# Sheila M. Gaynor, PhD

Biostatistics Postdoctoral & NHLBI BioData Catalyst Fellow

Harvard T.H. Chan School of Public Health · 677 Huntington Avenue, Boston MA 02115

(919) 656-8433 | sheilagaynor@hsph.harvard.edu | sheilagaynor.com | sheilagaynor


## Summary


Biostatistician working at the intersection of statistical genetics and genomics, data science, and statistical computing.

Expertise in data-driven science (*high-dimensional data modeling, large-scale sequencing analyses, functional data integration, genomic networks*), statistics (*mixed models, statistical learning, multivariate analysis*), cross-functional collaboration, and research leadership.

Technical skills including R, WDL, Linux, cluster and cloud computing, GitHub.


## Experience

**Harvard University** *Boston, MA*

Postdoctoral Fellow in the Department of Biostatistics *2018 - Present*

- Develop cloud-based computational pipelines for large-scale, reproducible rare variant analyses incorporating functional annotations
- Lead rare variant analysis of whole genome sequencing (WGS) data from the Trans-Omics for Precision Medicine (TOPMed) Program to study lung diseases, diabetes and glycemic traits, and inflammation biomarkers
- Develop inferential methods to improve statistical power for rare variant analysis using auxiliary phenotypes and functional data
- Improve fine-mapping methods and applications to identify likely causal genetic variants associated with lung cancer
- Mentor students, serve on thesis committees, prepare grants, and build scientific collaborations

**Duke University** *Durham, NC*

Statistical Consultant for the Center for Translational Pain Medicine *2017 - 2020*

- Conducted and validated clustering analyses on individuals from multiple cohorts with broad pain phenotyping
- Created clinically-implemented algorithm for pain subtyping to identify patients with increased risk status

**Boston University** *Boston, MA*

Visiting Researcher in Behavioral Science Research *2015 - 2017*

- Led statistical analysis of a cohort study to identify latent classes of smokers not motivated to quit within 30 days
- Executed analysis plan in R and identified three distinct subtypes of unmotivated smokers

**McLean Hospital** *Belmont, MA*

Visiting Researcher in Neurobiology of Fear Laboratory *2015 - 2017*

- Performed network and mediation analyses to identify biological mechanisms and mutations characterizing PTSD
- Contributed to book chapter on gene-environment interaction with applications to trauma disorders

## Education

**Harvard University** *Cambridge, MA*

Ph.D. in Biostatistics, A.M. in Biostatistics *2013 - 2018*

- Dissertation entitled “Statistical Methods for Integratively Characterizing Genetic and Genomic Data”
- Co-advised by Dr. Xihong Lin and Dr. John Quackenbush
- Developed statistical methods to perform mediation analysis with common binary outcomes, build and analyze tissue-specific eQTL networks, and detect community structure in regression-based networks

**University of North Carolina** *Chapel Hill, NC*

B.S.P.H. in Biostatistics, B.A. in Mathematics *2009 - 2013*

- With highest honors and highest distinction

## Awards

Fellowships

2020 **BioData Catalyst Fellowship in Cloud-based Biomedical Research**, NHLBI

2017 **F31 Kirschstein Predoctoral Individual National Research Service Award**, NHLBI

2013 **National Science Foundation Graduate Research Fellowship**, NSF

2013 **T32 NIH HIV/AIDS Training Grant Fellowship**, NIAID

### Honors

2020 **Rising Star in Computational and Data Science Award**, University of Texas at Austin
2017 **Summer Institute in Statistical Genetics Scholarship**, University of Washington
2017 **Program in Quantitative Genomics Travel Award**, Harvard University
2016 **XSEDE Computation Allocation**, NSF
2016 **Certificate of Distinction in Teaching**, Harvard University Department of Biostatistics

## Selected publications

**Gaynor, S.M.**, Fillingim, R.B., Zolnoun, D.A., Slade, G.D., ... & Bair, E. (2021). Association between craniofacial pain and hormonal contraceptive use: The OPPERA study. *Journal of Oral & Facial Pain and Headache*.

