JASON MASSEY

919.631.4308
jasonleemassey1991@gmail.com
GitHub: https://github.com/jlmassey1991
Visual Portfolio: https://jasonmassey.crevado.com/

PROFESSIONAL SUMMARY

Over a decade of experience in data science, statistics, and mathematics. Expert in the analytic data cycle: performing complex data linkages from many sources, data management, using advanced machine learning methods, conducting analyses, and creating professional visualizations. An adaptable team lead and communicator with experience tackling complex public health problems.

Relevant Skills: R programming, SAS, ArcGIS, SQL, Matlab, Python, Macros, PowerBI, Linear Regression, Logistic Regression, Longitudinal Data, Automation, Fixed/Random Effects, Time Series, Imputation, Bootstrapping, Clustering, QA, Language Learning Models (LLM), GitHub, Databricks, GAMS, Predictive modeling, Relational databases, Agile, Deep learning, Model Training, Validation, Testing, Electronic medical records/claims, Hierarchal models/Mixed Models/Multilevel Models, Supervised/Unsupervised Learning, Decision Trees, Random Forests, Causal Inference, Propensity Scores, AWS, Azure, Public Speaking, Gradient Boosting

Favorite Packages: Tidyverse, Lubridate, mlr3, Caret, Janitor, Renv, NumPy, Pandas, Matplotlib, TensorFlow, PROC SQL

PROFESSIONAL EXPERIENCE

Senior Mathematical Statistician (Surveillance, Informatics, and Statistics Office)

Centers for Disease Control and Prevention | April 2024-Present

- Lead analyst measuring the effect of diagnostic delays of fungal pathogens on healthcare costs. Performed data linkages to capture Marketscan ICD-10 insurance claims data and used quantile-binning probability weighting to account for the causal relationship for a zero-inflated exposure. Models showed increased healthcare costs to influence future policy implications.
- Acted as the main statistical resource to the Mycotic Diseases Branch and others. Included consulting clients, training colleagues, performing regressions, running statistical tests, and validating models for manuscripts. (SQL/SAS/R/Python)
- Leader to multiple fellows in statistical analysis, coding, and best data practices. Managed professional development.
- Created/led data group holding biweekly meetings to foster upskilling in data science, machine learning, statistics, creating reproducible code, and informatics. Topics include choosing models, tuning hyperparameters etc. (SQL/SAS/R/Python)
- Used cutting-edge DCIPHER software to rapidly link vast amounts of disparate environmental outbreak data from various sources and design data pipelines for public health needs. Included generating reusable frameworks and leveraging language learning models to increase efficiency for performing workflow tasks related to informatics and analytics.

Biostatistician II (Surveillance Branch)

Centers for Disease Control and Prevention | February 2023- April 2024

- Used the National Healthcare Safety Network to translate multileveled healthcare-associated infection surveillance data of more than 65,600 Long Term Care (LTC) health facilities into meaningful insights and published research. (Python/SQL/R)
- Led project that geocoding of LTC facilities including planning, organizing meetings, acquiring data access, validation of geolocations, and conversion of coordinates to census tracts. Linked demographic data include SVI and other metrics.
- Used machine learning package kmlshape to create novel cluster analysis that identified omicron curve peaks in LTC facilities via complex data linkages to the NHSN and performed bootstrapping to account for instability. Ran logistic regression to assess the odds of association between high and low-peak LTC facilities and various characteristics.
- Conducted a survival analysis that investigated the hazard of being infected by COVID among varying vaccination doses and types in LTC facilities; linking event-level, person-level and facility-level data, creating a nested control group of those who did not receive the bivalent booster, and running propensity score models in a cloud computing environment. (SAS/SQL/R)
- Automated pulling from surveillance databases. Merged RWE surveillance data to census, geolocation, CMS, NHSN, and other EMR/ERH insurance claims data to create analytic datasets for analysis and created interactive maps and dashboards.
- Oversaw coding training presentations and provided expertise and guidance via tutorials on high-level data science information. Advises colleagues on coding and data science topics like random forests and Bayesian statistics. (SAS/SQL/R)

