Adler Perotte, MD, MA

Columbia University
Department of Biomedical Informatics
622 West 168th Street, PH20
New York, NY 10032

Office: (212) 342-1633
Email: adler.perotte@columbia.edu

Research Overview

My primary research areas are Bayesian inference and prediction related to electronic health record data and metabolomics. My research interests are both in the development of novel statistical machine learning methods with broad applicability across domains and the use of such methods towards solving biomedical problems. Specific areas of interest include probabilistic phenotyping, statistical natural language processing, time series modeling, cohort selection, causal inference and analysis of high frequency monitoring data.

As a collaborator in the Observation Health Data Sciences and Informatics (OHDSI) collaborative I have been involved in the design and implementation of the infrastructure and methods for this international, multidisciplinary research network which in its decentralized model currently hosts more than 600 million patient records.

Education

M.A. Biomedical Informatics, Columbia University, 2012.

M.D. Columbia University, College of Physician and Surgeons, 2008.

A.B. Psychology, Neuroscience Certificate, *Magna Cum Laude*, Princeton University, 2004.
Thesis: A Solution to Catastrophic Interference: A Neural Network Model of Consolidation and Cortical Memory Protection During Sleep.

Positions

Assistant Professor, Department of Biomedical Informatics, Columbia University 2016–present.

Associate Research Scientist, Department of Biomedical Informatics, Columbia University 2013–2016.

Assistant Director for Technology, Center for Advanced Technology, Columbia University 2012–present.

Adjunct Associate Research Scientist, Department of Biomedical Informatics, Columbia University 2012–2103.

National Library of Medicine Postdoctoral Fellow (George Hripcsak, Frank Wood), Columbia University 2009–2012.

Reseach Specialist, Computational Memory Laboratory (Kenneth Norman, David Blei), Princeton University 2008–2009.

Medical Student Intern, Science and Technology Ventures, Columbia University, 2007.

Short-term Research Fellow, Center for Theoretical Neuroscience (Kenneth Miller), Columbia University 2005.

Research Assistant, Mechanical and Aerospace Engineering (Wale Soboyejo), Princeton University 2001.

Research Assistant, Computational Memory Laboratory (Kenneth Norman), Princeton University 2000.

Professional Organizations and Societies

Founder and Faculty Advisor, Columbia University Health Tech Assembly (www.healthtechassembly.com), 2012–present

Coordinating Center Collaborator, Observational Health Data Sciences and Informatics (OHDSI), 2014–present

Collaborator, New York City Clinical Data Research Network (NYC-CDRN), 2014–present

Member, American Medical Informatics Association 2009–2013, 2015–present

Member, Student National Medical Association 2004–2008

Honors

American Medical Informatics Association Distinguished Paper Reviewer, 2013

Malcolm X Scholar (merit-based tuition scholarship), Columbia University, College of Physicians and Surgeons, 2004–2008

J.F. Bohmfalk Scholar (merit-based tuition scholarship), Columbia University, College of Physicians and Surgeons, 2004–2008

Committees

College of Physicians and Surgeons Biomedical Informatics Thread Steering Committee at Columbia University, 2014–2016

University Research Misconduct Standing Committee at Columbia University, 2007–2008

Search Committee for the Dean of the Faculty of Medicine and Executive Vice President of Health and Biomedical Sciences at Columbia University, 2005–2006

Fellowships and Grant Support

Principal Investigator - Defense Advanced Research Projects Agency, BTO - "Combining clinical data and transdermal analysis of whole blood to create a next generation wearable device," 2016 – present

Assistant Director for Technology - New York State Foundation for Science, Technology and Innovation (NYSTAR) C070123 - "Center for Advanced Information Management," 2012–Present

Investigator - Astrazeneca, ASTRZEN CU15-1758 - "New p-value calibration based on the findings from the distributions," 2015–2015

