 2014–2018

Honors and Awards

- Barry H. Margolin Dissertation Award, Department of Biostatistics at UNC — 2024
- Travel award, ICSA 2023 China Conference — 2023
- Harry A. Guess Scholarship, Department of Biostatistics at UNC — 2021
- Special Commendation for Doctoral Qualifying Exams, Department of Biostatistics at UNC — 2019
- Outstanding Graduate Award, University of Science and Technology of China — 2018
- Outstanding Student Award, University of Science and Technology of China — 2014–2017

Teaching Experience

- **Instructor** — Spring 2024–2026
  [STAT3655] Survival Analysis
- **Instructor** — Spring 2025–2026
  [STAT6018] Research Frontiers in Data Science

- **Project Supervisor** 2024–2026
  [STAT3799] Directed Studies in Statistics
- **Project Supervisor** 2025
  [STAT4710] Capstone Experience for Statistics Undergraduates
- **Project Supervisor** 2024
  [ARIN7600] Artificial Intelligence Project

Research Grants

RGC–GRF 17308126 2027–2029
Dynamic prediction of Alzheimer's disease progression using multi-state and imaging data
- Role: Principal Investigator
- Amount: HKD 616,160

RGC–ECS 27303624 2025–2027
Joint analysis of interval-censored multi-state data, longitudinal data, and competing risks data
- Role: Principal Investigator
- Amount: HKD 831,657

Publications

[*corresponding author] [#equal contribution] [student author under my supervision]

Refereed Journal Publications

1. Ye, Z., Wang, Y., **Gu, Y.**, & Kang, K. (2026). Beyond Overall Diagnosis: A Longitudinal Multi-Domain Latent Structure Model for Domain-specific Parkinson's Disease Progression. ***Journal of the American Statistical Association***, (just-accepted).

2. Lyu, R., Jiang, J., Lam, K. F., & **Gu, Y.*** (2026). Evaluating Time-Varying and Heterogeneous Efficacy of COVID-19 Antiviral Drugs. ***Statistics in Biopharmaceutical Research***, (just-accepted), 1-13.

3. Fang, J.#, **Gu, Y.#**, & Yin, G. (2026). Pseudo-death: a new measure to evaluate composite outcomes in survival analysis. ***Journal of the Royal Statistical Society Series C: Applied Statistics***, qlag030.

4. **Gu, Y.***, Zeng, D., & Lin, D. Y. (2025). Semiparametric Regression Analysis of Interval-Censored Multi-State Data with An Absorbing State. ***Journal of the American Statistical Association***, 1-21.

5. Shrestha, P., Graff, M., **Gu, Y.**, Wang, Y., Avery, C. L., Ginnis, J., Simancas-Pallares, M. A., Ferreira Zandoná, A. G., Alotaibi, R. N., Orlova, K., Ahn, H. S., Nguyen, K. N., Highland, H. M., Lin, D. Y., Preisser, J. S., Slade, G. D., Marazita, M. L., North, K. E., & Divaris, K. (2025). Multi-ancestry Genome-Wide Association Study of Early Childhood Caries. ***Journal of Dental Research***, 104(3), 280-289.

6. Lin, D. Y., Xu, Y., **Gu, Y.**, Sunny, S. K., Moore, Z., & Zeng, D. (2024). Impact of booster vaccination interval on SARS-CoV-2 infection, hospitalization, and death. ***International Journal of Infectious Diseases***, 145, 107084.

7. Lin, D. Y., Wang, J., **Gu, Y.**, & Zeng, D. (2024). Evaluating treatment efficacy in hospitalized COVID-19 patients, with applications to Adaptive COVID-19 Treatment Trials. ***Clinical Trials***, 21(4), 500-506.

8. **Gu, Y.**, Zeng, D., Heiss, G., & Lin, D. Y. (2024). Maximum Likelihood Estimation for Semiparametric Regression Models with Interval-Censored Multi-State Data. ***Biometrika***, 111(3), 971-988.

9. Lin, D. Y.[#], Xu, Y.[#], **Gu, Y.**[#], Zeng, D., Wheeler, B., Young, H., Moore, Z., & Sunny, S. K. (2023). Effects of COVID-19 vaccination and previous SARS-CoV-2 infection on omicron infection and severe outcomes in children under 12 years of age in the USA: an observational cohort study. ***The Lancet Infectious Diseases***, 23(11), 1257-1265.

10. Lin, D. Y., Xu, Y., **Gu, Y.**, Zeng, D., Sunny, S. K., & Moore, Z. (2023). Durability of bivalent boosters against Omicron subvariants. ***New England Journal of Medicine***, 388(19), 1818-1820.

11. Lin, D. Y., Xu, Y., **Gu, Y.**, Zeng, D., Wheeler, B., Young, H., Sunny, S. K., & Moore, Z. (2023). Effectiveness of bivalent boosters against severe Omicron infection. ***New England Journal of Medicine***, 388(8), 764-766.

