Jessica Gronsbell

Contact Information

Verily Life Sciences
355 Main Street
Cambridge, MA 02142

Phone: (510) 599-6722
E-mail: jgronsbell@google.com

Education

Harvard University Ph.D., Biostatistics, May 2017
Thesis Title: Robust and Efficient Machine Learning Methods for the Analysis of Electronic Medical Records Data
Advisor: Professor Tianxi Cai

University of California at Berkeley B.A., Applied Mathematics, May 2012

Professional Experience

Verily Life Sciences (an Alphabet company) Cambridge, Massachusetts
Data Scientist **Oct 2018 -**
Health Informatics Intern **June - Sep 2018**

Frontier Science Boston, Massachusetts
Statistical Consultant **Nov 2016 - June 2017**

MIT Lincoln Laboratory Lexington, Massachusetts
Intern **May – Aug 2015**
Intern **May – Aug 2013**

Academic Experience

Stanford School of Medicine Stanford, California
Department of Biomedical Data Science **June 2017 - May 2018**
Postdoctoral Research Fellow
Developed methods for meta-analysis and the analysis of electronic health records (EHR) data.

Harvard T.H. Chan School of Public Health Boston, Massachusetts
Department of Biostatistics **July - Aug 2016**
Instructor
Redesigned and delivered a statistical methods and computing course for incoming Ph.D. students.

University of Bordeaux Bordeaux, France
Department of Statistics in Systems Biology **June - July 2016**
Visiting Scholar
Developed weakly supervised learning methods for predictive modeling with EHR data.

Harvard T.H. Chan School of Public Health Boston, Massachusetts
Department of Biostatistics **July 2013 - Dec 2016**
Teaching Assistant
Assigned to a masters-level biostatistics course and an introductory biostatistics course.

- BIST222: Basics of Statistical Inference (Fall 2013, 2014, 2015, 2016)
- BIST203: Principles of Biostatistics II (Summer 2013)

University of California at Berkeley Berkeley, California
Democratic Education at Cal **Dec 2009 - May 2012**
Instructor
Designed and taught an outreach-based course to assist high schoolers with college applications.

Papers Accepted or Invited for Revision

Gronsbell J*, Liu M*, Tian L, and Cai T. Efficient Estimation and Evaluation of Prediction Rules in Semi-Supervised Settings under Stratified Sampling. *Under revision at the Journal of the American Statistical Association.*

Gronsbell J, Hong C, Lie N, Lu Y, and Tian L. Exact Inference for the Random-Effects Model for Meta-Analyses with Rare Events. *Accepted at Statistics in Medicine*.

Liao KP, Sun J, Cai T, Link NB, Hong C, Huang J, Huffman J, **Gronsbell J**, Zhang Y, Ho Y-L, Casto VM, Gainer VS, Murphy SN, O'Donnell CJ, Gaziano JM, Cho K, Szolovits P, Kohane IS, Yu S, and Cai T. (2019) High-throughput Multimodal Automated Phenotyping (MAP) with Application to PheWAS. *Journal of the American Medical Informatics Association*, 26(11), 1255-1262.

Gronsbell J*, Minnier J*, Yu S, Liao KP, and Cai T. (2019) Automated Feature Selection of Predictors in Electronic Medical Records Data. *Biometrics*, 75(1), 268-277.

Gronsbell J, and Cai T. (2018) Semi-Supervised Approaches to Efficient Evaluation of Model Prediction Performance. *Journal of the Royal Statistical Society: Series B*, 80(3), 579-594.

Yu S, Ma Y, **Gronsbell J**, Cai T, Ananthakrishnan AN, Gainer VS, Churchill SE, Szolovits P, Murphy SN, Kohane IS, Liao KP, and Cai T. (2018) Enabling phenotypic big data with PheNorm. *Journal of the American Medical Informatics Association*, 25(1), 54-60.

Aghayev A, Giannopoulos A, **Gronsbell J**, George E, Cai T, Steigner ML, Mitsouras D, and Rybicki FJ. (2018) Common First-Pass CT Angiography Findings Associated with Rapid Growth Rate in Abdominal Aorta Aneurysms between 3 and 5 cm in Largest Diameter. *American Journal of Roentgenology*, 210(2), 431-437.

Geva A, **Gronsbell J**, Cai T, Cai T, Murphy SN, Lyons JC, Heinz MM, Natter MD, Patibandla N, Bickel J, Mullen MP, and Mandl KD. (2017) A Computable Phenotype Improves Cohort Ascertainment in a Pediatric Pulmonary Hypertension Registry. *The Journal of Pediatrics*, 188(5), 224-231.

Kumamaru H, Kumamaru K, Bateman BT, **Gronsbell J**, Cai T, Liu J, Higgins LD, Aoki S, Ohtomo K, Rybicki FJ, and Patorno E. (2016) Limited Hospital Variation in the Utilization and Yield of CT for Pulmonary Embolism in Patients Undergoing Total Hip or Knee Replacement Surgery. *Radiology*, 281(3), 826-834

Working Papers

Gronsbell J, Hong C, Gainer VS, Zhang Y, Geva A, Mandl K, Yu S, Liao KP, and Cai T. ssROC: Efficient Validation of EHR Algorithms with Semi-Supervised Learning.

Hong C, **Gronsbell J**, Castro VM, Gainer VS, Zhang Y, Murphy SN, Weiss S, Liao KP, and Cai T. Weakly Supervised Phenotyping using Patient-Reported Outcomes as Silver-Standard Labels.