Investigator - Astrazeneca, ASTRZEN CU15-1759 - "Semantic Browser," 2015–2015

Local Principal Investigator - National Institutes of Health, 1 R41 HL126568-01A1 - "Vulnerable Plaque Optical Analyzer," 2014–2015

Investigator - Patient Centered Outcomes Research Institute - "New York City Clinical Data Research Network," 2014–2015

Investigator - National Science Foundation, IIS-1344668 - "SCH INT: Large-Scale Probabilistic Phenotyping Applied to Patient Record Summarization," 2014–2014

Investigator - Telemedicine & Advanced Technology Research Center - "Pattern Representation and Evaluation of Data through Integration, Correlation, and Transformation," 2012-2013

Postdoctoral Fellow - National Library of Medicine, T15 LM007079-14 "Training in Biomedical Informatics at Columbia University," 2009–2012

Teaching

Guest Lecturer, BINF G4001: Introduction to Machine Learning, Biomedical Informatics, Columbia University, Fall 2016

Course Director, BINF G4099: Research Seminar in Biomedical Informatics, Biomedical Informatics, Columbia University, Fall 2016 –

Course Director, BINF G8001: Readings in Biomedical Informatics - Probabilistic Graphical Models, Biomedical Informatics, Columbia University, Spring 2016

Course Director, BINF G4000: Acculturation to Programming and Statistics, Biomedical Informatics, Columbia University, Fall 2015

Guest Lecturer, Privacy and Security in Biomedical Informatics, Columbia University College of Physicians and Surgeons, 2015 – 2016

Guest Lecturer, Decision Analysis in Biomedical Informatics, Columbia University College of Physicians and Surgeons, 2014–2015

Guest Lecturer, Principal Components Analysis in Computational Methods, Columbia University College of Physicians and Surgeons, 2012–2014

Teaching Assistant, BINF G4002: Computational Methods, Biomedical Informatics, Columbia University, 2011

Postdoctoral Curriculum Developer, Cornell/Columbia Health Informatics Certificate Program, 2010

Teaching Assistant, BINF G4003: Symbolic Methods, Biomedical Informatics, Columbia University, 2010

Teacher, Student Success Network, Columbia University College of Physicians and Surgeons, 2005–2006

Teaching Assistant, Organic Chemistry II, Chemistry, Princeton University, 2003

Teaching Assistant, Organic Chemistry I, Chemistry, Princeton University, 2003

Other Professional Activities

Scientific Programming Committee Member, American Medical Informatics Association, 2016

Member, Biomedical Informatics Data Mining Group, Columbia University, 2009–present

Reviewer, Journal of American Medical Informatics Association, 2012–present

Reviewer, Journal of Biomedical Informatics, 2015–present

Reviewer, Journal of Biomedical Semantics, 2016–present

Reviewer, PeerJ, 2016–present

Reviewer, Statistics in Medicine, 2015–present

Reviewer, American Medical Informatics Association Annual Symposium, 2012-2016

Reviewer, BMC Medical Informatics and Decision Making, 2014-2015

Member, ICML: ML for Clinical Data Anlaysis Workshop Program Committee, 2012

Organizer, Health 2.0/Columbia Code-a-thon, 2012

InSITE Fellow and Consultant for deTect Biosciences LLC, imup4, and BestVendor, 2010–2012

Founder, Biomedical Informatics Machine Learning Reading Group, 2009–2011

Research advisees

Tian Kang, PhD Student Rotation, Spring 2017

Elliot Mitchell, PhD Student Rotation, Spring 2017

Guillaume David, Postdoctoral Research Scientist, 2016 – present

Liana Tascau, Masters Student, 2016 – present

Joongheum Park, Medicine Resident Research Intern, 2016 – present

Aras Curukcu, High School Intern, Summer 2016

Amelia Averitt, PhD Student, 2016 – present

Publications

Journal Articles

Hripcsak G, Ryan P, Duke J, Shah NH, Park RW, Huser V, Suchard MA, Schuemie M, DeFalco F, **Perotte A**, Banda J, Reich C, Schilling L, Matheny M, Meeker D, Pratt N, Madigan D. Addressing Clinical Questions at Scale: OHDSI Assessment of Treatment Pathways. Proc Natl Acad Sci U S A. 2016 Jul 5;113(27):7329-36. doi: 10.1073/pnas.1510502113.