12. **Gu, Y.***, Preisser, J. S., Zeng, D., Shrestha, P., Shah, M., Simancas-Pallares, M. A., Ginnis, J., & Divaris, K. (2022). Partitioning around medoids clustering and random forest classification for GIS-informed imputation of fluoride concentration data. ***The Annals of Applied Statistics***, 16(1), 551-572.

13. Lin, D. Y., **Gu, Y.**, Wheeler, B., Young, H., Holloway, S., Sunny, S. K., Moore, Z., & Zeng, D. (2022). Effectiveness of Covid-19 vaccines over a 9-month period in North Carolina. ***New England Journal of Medicine***, 386(10), 933-941.

14. Lin, D. Y., **Gu, Y.**, Xu, Y., Wheeler, B., Young, H., Sunny, S. K., Moore, Z., & Zeng, D. (2022). Association of primary and booster vaccination and prior infection with SARS-CoV-2 infection and severe COVID-19 outcomes. ***JAMA***, 328(14), 1415–1426.

15. Lin, D. Y., **Gu, Y.**, Xu, Y., Zeng, D., Wheeler, B., Young, H., Sunny, S. K., & Moore, Z. (2022). Effects of vaccination and previous infection on Omicron infections in children. ***New England Journal of Medicine***, 387(12), 1141-1143.

16. Lin, D. Y., **Gu, Y.**, Zeng, D., Janes, H. E., & Gilbert, P. B. (2022). Evaluating vaccine efficacy against severe acute respiratory syndrome coronavirus 2 infection. ***Clinical Infectious Diseases***, 74(3), 544-552.

17. Lin, D. Y., Zeng, D., **Gu, Y.**, Krause, P. R., & Fleming, T. R. (2022). Reliably Assessing duration of protection for coronavirus disease 2019 vaccines. ***The Journal of Infectious Diseases***, 226(11), 1863-1866.

18. Divaris, K., Slade, G. D., Ferreira Zandona, A. G., Preisser, J. S., Ginnis, J., Simancas-Pallares, M. A., Agler, C. S., Shrestha, P., Karhade, D. S., Ribeiro, A. de A., Cho, H., **Gu, Y.**, Meyer, B. D., Joshi, A. R., Azcarate-Peril, M. A., Basta, P. V., Wu, D., & North, K. E. (2020). Cohort profile: ZOE 2.0—a community-based genetic epidemiologic study of early childhood oral health. ***International Journal of Environmental Research and Public Health***, 17(21), 8056.

Book Chapters

1. **Gu, Y.***, Divaris, K., & Gansky, S. (2026+). Imputation of Missing Data. In K. Divaris (Ed.). *Oral and Craniofacial Health Research: Principles, Methods and Applications.* Springer. In press.

Papers Under Review or In Revision

1. Hu J, **Gu Y***. SPOT-IC: Improving Prediction for Interval-Censored Data via Survival Probability Transfer.

2. Qi C, Kang K, **Gu Y***. Semiparametric Functional Multistate Modeling of Alzheimer's Disease Progression with Imaging Biomarkers.

3. **Gu Y**, Xu Y, Sang P. Varying Coefficient Model for Longitudinal Data with Informative Observation Times. Major revision requested by *The Annals of Applied Statistics*.

4. Wang Y, Deng Y, **Gu Y**, Wang Y, Zeng D. Debiased Inference for High-Dimensional Regression Models Based on Profile M-Estimation.

5. **Gu, Y.***, Zeng, D., Lin, D. Y. Prediction-oriented transfer learning for survival analysis. Major revision requested by *Biometrics*.

6. Simancas-Pallares, M. A., Gormley, A., Shrestha, P., **Gu, Y.**, Cho, H., Spangler, H. D., Burk, Z., Smith, M., Dashper, S., Burgner, D., Ferreira Zandoná, A. G., Ginnis, J., Vann, W. F., Esberg, A., Roach, J., Ribeiro, A. A., Wu, D., Silva, M. J., Holgerson, P. L., Haworth, S., Johansson, I., North, K. E., & Divaris, K. Evidence for Clinical Subtypes of Early Childhood Caries.

7. Shrestha P., Graff M., **Gu Y.**, Wang Y., Ahn H. S., Nguyen K. N., Khanna A., Avery C. L., Highland H. M., Ginnis J., Simancas-Pallares M. A., Ferreira Zandoná A. G., Alotaibi R. N., Lin D. Y., Preisser J. S., Slade G. D., Marazita M. L., North K. E., Divaris K. Genomic basis of the developmental defects of enamel and sex-specific effects.

Presentations

1. "Semiparametric Functional Model for Imaging and Interval-Censored Multistate Data", 6th Int. Workshop on Statistical Analysis on Multi-Outcome Data, Bordeaux, France, June 2026.

2. "SPOT-IC: Improving Prediction for Interval-Censored Data via Survival Probability Transfer", 2026 ICSA China Conference, Shenzhen, China, June 2026.

3. "Semiparametric Functional Model for Imaging and Interval-Censored Multistate Data", Institute of Mathematical Statistics Asia Pacific Rim Meeting 2026, Hong Kong, June 2026.

