

Research Interests

Statistical Methodology: environmental epidemiology, capture-recapture methods for disease surveillance, Bayesian methods, spatial-temporal modeling.

Applications: infectious disease, Alzheimer's disease, preterm birth.

Education

Emory University Atlanta, GA
Ph.D., Biostatistics 08/2018 - 07/2023
– Advisors: Howard H. Chang and Robert H. Lyles
– Dissertation: Statistical Methods for Disease Surveillance Based on Multiple Data Streams

Emory University Atlanta, GA
MSPH, Biostatistics 08/2016 - 05/2018

China Pharmaceutical University Nanjing, China
B.E., Biopharmaceutical Technology 09/2012 - 06/2016

Professional Experience

Division of Biostatistics College of Public Health, The Ohio State University Columbus, OH
Assistant Professor 08/2024 - Present

Department of Biostatistics and Bioinformatics, Emory University Atlanta, GA
Postdoctoral Research Fellow 08/2023 - 08/2024
– Advisors: Howard H. Chang and Robert H. Lyles
– High-dimensional mediation analysis, statistical methods for disease surveillance.

Publications

Statistical Methodology

- Zhang Y**, Ge L, Waller LA, Shah S, Lyles RH (2024). A capture-recapture modeling framework emphasizing prior information in disease surveillance. *Statistical Methods in Medical Research*, 2024;0(0), DOI: [10.1177/09622802241254217](https://doi.org/10.1177/09622802241254217).
★ This work was recognized with an ENAR Distinguished Student Paper Award for the 2024 Spring Meeting.
- Zhang Y**, Chang HH, Warren JL, Ebel ST (2024). A scalar-on-quantile function approach for estimating short-term health effects of environmental exposures. *Biometrics*, Volume 80, Issue 1, March 2024, ujae008, DOI: [10.1093/biometc/ujae008](https://doi.org/10.1093/biometc/ujae008) (Associated R package `nbRegQF` is available on [GitHub](https://github.com))
★ This work was recognized with a Young Investigator Award from the Section on Statistics in Epidemiology at the JSM 2023.
- Lu J*, **Zhang Y***, Cui Y, Peng L, Chen Z (2024). A novel phase II hybrid design to minimize trial duration and improve the success rate of follow-up phase III trial. *Journal of Applied Statistics*, 1–17, DOI: [10.1080/02664763.2024.2382135](https://doi.org/10.1080/02664763.2024.2382135)
- Ge L, **Zhang Y**, Waller LA, Lyles RH. Utilizing a capture-recapture strategy to accelerate infectious disease surveillance. (Accepted by *Annals of Applied Statistics*; Preprint in [arXiv](https://arxiv.org))
- Zhang Y**, Ge L, Waller LA, Lyles RH (2023). On some pitfalls of log-linear modeling framework for capture-recapture studies in disease surveillance. *Epidemiologic Methods*, 12(1), 20230019, DOI: [10.1515/em-2023-0019](https://doi.org/10.1515/em-2023-0019).
- Ge L, **Zhang Y**, Waller LA, Lyles RH (2023). Enhanced inference for finite population sampling-based prevalence estimation with misclassification errors. *The American Statistician*, DOI: [10.1080/00031305.2023.2250401](https://doi.org/10.1080/00031305.2023.2250401).
- Zhang Y**, Chen J, Ge L, Williamson JM, Waller LA, Lyles RH (2023). Sensitivity and uncertainty analysis for capture-recapture methods in disease surveillance. *Epidemiology*, 34(4), 601-610, DOI: [10.1097/EDE.0000000000001614](https://doi.org/10.1097/EDE.0000000000001614).
- Ge L, **Zhang Y**, Lash TL, Ward KC, Waller LA, Lyles RH (2023). Tailoring capture-recapture methods to estimate registry-based case counts based on error-prone diagnostic signals. *Statistics in Medicine*, DOI: [10.1002/sim.9759](https://doi.org/10.1002/sim.9759).
- Lyles RH, **Zhang Y**, Ge L, Waller LA (2023). A design and analytic strategy for monitoring disease positivity and case characteristics in accessible closed populations. *American Journal of Epidemiology*, kwad177, DOI: [10.1093/aje/kwad177](https://doi.org/10.1093/aje/kwad177).
- Zhang Y**, Chang HH, Cheng Q, Collender PA, Li T, He J, Remais JV (2022). A hierarchical model for analyzing multisite individual-level disease surveillance data from multiple systems. *Biometrics*, DOI: [10.1111/biom.13647](https://doi.org/10.1111/biom.13647).
- Lyles RH, **Zhang Y**, Ge L, England C, Ward K, Lash TL, Waller LA (2022). Using Capture-Recapture Methodology to Enhance Precision of Representative Sampling-Based Case Count Estimates. *Journal of Survey Statistics and Methodology*, DOI: [10.1093/jssam/smab052](https://doi.org/10.1093/jssam/smab052).

12. **Zhang Y**, Kutner M, Chen Z (2021). Adaptive Bayesian phase I clinical trial designs for estimating the maximum tolerated doses for two drugs while fully utilizing all toxicity information. *Biometrical Journal*, 63(7), 1476-1492, DOI: [10.1002/bimj.202000142](https://doi.org/10.1002/bimj.202000142).

