

# Nima Hejazi

✉ [nhejazi@berkeley.edu](mailto:nhejazi@berkeley.edu)  
🌐 [www.nimahejazi.org](http://www.nimahejazi.org)  
in [nshejazi](#)  
🐦 [nshejazi](#)  
🌐 [nhejazi](#)

## Curriculum Vitae

---

updated: July 2017

### Current Positions

*Doctoral Student*, Division of Biostatistics, University of California, Berkeley

Advisors: Mark J. van der Laan and Alan E. Hubbard

*Biomedical Big Data Predoctoral Fellowship Trainee*, Biomedical Big Data Training Program, University of California, Berkeley

### Research Interests

**Methodology** nonparametric statistics, causal inference, survival analysis, semiparametric theory, targeted learning, statistical machine learning, data adaptive estimation with inference, sieve analysis, statistical genomics, reproducible research, statistical computing.

**Applications** precision medicine, clinical trials, dynamic treatment regimes, epidemiology, computational biology and bioinformatics, healthcare policy.

### Education

2016–present **Doctor of Philosophy (Ph.D.) - Biostatistics**, *University of California, Berkeley*.  
Designated Emphasis (graduate minor) in Computational and Data Science and Engineering

2016–2017 **Master of Arts (M.A.) - Biostatistics**, *University of California, Berkeley*.  
*Committee*: Mark J. van der Laan (co-chair), Alan E. Hubbard (co-chair), Martyn T. Smith  
*Thesis*: Generalized Application of Empirical Bayes Statistics to Asymptotically Linear Parameters

2011–2015 **Bachelor of Arts (B.A.)**, *University of California, Berkeley*.  
Molecular & Cell Biology, Psychology, Public Health (triple major)

#### Accreditation, Open Education, and Training

2017-2018 **Biomedical Big Data Training Program**, *University of California, Berkeley*.

04–11/2016 **Genomic Data Science Specialization**, *Johns Hopkins University via Coursera*.

02–12/2015 **Data Science Specialization**, *Johns Hopkins School of Public Health via Coursera*.

### Research Experience

01/2017–**Graduate Student Researcher**, *University of California, Berkeley*, Berkeley, CA.

present Advisor: Prof. Mark J. van der Laan, Division of Biostatistics

- Development of techniques for data-adaptive estimation, nonparametric hypothesis testing, and causal inference for applications in vaccine sieve analysis, computational biology, and clinical trials.
- Software development for an ecosystem of tools designed to facilitate the statistical estimation of a diverse set of causal parameters within the framework of targeted minimum loss-based estimation.
- Large-scale, team-driven analysis of epidemiologic data, leveraging Targeted Learning to characterize factors contributing to and the effects of nutritional interventions on childhood “wasting”.

- 08/2015–  
present **Graduate Student Researcher**, *University of California, Berkeley*, Berkeley, CA.  
Advisor: Prof. Alan E. Hubbard, Division of Biostatistics
- Development of statistical methods for robust inference in high-dimensional biological problems, including data-adaptive methods for hypothesis testing and moderated statistics for targeted learning.
  - Statistical analysis of high-dimensional genomics and epigenetics data, as well as bioinformatical support for experimental work in molecular biology and environmental epidemiology.
- 03/2014–  
05/2015 **Research Assistant (undergraduate)**, *University of California, Berkeley*, Berkeley, CA.  
Supervised by Prof. Mark T. D’Esposito, Helen Wills Neuroscience Institute
- Study of the properties of resting-state voxel connectivity in the visual system using fMRI data.
- 05/2012–  
02/2015 **Research Assistant (undergraduate)**, *UCSF Medical Center*, San Francisco, CA.  
Supervised by Dr. Paul S. Larson, Department of Neurological Surgery
- Analysis of outcomes from deep brain stimulation surgery using interventional MR imaging.
- 08/2012–  
05/2014 **Research Assistant (undergraduate)**, *University of California, Berkeley*, Berkeley, CA.  
Supervised by Prof. Robert T. Knight, Helen Wills Neuroscience Institute
- Study of disruptions in lexical selection from neurological disease based on human EEG recordings.
- 08/2011–  
05/2013 **Research Assistant (undergraduate)**, *University of California, Berkeley*, Berkeley, CA.  
Supervised by Prof. Leslea J. Hlusko, Department of Integrative Biology
- Data collection for studies in quantitative genetics concerning the evolution of hominid dentition.

## Industry Experience

- 01/2016–  
03/2016 **Data Scientist (intern)**, *Cogitativo, Inc.*, Berkeley, CA.
- Applied modern statistical methods, data analytic tools, and data product design to problems in the healthcare services space, producing client-side tools for interactive data mining and visualization.

