Jie He

Advanced – R/SAS; Proficient – Python/SQL/Shell Scripting

RESEARCH INTERESTS

The design of clinical trials and the development of analytic methods, to extend computational frameworks and tools for both statisticians and medical professionals to work with longitudinal, multicenter data or large-scale, high-dimensional data.

EDUCATION

2024-2028 **Boston University**, Boston, MA
Ph.D. Student in Biostatistics: Research on the Framingham Heart Study under the supervision of Prof. Chunyu Liu.
[R, SAS, and shell scripting; longitudinal analysis; sampling-based models]

2015-2020 **University of North Carolina at Chapel Hill**, Chapel Hill, NC
M.S., Biostatistics, *Not Providing Point-scaled Grades*
B.S., Biostatistics with a second major in Computer Science and a math minor, *Distinction*
Advisor: Di Wu
Thesis: Weighted inference of gene expression variability in single cell RNAseq data for gene set tests: a replication study. [R; high-dimensional scRNA-seq data; permutation-based models]

PROFESSIONAL EXPERIENCE

2021-2024 **Boston Children's Hospital**, Boston, MA
Biostatistician – Biostatistics and Research Department, supervised by Dr. Edie Weller
- Simulation study: Based on structures resembling clinic data from the Extracorporeal Life Support Organization (ELSO) Registry to evaluate model performances regarding different types of data and correlation strengths. [R; electronic health records (EHRs); LASSO regression]
- Prediction: Employed random survival forests on clustered data to identify key features predicting in-hospital mortality following ECPR, and to assess the prediction accuracy of current ELSO registry data. [R; EHRs]
- Experimentation and AB tests: Used Cox models with mixed effects on clustered right-censored data for association analysis. Also validated the theoretical framework of optimizing partial likelihood for time-dependent covariates [1, 2, 3]. Analyzed complex survey data containing six imputations of the missing data to assess the impact of federal welfare programs on children's physical and mental health indicators [4]. Performed chi-squared tests for association analysis of a single center, retrospective study on the onset of epileptic spasms [5].
- Tool development: Developed an R package with custom functions and a Shiny app to streamline the data cleaning and descriptive statistics presentation of extensive, yet unorganized, patient data containing clinical notes from the REDCap database. Enabled researchers to efficiently generate metadata via API [6].

ACADEMIC PROJECTS

Fall 2020 **University of North Carolina at Chapel Hill**, Chapel Hill, NC
Graduate Research Assistant – Biostatistics, supervised by Dr. Donglin Zeng
- Used R to combine national data of social determinants including social vulnerability index, dates of COVID policies such as shut-downs and facial-masks-in-public, U.S. mobility changes provided by Google Map, and demographics. The aim is to assist a doctoral researcher to use them as covariates in trend prediction models.

Spring 2019 **University of North Carolina at Chapel Hill**, Chapel Hill, NC
Graduate Teaching Assistant – BIOS 661, supervised by Dr. Fengchang Lin
- Graded homework and answered questions for graduate-level courses in Probabilities and Statistical Inferences.

2018-2019 **University of North Carolina at Chapel Hill**, Chapel Hill, NC
Undergraduate Honors Thesis – Biostatistics, advised by Dr. Di Wu
- Conducted a hypothesis test for differential gene expression of single cell RNA sequencing where gene level variability was defined as the test measure.
- Wrote R functions to address the mean-variance relationship for zero-inflated RNA-seq counts (32,738 genes and 2,692 cells).

AWARDS

2019   **Undergraduate Honors Thesis with Distinction, UNC**

2018   **Phi Beta Kappa, the Phi Beta Kappa Society**

2018   **Carolina Research Scholar Award, UNC**

2017   **GHC Student Scholarship, AnitaB.org** (Grace Hopper Celebration of Women in Computing)

2013   **Second Prize at National Olympiad in Informatics (NOI), China** (Language Used: C++)

PUBLICATIONS

Accepted or In Review

2025   [1] S. Taneja, J. He, L. Marcil, E. Weller, L. K. Lee. "Association of Social Program Participation with Health, Healthcare Access, and Utilization Among Children in Low-Income Families." In Review 2025.

2024   [2] W. D. Soulsby, R. Olveda, J. He, L. Berbert, E. Weller, K. Barbour, K. Greenlund, L. Schanberg, E. v. Scheven, A. Hersh, M. B. Son, J. Chang, A. Knight. "Factors Driving Disparities in Glucocorticoid Exposure among Children with SLE in the CARRA Registry." Accepted 2024.

[3] W. D. Soulsby, R. Olveda, J. He, L. Berbert, E. Weller, K. Barbour, K. Greenlund, L. Schanberg, E. v. Scheven, A. Hersh, M. B. Son, J. Chang, A. Knight. "Racial Disparities and Achievement of the Low Lupus Disease Activity State (LLDAS): A CARRA Registry Study." Accepted 2024.

Full Publications

2024   [4] Hadjinicolaou, A., Briscoe Abath, C., Singh, A., Donatelli, S., Salussolia, C.L., Cohen, A.L., He, J., Gupta, N., Merchant, S., Zhang, B. and Olson, H., 2024. Timing the clinical onset of epileptic spasms in infantile epileptic spasms syndrome: A tertiary health center's experience. *Epilepsia*.

2023   [5] Mamolea, C., He, J., Weller, E., Hoskote, A., Ryan, K., Cooper, D., Bhaskar, P., Thiagarajan, R. and Vitali, S., 2023. 207: Factors and Outcomes Associated with Circuit Change in the Neonatal and Pediatric ECLS populations: An Analysis of the ELSO Registry. ASAIO Journal, 69(Supplement 3), p.54.

2022   [6] Maddux, A.B., Berbert, L., Young, C.C., Feldstein, L.R., Zambrano, L.D., Kucukak, S., Newhams, M.M., Miller, K., FitzGerald, M.M., He, J. and Halasa, N.B., 2022. Health impairments in children and adolescents after hospitalization for acute COVID-19 or MIS-C. *Pediatrics, 150*(3), p.e2022057798.