Gronsbell J, Lie N, Lu Y, and Tian L. Exact Inference for Random-Effects Meta-Analysis of the Risk Difference.

Conference Papers

Attarian A, Danis G*, **Gronsbell J***, Iervolino G*, Layne L, Padgett D, and Tran H. (2013). Baseball Pitch Classification: A Bayesian Method and Dimension Reduction Investigation. *IAENG Transactions on Engineering Sciences: Special Issue of the International MultiConference of Engineers and Computer Scientists and World Congress on Engineering*, 393-399.

Attarian A, Danis G*, **Gronsbell J***, Iervolino G*, Layne L, Padgett D, and Tran H. A Comparison of Feature Selection and Classification Algorithms in Identifying Baseball Pitches. *Proceedings of the International MultiConference of Engineers and Computer Scientists: Lecture Notes in Engineering and Computer Science*, March 13-15, 2013, Hong Kong, 263-268.

*Indicates equal contribution as first authors.

Honors and Awards

Gertrude M. Cox Scholarship Honorable Mention, American Statistical Association, 2016

MIT IMPACT Fellow, 2016

Young Researchers Award, International Society of Nonparametric Statistics, 2016

Rose Traveling Fellowship, Harvard T.H. Chan School of Public Health, 2015

Phi Beta Kappa Honors Society, 2012

High Distinction in General Scholarship, UC Berkeley, 2012

High Honors in Mathematics, UC Berkeley, 2012

Biology Scholars Program, UC Berkeley, 2009

Pre-Medical Honors Society, UC Berkeley, 2009

Grant Support

NIH F31 GM119263 Semi-Supervised Learning with Electronic Medical Records, 2016-2017 (PI)
NIH T32 NS048005-11 Training in Neurostatistics and Neuroepidemiology, 2014-2016 (Trainee)
NIH T32 AI007358-26 Biostatistics/Epidemiology Training Grants in AIDS, 2013-2014 (Trainee)

Invited Talks

My Path to Biostatistics. Data Science in Action Course at Harvard Biostatistics. Boston, Massachusetts. August 2019.

An Introduction to Digital Phenotyping. TechTalk at Verily Life Sciences. Cambridge, Massachusetts. September 2018.

Statistical Learning Methods for Efficient EHR-based Phenotyping. TechTalk at Verily Life Sciences. Cambridge, Massachusetts. June 2018.

My Path to Biostatistics. Career Development Seminar at Harvard Biostatistics. Boston, Massachusetts. October 2017.

Applying for the NIH F31 Predoctoral Fellowship. Neurostatistics Working Group at Harvard Biostatistics. Boston, Massachusetts. October 2016.

Weakly Supervised Prediction Modeling with EHR Data. Bioinformatics Working Group at University of Bordeaux. Bordeaux, France. July 2016.

Efficient Estimation of Prediction Performance Measures in Semi-Supervised Settings. Biostatistics Working Group at University of Bordeaux. Bordeaux, France. June 2016.

Efficient Estimation of Prediction Performance Measures in Semi-Supervised Settings. Neurostatistics Working Group at Harvard Biostatistics. Boston, Massachusetts. October 2015.

Applying for the NIH F31 Predoctoral Fellowship. Neurostatistics Working Group at Harvard Biostatistics. Boston, Massachusetts. September 2015.

Adaptive Radar Waveform Design. Technical Presentation at MIT Lincoln Laboratory. Lexington, Massachusetts. July 2015.

Radar Classification Bounds. Technical Presentation at MIT Lincoln Laboratory. Lexington, Massachusetts. July 2013.

Contributed Talks

Challenges and Rewards of Working with Healthcare Data. Richard Tapia Celebration of Diversity in Computing. San Diego, California. September 2019.

Leveraging Technology to Improve Healthcare. Richard Tapia Celebration of Diversity in Computing. San Diego, California. September 2019.

Efficient Estimation of Prediction Performance Measures in Semi-Supervised Settings. Eastern North Atlantic Region Spring Meeting. Washington, District of Columbia. March 2017.

Efficient Estimation of Prediction Performance Measures in Semi-Supervised Settings. International Society of Nonparametric Statistics Conference. Avignon, France. June 2016.

Professional Service

Member, Google K-12 Visit Committee, 2019-
Member, Google EngEdu, 2019-
Organizer, Harvard Biostatistics Student Seminar, 2016-2017
Chair, Harvard Biostatistics Student Committee, 2015-2017
Director, Prepare to Achieve a College Education, 2008-2012

Professional References

Tianxi Cai, Professor
Departments of Biostatistics and Biomedical Informatics
Harvard T.H. Chan School of Public Health and Harvard Medical School
email: tcai@hsph.harvard.edu

Lu Tian, Professor
Department of Biomedical Data Science
Stanford School of Medicine
email: lutian@stanford.edu

Menachem Fromer, Adjunct Professor
Department of Psychiatry
Icahn School of Medicine at Mount Sinai
Head of Mental Health Data Science and R&D at Verily Life Sciences
email: fromer@google.com

Paul Varghese, Head of Health Informatics
Verily Life Sciences
email: paulvarghese@google.com

Teaching Reference

David Wypij, Senior Lecturer in Biostatistics
Department of Biostatistics
Harvard T.H. Chan School of Public Health
email: wypij@hsph.harvard.edu

Prepared: November 17, 2019