Associate Scientist II (Cancer Disparity Team)

American Cancer Society | January 2021- February 2023

- Used a hierarchal Bayesian model in R to assess the spatial association between healthy food access and life expectancy in U.S. Census Tracts and health metrics. Findings showed an independent association between life expectancy and access to healthy foods especially when in the presence of reliable public transit. Helped influence public health policy. (R/ArcGIS)
- In partnership with the state of Idaho's cancer registry, developed a data sharing agreement, data dictionary, and preprocessing dataset for streamlining multiple analytic datasets using real-world evidence (RWE) data.
- Used a random effects mixed multinomial logistic regression model and principal component analysis to find the impact of inflation-adjusted economic segregation on the stage of cancer diagnosis within Idaho. Findings show that economic segregation resulted in an increase in the odds of late-stage cancer diagnosis among all cancer sites. (R/CDISC/ArcGIS)
- Helped to create a nationwide Cancer Atlas for ACS staff; an interactive Web Map Dashboard with oncology statistics, socioeconomic status, office locations, congressional districts. (R-Shiny/ArcGIS Pro)
- Led the monitoring, evaluation, and management of the Cancer Atlas/SEER surveillance systems. (SQL)
- Conducted weighted multinomial logistic regression on health disparities among various cancer sites (SAS/SQL/R)
- Used Power BI, Tableau, and ArcGIS to develop dashboards and web mapping tools with cancer statistics.

Surveillance Epidemiology Analyst
Oak Ridge Institute | 2020-2021

- Led data management for over 44,000 mothers in the BD-Steps database; birth defect case control. This involved using queries to store and organize data to be used in future analyses. (SAS/SQL/Python)
- Evaluated several public health departments' surveillance systems for federal funding approval.
- Led ongoing development and implementation of a nationwide surveillance system by strategizing with stakeholders.
- Provided consultations, training, and workshops with state and local health departments on survey methods, data collection, and data analysis working with real-world evidence (RWE) and administrative claims data.
- Conducted replications, and validations, created visualizations for longitudinal analyses, and automated monthly data reports on birth defects data. (SAS/SQL/Power BI/Tableau)

Biostatistics Research Assistant
Emory University | 2018-2020

- Created a literature review and analysis plan on the burden of influenza from hospital morbidity data in LA county. (Excel)
- Linked ICD 9/10 code data from large datasets from CDC, NOAA, and Los Angeles Health Department. (SQL/SAS)
- Estimated the Burden of Influenza with Quasi-Poisson Time Series Spline Models via primary and secondary diagnoses. (R)

Mathematics and Statistics Tutor
North Carolina State University | 2011-2017

- Tutored mathematical and statistical modeling, probability, linear algebra, calculus, differential equations, real analysis.

Electronic Data Interchange (EDI) Analyst
Advance Auto Parts | 2016-2017

- Stored manipulated, deduplicated, flagged anomalies, and analyzed EDI files that tracked purchase orders and shipping records; Consulted with clients to fix clerical errors and train in EDI formatting. (Excel/ SQL)

ACADEMIC EDUCATION

***Emory University Rollins School of Public Health,* (May 2020)**
Master of Public Health, Epidemiology (3.8)

***North Carolina State University*, (May 2013)**
Bachelor of Science in Applied Mathematics (3.2)

VOLUNTEER EXPERIENCE

Raleigh LGBT Center
Provided public health educational resources regarding HIV and sexual health to men who have sex with men.
Rock Club Rock Climbing Belay Coordinator
As a part of my minor practicum helped to establish and facilitate an instructional rock-climbing belaying clinic in the community.