Applications

1. **Zhang Y**, Chang HH, Iuliano AD, Reed C. A Bayesian spatial-temporal varying coefficients model for estimating excess deaths associated with respiratory infections. (Accepted by *Journal of the Royal Statistical Society, Series A (Statistics in Society)*); Associated R package NBRegAD is available on [GitHub](#)
2. Mehta P, Raymond J, **Zhang Y**, Punjani R, Han M, Larson T, Muravov O, Lyles RH, Horton DK (2023). Prevalence of amyotrophic lateral sclerosis in the United States, 2018. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, pp. 1-7, DOI: [10.1080/21678421.2023.2245858](https://doi.org/10.1080/21678421.2023.2245858).
3. **Zhang Y**, Ebelt ST, Shi L, Scovronick NC, D'Souza RR, Steenland K, Chang HH (2023). Short-term associations between warm-season ambient temperature and emergency department visits for Alzheimer's disease and related dementia in five US states. *Environmental Research*, DOI: [10.1016/j.envres.2022.115176](https://doi.org/10.1016/j.envres.2022.115176).
4. **Zhang Y**, Chang HH, Iuliano AD, Reed C (2022). Application of Bayesian spatial-temporal models for estimating unrecognized COVID-19 deaths in the United States. *Spatial Statistics*, 100584, DOI: [10.1016/j.spasta.2021.100584](https://doi.org/10.1016/j.spasta.2021.100584).
5. Bidot S, Monsrud A, Kline M, Speak A, Martini D, Bilen MA, Switchenko JM, **Zhang Y**, Gerges AG, Farhat GN, Dent EA, Viraj A, Tinsley ML, Harik LR (2022). Risk stratification of prostatic adenocarcinoma metastatic to the lymph nodes. *Archives of Pathology & Laboratory Medicine*, DOI: [10.5858/arpa.2021-0247-OA](https://doi.org/10.5858/arpa.2021-0247-OA).
6. Robinson Jr WE, **Zhang Y**, Chen Z, Calzada O, Churnetski MC, Flowers C, Cohen JB (2017). An investigation of potential prognostic factors in late relapse diffuse large B-Cell lymphoma patients. *Blood*, 130, 5220.

Manuscripts under Review

- **Zhang Y**, Chang HH. Time-to-event analysis of preterm birth accounting for gestational age uncertainties.

Honors and Awards

- **ENAR Distinguished Student Paper Award**, International Biometric Society Eastern North American Region, 2024
- **Statistics In Epidemiology Young Investigator Award**, Joint Statistical Meetings, 2023
- **The Livingston Fellow Award**, Emory University, 2022
— Awarded to PhD candidates for strong leadership skills and academic contributions to public health

Software Developed

- **nbRegQF**: Estimate health effects of environmental exposures accounting for exposure heterogeneity. Available on <https://github.com/YZHA-yuzi/nbRegQF>.
- **NBRegAD**: Estimate excess deaths associated with respiratory infections. Available on <https://github.com/YZHA-yuzi/NBRegAD>.

Presentations

- Oral presentation, ENAR 2024, Baltimore, MD., March, 2024.
— A capture-recapture modeling framework emphasizing prior information in disease surveillance.
- Oral presentation, JSM 2023, Toronto, Ontario, Canada, August, 2023.
— A scalar-on-quantile function approach for estimating short-term health effects of environmental exposures.
- Oral presentation, ENAR 2023, Nashville, TN., March, 2023.
— A scalar-on-quantile function approach for estimating short-term health effects of environmental exposures.
- Oral presentation, GEOMED 2022, Irvine, CA., October, 2022.
— A scalar-on-quantile function approach for estimating short-term health effects of environmental exposures.
- Oral presentation, ENAR 2022 Spring Meeting, Houston, TX., March, 2022.
— A hierarchical model for analyzing multisite individual-level disease surveillance data from multiple systems.
- Oral presentation, JSM 2021 Virtual Conference, Online, August, 2021.
— Sensitivity and uncertainty analysis for capture-recapture methods in disease surveillance.
- Poster, ENAR 2018 Spring Meeting, Atlanta, GA., March, 2018.
— Adaptive Bayesian phase I clinical trial designs for estimating the maximum tolerated doses for two drugs while fully utilizing all toxicity information.

Teaching Experience

Teaching Assistant

- BIOS 500: Biostatistics Methods I - Lecture, Lab. 2017 Fall, Emory University.
- BIOS 520: Clinical Trials. 2019 Spring and 2021 Spring, Emory University.
- BIOS 525: Longitudinal and Multilevel Data Analysis. 2019 Fall and 2020 Fall, Emory University.
- BIOS 544: Introduction to R programming for Non-BIOS students. 2020 Spring, Emory University.

Editorial Activities

Reviewer for *Epidemiology*, *Biometrics*, *Statistics in Medicine*, *BMJ Open*, *Occupational and Environmental Medicine*, *PLOS Neglected Tropical Diseases*, *GeoHealth*.

Skills and Certificates

- **Software Skills:** R, Python, SAS
- **Platform:** Linux-cluster
- **Certification:** SAS Certified Base Programmer