Email: babijah@umass.edu, benabijah@gmail.com; Website: benabijah.github.io

Benjamin A. Abijah

Statistical and data science methods for large-scale biomedical data

1. EDUCATION

University of Massachusetts, Amherst — Expected Sep 2026
Ph.D. Biostatistics

University of Massachusetts, Amherst — May 2023
M.S. Statistics

Kwame Nkrumah Univ. of Sci. & Tech., Ghana — Aug 2020
B.Sc. Statistics

2. PROFESSIONAL MEMBERSHIP/ACCREDITATION/CERTIFICATION

Certificate, Statistical and Computational Data Science, University of Massachusetts, Amherst — Expected May 2026
GStat, American Statistical Association — 2024 – present
Member, American Statistical Association (ASA) — 2024 – present

3. SKILLS

* *denotes "intermediate"*
Computing: R, Python, git, SAS*, SPSS, SQL*, MATLAB*
Methods: Multiomics, Networks, Bioinformatics, Bayesian and Survival models, Machine Learning applications in health.
Teaching: 5 years teaching – including assisting statistics courses of over 300 students and mentoring 5 undergraduates.
Languages: English (fluent), West African Pidgin (fluent), Fante (native), and Twi (fluent)

4. WORK EXPERIENCE

Research Assistant, Raji Lab, UMass Amherst — Aug 2023 – present

- Applying advanced statistical methods to analyze association of proteomics with CVD outcomes and detect critical sex differences.
- Comparing a suite of multitask learning methods to identify selectively advantageous ones for feature selection of proteomics associated with CVD outcomes.
- Developing machine learning methods for multitask feature selection of protein biomarkers for CVDs.
- Predicting particulate matter pollutant using land use models to estimate air pollution exposure.

Consulting Assistant, Statistical Consulting & Collaboration Services, UMass Amherst — May – Aug 2022

- Provided statistical expertise for clients' biomedical, pharmaceutical, and entomology projects.
- Directly supported clients' research from data ingestion through analysis implementation and helped to translate findings into actionable health solutions.

5. PUBLICATIONS/MANUSCRIPTS

1. Frempong, N.K., Berchie, R.O., Baidoo, R., **Abijah, B.A.**, & Oforiwaa-Amanfo, O.Y. (2021). A Simulation Study to Examine the Bias of Some Sample Measures of Skewness. Applied Mathematical Sciences, 15(4), 189-200. https://doi.org/10.12988/ams.2021.914276
2. **Abijah, B.A.**, Sanchez, K., Janiczek, M.L., Spracklen, C.N., Zeleznik, O.A., DeMeo, D.L., Rexrode, K.M., & Balasubramanian, R. Sex Differences in Proteins Associated with Incident Ischemic Stroke in the UK Biobank. (**in progress**)
3. **Abijah, B.A.** Balasubramanian, R. & Tadesse, M.G. A Tutorial on Multitask Learning Methods for Proteomics Feature Selection for CVD outcomes (**in progress**)
4. **Abijah, B.A.**, Mottey, B., Janiczek, M., Balasubramanian, R. & Arku, R. Land Use Regression Models for Predicting $PM_{2.5}$ for Epidemiologic Studies in Africa: A Comparative Analysis from Accra and Kigali. (**in progress**)

6. CONFERENCE PRESENTATIONS

1. Multitask Learning Methods for Predicting Coronary Vascular Disease Outcomes Using Proteomics in the UK Biobank. 38th New England Statistics Symposium. New Haven, CT. (Jun 2025)
2. Land Use Regression Models for Predicting $PM_{2.5}$: A Comparative Analysis from the Accra Birth Cohort. Joint Statistical Meetings. Nashville, TN. (Aug 2025) (**poster**)
3. Identifying Proteins associated with Stroke Outcomes in the UK Biobank Study. UMass Amherst 28th SPHHS Research Day. Amherst, MA. (Apr 2025) (**poster**)