PUBLICATIONS

Jason Massey, Daniel Wiese, Farhad Islami, Ahmedin Jemal, Marjorie McCullough (2023): *"The Association Between Census Tract Healthy Food Accessibility and Life Expectancy in the United States"*. Journal of Urban Health. *The Association Between Census Tract Healthy Food Accessibility and Life Expectancy in the United States - PubMed (nih.gov)*

Jeneita Bell **et al**. (2023): "*Influenza and Up-to-Date COVID-19 Vaccination Coverage Among Health Care Personnel — National Healthcare Safety Network, United States, 2022–23 Influenza Season"*. Centers for Disease Control and Prevention. MMWR. 72(45);1237–1243. Influenza and Up-to-Date COVID-19 Vaccination Coverage Among Health Care Personnel — National Healthcare Safety Network, United States, 2022–23 Influenza Season | MMWR (cdc.gov)

Farhad Islami MD Daniel Wiese, Emily C Marlow, Tyler B Kratzer, **Jason Massey**, Hyuna Sung, Ahmedin Jemal (2023): *Progress in reducing cancer mortality in the United States by congressional district, 1996–2003 to 2012–2020.* Cancer. *https://acsjournals.onlinelibrary.wiley.com/doi/10.1002/cncr.34808*

Jason Massey, Daniel Wiese, Farhad Islami, Ahmedin Jemal (expected 2025): *"Index of Concentration of Extremes and Cancer Survival in Idaho Census Tracts"*

Lu Meng*, Jeneita M. Bell*, Sydney Guthrie, Kira Barbre, Larry Mason**, Jason Massey**, Ryan Wiegand, Theresa Rowe, Austin Woods, Iram Qureshi, Hannah Reses, Alfonso Hernandez-Romieu, Matthew J. Stuckey, David Kuhar, Megan C. Lindley, Andrea Benin (expected 2025): *"High versus low SARS-CoV-2 infection peak and resilience to COVID-19 surge among nursing home residents during Omicron variant BA.1 wave in the United States, December 20, 2021-March 20, 2022, National Healthcare Safety Network, United States"*

Lu Meng **et al**. (expected 2025): *Vaccination and risk of COVID infection in State Veteran Home Facilities*

Kaitlin Benedict, **Jason Massey**, Michelle Fearon Scales, Ian Hennessee, Samantha L. Williams, Mitsuru Toda (expected 2025): *"Impact of delays in diagnosis on healthcare costs associated with blastomycosis, coccidioidomycosis, and histoplasmosis in a commercially insured population"*

L Mason, BS[1,2];W Koech, PhD[1,3]; Young, Janine, MD1; S Guthrie, MPH[1,3]; K Barbre, MPH[1,3]; **J Massey, MPH[1]**; L Meng, PhD[1]; T Rowe, MD[1]; L A Benin, MD[1], J Bell, PhD[1] (expected 2025) *"Using Index of Concentration at the Extremes to Examine Skilled Nursing Home Characteristics and Community Dynamics, United States, August 2024"*

REFERENCES

Lu Meng, PhD
Deputy Team Lead
CDC | Division of Healthcare Quality and Promotion
prj8@cdc.gov
404.498.2905

Lawrence McDonald, MD
Associate Director for Science
CDC | Division of Healthcare Quality and Promotion
ljm3@cdc.gov
404.639.3833

Preeta Kutty, MD, MPH
Medical Epidemiologist
CDC | Division of Healthcare Quality and Promotion
chz6@cdc.gov
404.639.0470

Theresa Rowe, DO
Unit Team Lead (Long-Term Care)
CDC | Division of Healthcare Quality and Promotion
nnq1@cdc.gov
404.498.2905

Logan Ray, MPH
Associate Director of Surveillance
CDC | Surveillance, Informatics, and Statistics Office
prj8@cdc.gov
404.498.2905

Lyndsay Bottichio, MPH, DrPH
Principal Deputy Director
CDC | Division of Readiness and Response Science
prj8@cdc.gov
404.498.2905

Daniel Wiese, PhD
Principal Scientist (Cancer Disparities)
American Cancer Society | Surveillance
Daniel.wiese@cancer.org
215.459-9483

Farhad Islami, MD, PhD
Senior Scientific Director (Cancer Disparities)
American Cancer Society | Surveillance
Farhad.islami@cancer.org
678.528-3133