## Teaching Experience

- Aug. 2017 Module Instructor, R Bootcamp, Department of Statistics, University of California, Berkeley
- 2015–2016 Tutor in Mathematics & Statistics, Athletic Study Center, University of California, Berkeley
- 2013–2015 Tutor in General Chemistry, Student Learning Center, University of California, Berkeley

## Fellowships and Scholarships

- 2017-2018 Biomedical Big Data Predoctoral Fellowship, Biomedical Big Data Training Program (NIH BD2K), University of California, Berkeley
- 2017 Travel and Registration Scholarships, Summer Institute in Statistics for Big Data, Department of Biostatistics, University of Washington
- 2015 Travel and Registration Scholarships, Summer Institute in Biostatistics, Department of Biostatistics & Medical Informatics, University of Wisconsin — Madison

## Awards and Honors

- 2017 Honorable mention (rank 3 – 6), Thomas R. Ten Have Award for junior researchers, for the poster “Nonparametric variable importance for continuous exposures with applications in high-dimensional biology”, Atlantic Causal Inference Conference

## Conferences and Workshops

*Contributed or Invited*

- May 2017 Atlantic Causal Inference Conference; Chapel Hill, NC, USA  
July 2016 Statistical Causal Inference and Applications to Genetics, Centre de Recherches Mathématiques, Université de Montréal; Montréal, QC, Canada

*Attended or Visited*

- July 2017 Summer Institute for Statistics in Big Data, Department of Biostatistics, School of Public Health, University of Washington; Seattle, WA, USA  
July 2016 Computational Genomics Summer Institute, Institute for Pure and Applied Mathematics, University of California, Los Angeles; Los Angeles, CA, USA  
June 2016 BioC: Where Software and Biology Connect; Stanford, CA, USA  
June 2015 Summer Institute in Biostatistics, Department of Biostatistics & Medical Informatics, University of Wisconsin — Madison; Madison, WI, USA

---

## Professional Affiliations

- 2017–present Western North American Region of the International Biometrics Society (member)  
2016–present Berkeley Institute for Data Science (member)  
2016–present American Statistical Association (member)

---

## Computing Skills

- Systems: macOS (Mac OS X), Linux (Ubuntu, Red Hat), Chrome OS, UNIX  
Languages: R, PYTHON, Shell Scripting, L<sup>A</sup>T<sub>E</sub>X, JULIA, MATLAB, SQL  
Apps/Other: Git, Amazon EC2, Microsoft Office

---

## Publications

### Refereed Journal Publications

- [1] W. Cai, A. Hubbard, and **N. Hejazi**. adaptest: Data-adaptive statistics for high-dimensional testing. *The Journal of Open Source Software*, submitted, 2017.
- [2] J. R. Coyle and **N. S. Hejazi**. origami: A generalized framework for cross-validation in R. *The Journal of Open Source Software*, submitted, 2017.
- [3] **N. Hejazi**, C. Schiffman, and X. Zhou. Review of “statistical analysis of numerical preclinical radiobiological data”. *ScienceOpen Research*, 2017.
- [4] **N. S. Hejazi**, W. Cai, and A. E. Hubbard. biotmle: Targeted learning for biomarker discovery. *The Journal of Open Source Software*, 2(15), 2017.

### Manuscripts in Preparation

- [1] W. Cai, **N. S. Hejazi**, and A. E. Hubbard. Data adaptive statistical procedures for multiple hypothesis testing in high-dimensional settings.
- [2] **N. S. Hejazi** and N. P. Jewell. Immortal time bias analyses: Time-varying survival regression models and delayed entry approaches.
- [3] **N. S. Hejazi**, S. Kherad-Pajouh, M. J. van der Laan, and A. E. Hubbard. Variance moderation for estimators of asymptotically linear parameters by empirical Bayes shrinkage.

### Dissertations and Theses

- [1] **N. S. Hejazi**. Generalized application of empirical Bayes statistics to asymptotically linear parameters. Master’s thesis, University of California, Berkeley, 2017.

---

## Software

### Open Source Contributions

- [1] D. C. Benkeser and **N. S. Hejazi**. survtmle: Compute targeted minimum loss-based estimates in right-censored survival settings. <https://github.com/benkeser/survtmle>, 2017.
- [2] J. R. Coyle and **N. S. Hejazi**. origami: Generalized framework for cross-validation. <https://github.com/jeremyrcoyle/origami>, 2017.
- [3] **N. S. Hejazi** and A. E. Hubbard. biotmle: Targeted learning for biomarker discovery with moderated statistics. <https://github.com/nhejazi/biotmle>, 2017.

---

## Presentations

### Talks, Invited and Contributed

- [1] **N. Hejazi**. Empirical Bayes moderation of asymptotically linear parameters. Biostatistics Seminar Series, Division of Biostatistics, UC Berkeley, March 2017.
- [2] **N. Hejazi**. Targeted biomarker discovery. Graduate student admit day, Division of Biostatistics, UC Berkeley, March 2017.
- [3] **N. Hejazi** and E. Muzzall. Ensemble machine learning with Super Learner and h2o in R. The Hacker Within, Berkeley Institute for Data Science, UC Berkeley, December 2016.

- [4] A. E. Hubbard, **N. Hejazi**, W. Cai, and A. Decker. Targeted learning for high-dimensional variable importance. Montréal, QC, Canada, July 2016. Statistical Causal Inference and Applications to Genetics, Centre de Recherches Mathématiques.

*Posters, Invited and Contributed*

- [1] **N. Hejazi** and A. E. Hubbard. Super Learner for optimal prediction. Montréal, QC, Canada, July 2016. Statistical Causal Inference and Applications to Genetics, Centre de Recherches Mathématiques.
- [2] **N. S. Hejazi**, I. Malenica, A. K. Waschka, A. E. Hubbard, and M. J. van der Laan. Non-parametric variable importance for continuous exposures with applications in high-dimensional biology. Chapel Hill, NC, USA, May 2017. Atlantic Causal Inference Conference.