7. SELECTED PROJECTS

- Longitudinal Analysis of Diabetes Progression in Medicare Patients Using Claims Data. 2025.
  - Analyzed synthetic electronic health data (from the Centers for Medicare & and Medicaid Services) and tracked beneficiaries' progression from prediabetes diagnoses to diabetes mellitus with complications.
- Estimating Infertility Prevalence by Applying a Bayesian Current Duration Approach to Demographic and Health Survey Data. 2024.
  - Applied a flexible Bayesian model to estimate infertility prevalence in cross-sectional data from the 2018 Nigeria DHS, which offers improved and alternate infertility estimates over traditional parametric methods.
- Identifying Single-nucleotide polymorphisms (SNPs) associated with Asthma. 2024.
  - Identified SNPs associated with asthma, a chronic respiratory condition affecting millions worldwide, to understand some of the genetic factors involved in its risks.
- A Bayesian Simulation: the case of Mis-specified Data Generating Process. 2024.
  - Investigated the effects of mis-specifying the likelihood function in Bayesian survival analysis, particularly in cases where the data-generating process is uncertain.
- An Exploration of Heart Failure Clinical Data. 2023.
  - Analyzed clinical records of heart failure patients in Faisalabad, Pakistan, to identify key clinical features associated with heart failure-induced deaths, using a suite of machine learning methods.
- Mis-specifying the Variance: Does Clustering Affect Estimates? 2023.
  - Explored the effects of mis-specified variance structures in clustered/hierarchical data, specifically in models where separate variances are apparent for the two treatment groups.
- Cox Proportional Modeling of Cancer Patients' Recurrence-free Survival. 2022.
  - Examined the effects of treatment on recurrence-free and overall survival in cancer patients and underscored the importance of considering patient-specific factors, such as tumor grade and lymph node involvement, in treatment strategies.
- Socio-economic Inequalities in Childhood Mortality in Ghana. 2020.
  - Investigated some socio-economic determinants of childhood mortality in Ghana with focus on disparities in survival risks among children under five.

8. SELECTED AWARDS/HONORS

- SPHHS Dean's Fellowship Award, UMass Amherst. 2023 – 2025. **$30,000**
- Travel Grant, UMass Amherst. 2025. **$900**
- Barclays Bank Scholarship, Barclays Bank Ghana. 2017 – 2020. **$3,000**
- Valedictorian, 54th Congregation, College of Science, KNUST

9. TRAININGS/WORKSHOPS

- Practical Considerations for Adaptive Clinical Trials Using Bayesian and Frequentist Methods. Joint Statistical Meetings, American Statistical Association. 2025.
- Causal Inference in Randomized Controlled Trials. Joint Statistical Meetings, American Statistical Association. 2025.
- Statistical Considerations in Cell and Gene Therapy Development. Joint Statistical Meetings, American Statistical Association. 2025.
- Optimization for Data Science and Machine Learning Problems. New England Statistics Symposium. 2025.

- Fundamentals of Causal Inference With R. Boston Chapter, American Statistical Association. 2024.
- SHARP Training in Mendelian Randomization. Columbia University Mailman School of Public Health. 2024.

10. SELECTED COMMUNITY SERVICE

Leadership Activities:

Student Rep., Department of Biostatistics and Epidemiology, UMass Amherst 2025 – present
- Coordinate with department to organize student engagement programs like professional development seminars, social events, etc.

Chair, Education and Mentorship, EKO Global Foundation - Worldwide 2022 – present
- Coordinate career and professional mentorship for scholarship beneficiaries. Developed mentorship curriculum and match beneficiaries with mentors.

Chair, Social Media Committee, African Graduates and Scholars' Association, UMass Amherst 2022 – 2023
- Publicized, run, and managed the association's social media platforms.

Member, Strategic Planning Committee, Science Students Association, KNUST 2019
- Developed a 10-year Strategic Plan, "PLAN 2K29", for the association.

Chair, Academic Committee, Association of Mathematics and Statistics Students, KNUST 2018 – 2019
- Coordinated all academic mentorship activities of the association.

Mentoring Activities:

Graduate Mentor – 5 Students 2025

Summer Workshop in Biostatistics - SWIB25, UMass-Amherst

- Ana Hutchinson (Springfield College, MA), Belle Song (Mount Holyoke College, MA), & Monet Williams (UMass Amherst, MA). "Identifying Predictors of Type II Diabetes in the NHANES Study".
- Bao Nguyen (UMass Amherst, MA) & Sherry Zhang (Emory University, GA). "A Comparative Study of Machine-Learning Classifiers for Predicting HOLC Grades".