4. "Prediction-Oriented Transfer Learning for Survival Analysis", 2026 International Conference on Frontiers of Data Science, Hangzhou, China, May 2026.

5. "Prediction-Oriented Transfer Learning for Semiparametric Transformation Models with Survival Data", Joint Meetings of 2025 Taipei International Statistical Symposium and the 13th ICSA International Conference, Taipei City, Taiwan, December 2025.

6. "Prediction-Oriented Transfer Learning for Semiparametric Transformation Models with Survival Data", The 3rd Joint Conference on Statistics and Data Science, Hangzhou, China, July 2025.

7. "Prediction-Oriented Transfer Learning for Semiparametric Transformation Models with Survival Data", ICSA 2025 China Conference, Zhuhai, China, June 2025.

8. "Maximum Likelihood Estimation for Semiparametric Regression Models with Interval-Censored Multi-State Data", EcoSta 2024, Beijing, China, July 2024.

9. "Evaluating Treatment Efficacy in Hospitalized COVID-19 Patients", IMS-APRM 2024, Melbourne, Australia, January 2024.

10. "Semiparametric Regression Analysis of Interval-Censored Multi-State Data with An Absorbing State", ICSA 2023 China Conference, Chengdu, China, July 2023.

11. "Maximum Likelihood Estimation for Semiparametric Regression Models with Interval-Censored Multi-State Data", ENAR 2023 Spring Meeting, Nashville, TN, USA, March 2023.

12. "Semiparametric Regression Analysis of Interval-Censored Multi-State Data with An Absorbing State", Department of Biostatistics, Columbia University, New York, NY, USA, February 2023.

13. "Semiparametric Regression Analysis of Interval-Censored Multi-State Data with An Absorbing State", Department of Statistics, University of Pittsburgh, Pittsburgh, PA, USA, February 2023.

14. "Semiparametric Regression Analysis of Interval-Censored Multi-State Data with An Absorbing State", Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA, February 2023.

15. "Semiparametric Regression Analysis of Interval-Censored Multi-State Data with An Absorbing State", Department of Statistics, Chinese University of Hong Kong, Hong Kong, December 2022.

16. "Maximum Likelihood Estimation for Semiparametric Regression Models with Interval-Censored Multi-State Data", UNC Department of Biostatistics, Chapel Hill, NC, USA, November 2022.

17. "Durability of Covid-19 Vaccines", ICSA Applied Statistics Symposium, Gainesville, FL, USA, June 2022.

18. "Development of a Novel Imputation Method for Missing Fluoride Measurements in a Community-Based Epidemiologic Study", UNC Adams School of Dentistry, Chapel Hill, NC, USA, February 2019.

Software

- **DOVE:** Durability of Vaccine Efficacy
  - R package: https://cran.r-project.org/web/packages/DOVE/index.html
- **iDOVE:** Durability of Vaccine Efficacy Against SARS-CoV-2 Infection
  - R package: https://cran.r-project.org/web/packages/iDOVE/index.html
- **DOVE3:** Durability of Effectiveness of Vaccination and Prior Infection
  - R package: https://yugu-stat.github.io/software/dove3
- **COVID:** Assessing Totality of Evidence for COVID-19 Treatment Effects
  - R package & SAS macro: https://yugu-stat.github.io/software/covid

Professional Memberships

- American Statistical Association (ASA)
- Eastern North American Region (ENAR) of International Biometric Society
- Institute of Mathematical Statistics (IMS)
- International Chinese Statistical Association (ICSA)

Professional Services

- Journal referee: *Annals of Applied Statistics*, *Biometrics*, *Biostatistics*, *Electronic Journal of Statistics*, *Journal of Data Science*, *Journal of the American Statistical Association*, *Lifetime Data Analysis*, *Statistics in Biosciences*, *Statistics in Medicine*, *Statistical Methods in Medical Research*
- Conference service: Organizer of invited sessions at ENAR 2025 Spring Meeting and IMS-APRM 2026
- Department/Faculty service:
  - Co-organizer of HKU SAAS Departmental Seminars, 2025–2026
  - Member of Organizing Committee, HKU Summer Workshop on Statistics and Data Science, 2024
  - Member of Staff–Student Consultative Committee, 2024
  - Member of Selection Panel, HKU Undergraduate Research Colloquium, 2023

Mentorship

Supervisor

- Jingyi Hu (PhD student, 2024–)
- Chenrui Qi (PhD student, 2024–)
- Ruoxi Lyu (MPhil student, 2024–2026)

Thesis Committee Member

- Dr. Yaxian Chen (Supervisor: Prof. Eddy K.F. Lam)
- Dr. Anqi Wang (Supervisor: Prof. Yan Dora Zhang)
- Dr. Zewei Xiong (Supervisor: Prof. Pak Chung Sham)
- Dr. Daoyuan Lai (Supervisor: Prof. Yan Dora Zhang)
- Dr. Jialiang Jiang (Supervisor: Prof. Guosheng Yin)