Diptavo Dutta

Dept. of Biostatistics
Johns Hopkins University
email: diptavo21@jhu.edu
Contact: 734-223-1009
Website

Current Position

Postdoctoral Fellow 2019 - Present
Dept. of Biostatistics, Johns Hopkins University
Advisors: Dr. Nilanjan Chatterjee and Dr. Alexis Battle

Education

PhD. in Biostatistics 2014 - 2019
Dept. of Biostatistics, University of Michigan

- Dissertation: Statistical Methods for Gene-based and Gene-Set Analysis
- Advisor: Dr. Seunggeun (Shawn) Lee
- Committee Members: Dr. Michael Boehnke, Dr. Laura J. Scott, Dr. Ananda Sen

M.Stat 2012 - 2014
Indian Statistical Institute

- Graduate with 1^{st} class Distinction
- Specialization: Biostatistics and Data Analysis

B.Sc (Statistics Hons.) 2009 - 2012
St. Xavier's College, University of Calcutta

- Graduate with 1^{st} class Distinction
- Auxiliary subjects: Mathematics, Computer Science

Professional Experience

Summer Intern, HEOR Oncology June, 2017 - August, 2017
AbbVie Inc.

1. Treatment Patterns and Survival in Patients with Glioblastoma: A Retrospective Database Analysis Using US Electronic Health Records (EHR)
2. Treatment Patterns in Patients with Chronic Lymphocytic Leukemia (CLL) Treated with B-Cell Receptor Inhibitors (BCRis) in Canada : A Medical Chart Review Study

Research Assistant Sept., 2014 - June, 2019
Dept. of Biostatistics, University of Michigan
Advisors: Dr. Seunggeun (Shawn) Lee & Dr. Laura J. Scott

Summer Intern May, 2013 - July, 2013
Central Statistical Organization, Govt. of India

Peer reviewed Publications

- **Dutta, D.**, Scott, L., Boehnke, M., Lee, S. (2019) Multi-SKAT: General framework to test multiple phenotype associations of rare variants. *Genetic Epidemiology*; 43(1), 1-20
- **Dutta, D.**, Gagliano, S. A., Weinstock, J., Zawistowski, Sidore, C., Fritsche, L., M., Cucca, F., Schlessinger, D., Abecasis, G., Brummett, C., Lee. S. (2019)

Meta-MultiSKAT: Region-based rare variant meta-analysis of multiple phenotypes using summary statistics. *Genetic Epidemiology*; 43(7), 800-814

- **Dutta, D.**, Brummett, C., Fritsche, L., Moser, S., Tsodikov, A., Lee, S., Clauw, D., Scott, L. (2020) Heritability of the fibromyalgia phenotype varies by age. *Arthritis & Rheumatology*; 72(5), 815-823
- Arvanitis, M., Tampakakis, E., Zhang, Y., Wang, W., Auton, A., 23andMe Research Team, **Dutta, D.**, $\cdots$, Battle, A. (2020) Genome-wide association and multi-omic analyses reveal ACTN2 as a gene linked to heart failure. *Nature Communications*; 11(1), 1-12
- **Dutta, D.**, VandeHaar, P., Fritsche, L. G., Zollner, S., Boehnke, M., Scott, L. J., Lee, S. (2021) A powerful subset-based gene-set analysis method identifies novel associations and improves interpretation in UK Biobank. *The American Journal of Human Genetics*; 108(4) 669-681
- Cox, C. K., Zawitowski, M. , Pandit, A., **Dutta, D.**, Narla, G., Swenson, C. W. (2021) Genome Wide Association Study of Pelvic Organ Prolapse using the Michigan Genomics Initiative. *Female Pelvic Medicine & Reconstructive Surgery*; 27(8) 502-506.
- Grams, M. E., Surapaneni, A., Chen, J., Zhou, L., Yu, Z., **Dutta,D.**, $\cdots$, Coresh, J. (2021) Proteins Associated with Risk of Kidney Function Decline in the General Population. *Journal of the American Society of Nephrology*; 32(9) 2291-2302 .
- Zhang, J., **Dutta, D.**, Kottgen, A., $\cdots$, Chatterjee, N. (2022) Plasma proteome analyses in individuals of European and African ancestry identify cis-pQTLs and models for proteome-wide association studies. *Nature Genetics*; 54(5) 593-602.
- Qi, G. †, **Dutta, D.**†, Leroux, A., Ray, D., Crainiceanu, C., Chatterjee, N. (2021) Genome-wide association studies of 27 accelerometry-derived physical activity measurements identifies novel loci and genetic mechanisms. *Genetic Epidemiology*; 46(2) 122-138.
- Rhee, E., Surapaneni, A., Zheng, Z., Zhou, L., **Dutta, D.**, $\cdots$, Grams, M. E. (2021) A Trans-ethnic Genome-wide Association Study of Blood Metabolites in the Chronic Renal Insufficiency Cohort (CRIC) Study. *Kidney International*; 101(4) 814-823.
- Shabani, M., **Dutta, D.**, $\cdots$, Lima, J. Rare genetic variants associated with Myocardial Fibrosis in Multi Ethnic Study of Atherosclerosis (MESA). *Frontiers in Cardiovascular Medicine*; 9: 804788.
- **Dutta, D.**, He, Y., Saha, A., Arvanitis, M., Battle, A., Chatterjee, N. (2020) Aggregative trans-eQTL analysis detects trait-specific target gene sets in whole blood. Preprint available on medRXiv. To appear in *Nature Communications*.

