

Ngoc Duong

Chicago, IL, 60610 | ngoc.duong@northwestern.edu | (909) 973-3401

EDUCATION

Northwestern University	09/2023 – Present
• PhD candidate in Biostatistics, GPA: 3.96/4.00	
Columbia University Mailman School of Public Health	09/2019 – 05/2021
• MS in Biostatistics, Theory & Methods Track, GPA: 3.93/4.00	
Vassar College	08/2014 – 05/2018
• BA in Economics and Chinese, General and Department Honors, GPA: 3.83/4.00	

RELEVANT DATA ANALYSIS PROJECTS

Marginal structural modeling with competing risks using g-formula. Northwestern University

- Integrated and cleaned intensive longitudinal PICU EHR data in R to build an analysis-ready dataset with standardized time-varying treatments, covariates, and organ dysfunction outcomes.
- Implemented parametric g-formula and marginal structural models to estimate the causal effect of mean arterial pressure (MAP) guideline complete adherence on 7-day multiple organ dysfunction, accounting for time-varying confounding and competing risks.

Topic modeling analysis of funded cardiovascular research grants. Northwestern University

- Performed text preprocessing on ~10,000 cardiovascular grant abstracts (tokenization, stemming, stop-word removal, TF and TF-IDF matrices).
- Applied Latent Dirichlet Allocation and spherical k-means clustering methods to identify research themes and funding patterns across global cardiovascular research organizations from 2017 to 2024.

Covariate-adaptive randomization algorithm for factorial design. Northwestern University

- Designed a modified covariate-adaptive randomization algorithm for sequential factorial clinical trials to improve covariate balance across treatment components.
- Conducted extensive R-based simulation studies to compare covariate imbalance, bias, coverage probability, and power versus standard simple and blocked randomization schemes.

Cognitive Remediation in Schizophrenia: Exploring Treatment Effect Heterogeneity to Improve Treatment Success. Columbia University

- Cleaned and harmonized data from three randomized clinical trials of cognitive remediation in schizophrenia including variable screening, factor encoding, random forest-based imputation, and feature scaling in R.
- Built a machine learning pipeline using SuperLearner and nested cross-validation to estimate individualized treatment effects, derive optimal individualized treatment rules, and evaluate the clinical value of those rules.

RELEVANT WORK EXPERIENCE

Division of Biostatistics and Informatics, Northwestern University

Graduate Research Assistant	07/2024 – Present
• Independently lead multiple biostatistics research projects from problem formulation to analysis, visualizations, and result dissemination, including defining research questions of interest, evaluating statistical strategies with faculty advisors, and implementing the chosen approaches with reproducible workflows.	
• Deliver weekly analytic updates, collaborate with data managers to resolve data issues if necessary, and lead development of manuscripts.	

Division of Biostatistics and Informatics, Northwestern University

Graduate Teaching Assistant	06/2024 – 12/2024
• Held weekly discussion sessions to review and explain epidemiology lab problem sets	
• Prepared and led coding labs on data cleaning, descriptive statistics, regression modeling and propensity score methods in R	

Department of Pediatrics, Columbia University

Statistical Analyst

05/2021 – 07/2023

- Performed data cleaning, visualizations, statistical analyses, and result interpretation using EHR data for research projects across various study designs
- Prepared statistical analysis plans for research projects; utilizing statistical knowledge to propose and perform additional analyses when appropriate
- Collaborated on preparation of grants, abstracts and research manuscripts

Department of Biostatistics, Columbia University

Graduate Teaching Assistant, Introduction to Biostatistics

09/2020 – 12/2020

- Held weekly office hours to help students understand lecture materials and biostatistical concepts
- Prepared and led discussions and demonstrations on applications of statistics and data analysis in R

ADDITIONAL

Quantitative skills: Statistical modeling (GLM), Longitudinal data analysis (mixed-effects models, GEE), Survival analysis, Causal Inference (propensity scores methods, marginal structural models, mediation analysis, double machine learning, meta-learners), Machine Learning, Deep Learning, Natural Language Processing

Programming: R (advanced), Python (advanced), SAS (intermediate), Stata (intermediate)

Other: Authorized to work for any US employer (no future sponsorship required)

PUBLICATIONS

1. **Duong N**, Lee JY, Peprah Y, Brown T, Linder JA, Doctor JN, Meeker D, Fox C, Goldstein N, Persell S, Petito LC. Changing Clinician Behavior in Geriatrics: Point-of-Care Alerts for Prostate-Specific Antigen Screening. Under review at *American Journal of Preventive Medicine*.
2. **Duong N**, Gauen A, Khan S, Petito LC. Charting the Currents of Private Cardiovascular Research: A Topic Modeling Analysis. Submitted to *Journal of the American Heart Association*.
3. **Duong N**, Pitts AJ, Kim S, Miles CH. Sensitivity analysis for transportability in multi-study, multi-outcome settings. *arXiv preprint*. arXiv:2301.02904; 2023.
4. Woo Baidal JA, **Duong N**, Goldsmith J, Hur C, Lauren BN, Partida I, Rosenthal A, Hulse E, Shea S, Cheung K, Meyer D. Association of a primary care-based mobile food pantry with child body mass index: A propensity score-matched cohort study. *Pediatr Obes*. 2023;18(6):e13023.
5. Weaver DJ, Dagher T, **Duong N**, Winfrey S, Koo A, Balach T. Assessing the experiences of sexual and gender minority applicants to orthopaedic surgery residency. *JBJS Open Access*. 2025;10(1):e24.00158.
6. Finkel MA, **Duong N**, Hernandez A, Goldsmith J, Rifas-Shiman SL, Dumitriu D, Oken E, Shechter A, Woo Baidal JA. Associations of infant sleep characteristics with childhood cognitive outcomes. *J Dev Behav Pediatr*. 2024;45(6):e560-e568.
7. Woo Baidal JA, Finkel MA, Kelman E, **Duong N**, Bien-Aime C, Goldsmith J, Albrecht SS, Hulse E, Rosenthal A, Reiss J, et al. Longitudinal associations of food security with health and dietary factors among Food FARMacy participants during COVID-19 in New York City. *Nutrients*. 2024;16(3):434.
8. Woo Baidal JA, Meyer D, Partida I, **Duong N**, Rosenthal A, Hulse E, Nieto A, HERALD Collaborative. Feasibility of Food FARMacia: Mobile food pantry to reduce household food insecurity in pediatric primary care. *Nutrients*. 2022;14(5):1059.
9. Woo Baidal JA, Nichols K, Charles N, Chernick L, **Duong N**, Finkel MA, Falbe J, Valeri L. Text messages to curb sugar-sweetened beverage consumption among pregnant women and mothers: A mobile health randomized controlled trial. *Nutrients*. 2021;13(12):4367.