**Gaynor, S.M.**, Bortsov, A.V., Bair, E., Fillingim, R.B., ... & Smith, S. (2021). Phenotypic profile clustering pragmatically identifies diagnostically and mechanistically informative subgroups of chronic pain patients. *PAIN*.

Sun, R., Xu, M., Li, X., **Gaynor, S.**, ... & Lin, X. (2021). Integration of multiomic annotation data to prioritize and characterize inflammation and immune-related risk variants in squamous cell lung cancer. Genetic epidemiology, 45(1), 99-114.

Raffield, L.M., Iyengar, A.K., Wang, B., **Gaynor, S.M.**, ... & Auer, P.L. (2020). Allelic Heterogeneity at the CRP Locus Identified by Whole-Genome Sequencing in Multi-ancestry Cohorts. *American Journal of Human Genetics*, 106(1), 112-120.

Li, X.*, Li, Z.*, Zhou, H., **Gaynor, S.M.**, ... & Lin, X. (2020). Dynamic incorporation of multiple in silico functional annotations empowers rare variant association analysis of large whole-genome sequencing studies at scale. Nature genetics, 52(9), 969-983.

**Gaynor, S.M.***, Sun, R.*, Lin, X, & Quackenbush, J. (2019). Identification of differentially expressed gene sets using the Generalized Berk-Jones statistic. *Bioinformatics*.

**Gaynor, S.M.**, Schwartz, J., & Lin, X. (2019). Mediation analysis for common binary outcomes. *Statistics in Medicine*, 38(4), 512-529.

Borrelli, B., **Gaynor, S.**, Tooley, E., Armitage, C.J., Wearden, A., & Bartlett, Y.K. (2018). Identification of three different types of smokers who are not motivated to quit: Results from a latent class analysis. *Health Psychology*, 37(2), 179.

**Gaynor, S.**, & Bair, E. (2017). Identification of relevant subtypes via preweighted sparse clustering. *Computational Statistics & Data Analysis*, 116, 139-154.

Bair, E., **Gaynor, S.**, Slade, G.D., Ohrbach, R., Fillingim, R.B., Greenspan, J.D., ... & Maixner, W. (2016). Identification of clusters of individuals relevant to temporomandibular disorders and other chronic pain conditions: the OPPERA study. *Pain*, 157(6), 1266.

### Under Review and Preprints

**Gaynor, S.M.***, DiCorpo, D.*, Westerman, K., Russell, E., ... & Manning, A. Whole Genome Sequence Association Analysis of Fasting Glucose and Fasting Insulin Levels in Diverse Cohorts from the NHLBI TOPMed Program. medRxiv [Preprint]. January 4, 2021. Available from: https://doi.org/10.1101/2020.12.31.20234310.

Li, D., **Gaynor, S.M.**, Quick, C., Chen, J. T., ... & Lin, X. Identifying US Counties with High Cumulative COVID-19 Burden and Their Characteristics. medRxiv [Preprint]. January 12, 2021. Available from: https://doi.org/10.1101/2020.12.02.20234989.

McCaw, Z.R., **Gaynor, S.M.**, Sun, R., and Lin, X. Cross-tissue eQTL Calling via Surrogate Expression Analysis.

**Gaynor, S.M.**, Fagny, M., Lin, X., Platig, J., Quackenbush, J. Connectivity of variants in eQTL networks dictates reproducibility and functionality. bioRxiv 515551 [Preprint]. January 9, 2019. Available from: https://doi.org/10.1101/515551.

**Gaynor, S.M.**, Lin, X., Quackenbush, J. Spectral clustering in regression-based biological networks. bioRxiv 651950 [Preprint]. May 27, 2019. Available from: https://doi.org/10.1101/651950.