Publications under revision or submitted

- **Dutta, D.**, Ray, D., Chatterjee, N. (2021) Genetic association tests with multiple phenotypes: A review of current methods. (*Invited review article*)
- **Dutta, D.**, Sen, A., Satagopan, J. (2021) Sparse canonical correlation to identify copy number aberration-regulated genes for multiple breast cancer outcomes. Preprint available on medRXiv. (*Under Review*)
- Maity, S., **Dutta, D.**, Terhorst, J., Sun, Y., Banerjee, M. A linear adjustment based approach to posterior drift in transfer learning. Preprint available on arXiv. (*Under Review*)

- Surapaneni, A., ⋯, **Dutta, D.**, Coresh, J., Rhee, E., Grams, M. E. 969 protein quantitative trait loci identified in an African-American population with kidney disease. (*Under Review*)
- Qi, G., Chhetri, S., Ray, D., **Dutta, D.**, ⋯, Chatterjee, N. Genome-Wide Large-Scale Multi-Trait Analysis Characterizes Global Patterns of Pleiotropy and Unique Trait-Specific Variants Preprint available on biorXiv. (*Under Review*)
- Kumthekar, P., Dixit, K. S., Kamalakar, R., **Dutta, D.**, Holen, K., Shaikh, N. I., Ganguli, A. (2020) Treatment Patterns and Survival in Patients with Glioblastoma Multiforme: A Retrospective Database Analysis Using US Electronic Health Records (EHR). (*Under Review*)

†: Joint first author

Papers in progress

- **Dutta, D.** & Chatterjee, N. Fast and efficient liability model for binary phenotypes in Biobanks.
- **Dutta, D.**, Zhang, J., ⋯, Chatterjee, N. Trans-regulated protein networks associated with known complex trait variants.
- Shabani, M., **Dutta, D.**, ⋯, Lima, J. Rare Genetic Variants in Individuals with Low ASCVD Risk and Hard CHD or High Coronary Artery Disease: Multi-Ethnic Study of Atherosclerosis.
- Chhetri, S., **Dutta, D.**, ⋯, Battle, A. A colocalization approach to determine shared heritability of diseases across ethnicities.

Presentations

- Genome Informatics: Poster (2020)
- American Society for Human Genetics Annual Meeting, Virtual: Poster (2020)
- International Genetics Epidemiology Society Annual Meeting, Houston: Platform talk (2019)
- American Society for Human Genetics Annual Meeting, Houston: Poster (2019)
- American Society for Human Genetics Annual Meeting, San Diego: Platform talk (2018)
- International Conference, Institute for Applied Statistics Sri Lanka, 2017, Colombo: Invited Talk
- International Genetics Epidemiology Society Annual Meeting, Cambridge: Poster (2017)
- MSSISS, Ann Arbor: Speed Oral presentation (2017)
- American Society for Human Genetics Annual Meeting, Vancouver: Poster (2016)
- Joint Statistical Meetings, Chicago: Contributed Talk (2016)
- MSSISS (Michigan Student Symposium for Interdisciplinary Statistical Sciences), Ann Arbor: Poster (2016).
- Projects at AbbVie presented at SNO, 2017 and ASH, 2017.

Awards & Achievements

Williams Award Finalist: IGES Meeting (2019)
Rackham Travel Grant: University of Michigan
INSPIRE Scholarship: Dept. of Science and technology, Govt. of India.

R-Packages & Software

- MultiSKAT: Rare variant association tests for multiple phenotypes
- Meta-MultiSKAT: Meta analysis of rare variant association tests for multiple phenotypes.
- GAUSS: Gene-set association analysis using GWAS summary statistics
- PathWeb: Visualization for results for application of GAUSS on UK Biobank data.
- PWAS: Visualization of proteome-wide association analysis on ARIC data for European and African Americans.
- ARCHIE: trans-regulated gene modules for known trait-related genetic variants using sparse canonical correlation

Professional Membership

American Society for Human Genetics, International Genetic Epidemiology Society, American Statistical Association.

Peer Review

The American Journal of Human Genetics, PLoS Genetics, PLoS Computational Biology, Annals of Applied Statistics, Annals of Human Genetics, Scientific Reports, NAR Genomics and Bioinformatics, PLoS One, BMC Medical Genomics, NPJ Genomic Medicine, International Journal of Cancer, Health Services and Outcomes Research Methodology, Journal for Trauma Nursing, International Journal of Biostatistics

Sub-reviewer: *Machine Learning in Computational Biology, International Conference on Machine Learning (Computational Biology workshop)*

Organization

Led the organization of:

- Statistical Genetics Journal Club, Dept. of Biostatistics, University of Michigan, Fall 2015
- Statistical Genetics Working Group Seminar, Dept. of Biostatistics, Johns Hopkins University, Fall 2021 and Spring 2022

Technical Skills

Programming Languages: R, Python, C++, C
Operations Technologies: Google Cloud, AWS
Operating Systems: Linux, Windows
Statistical Softwares: SAS, Minitab, SPSS
High-performance computing clusters

References

Available upon request.