Cole-Lewis H, **Perotte A**, Galica K, Dreyer L, Griffith C, Schwarz M, Yun C, Patrick H, Coa K, Augustson E. Social Network Behavior and Engagement Within a Smoking Cessation Facebook Page. J Med Internet Res. 2016 Aug 2;18(8):e205. doi: 10.2196/jmir.5574.

Pivovarov R, **Perotte A**, Grave E, Angiolillo J, Wiggins C, Elhadad N. Learning Probabilistic Phenotypes from Heterogeneous EHR Data. *J Biomed Inform.*, 2015 Oct 14. pii: S1532-0464(15)00223-3. doi: 10.1016/j.jbi.2015.10.001.

Hripcsak G, Albers DJ, **Perotte A**. Parameterizing time in electronic health record studies. *J Am Med Inform Assoc.* 2015 Feb 26. pii: ocu051. doi: 10.1093/jamia/ocu051. **Editor's Choice**

Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. *J Am Med Inform Assoc.* 2015 Apr 20. pii: ocv024. doi: 10.1093/jamia/ocv024. **AMIA Journal Club Invited Talk**

Hernández D, Carrión D, **Perotte A**, Fullilove R. Public health entrepreneurs: training the next generation of public health innovators. *Public Health Reports* 2014 Nov-Dec;129(6):477-81.

Albers DJ, Elhadad N, Tabak E, **Perotte A**, Hripcsak G. Dynamical phenotyping: using temporal analysis of clinically collected physiologic data to stratify populations. *PLoS One* 2014 Jun 16;9(6):e96443. doi: 10.1371/journal.pone.0096443. 2014.

Perotte A, Pivovarov R, Natarajan K, Weiskopf N, Wood F, Elhadad N. Diagnosis code assignment: models and evaluation metrics. *J Am Med Inform Assoc.*, Dec 2013. doi: 10.1136/amiajnl-2013-002159

Overby C, Pathak J, Gottesman O, Haerian K, **Perotte A**, Murphy S, Bruce K, Johnson S, Talwalker J, Shen Y, Ellis S, Kullo I, Chute C, Friedman C, Bottinger E, Hripcsak G, Weng C. A Collaborative Approach to Developing an Electronic Health Record Phenotyping Algorithm for Drug-Induced Liver Injury. *J Am Med Inform Assoc.*, Dec 2013;20(e2):e243-52. doi:10.1136/amiajnl-2013-001930

Perotte A and Hripcsak G. Temporal Properties of Diagnosis Code Time Series in Aggregate. *IEEE Transactions of Information Technology in Biomedicine*, March 2013:17(2):477-483.

Claassen J, **Perotte A**, Albers D, Kleinberg S, Schmidt JM, Tu B, Badjatia N, Lantigua H, Hirsch LJ, Mayer SA, Connoly ES, Hripcsak G. Nonconvulsive seizures after subarachnoid hemorrhage: multimodality detection and outcomes. *Annals of Neurology*, 2013 Jul;74(1):53-64. doi: 10.1002/ana.23859.

Hripcsak G, Albers D, **Perotte A**. Exploiting Time in Electronic Health Record Correlations. *J Am Med Inform Assoc.* 2011 Dec;18 Suppl 1:i109–15.

Norman K, Newman E, **Perotte A**. Methods for reducing interference in the Complementary Learning Systems Model: Oscillating Inhibitation and Autonomous Memory Rehearsal. *Neural Networks* 2005:18(9):1212–28. **Special Issue: Computational Theories of the Functions of the Hippocampus**

Paper Proceedings

Ranganath R, **Perotte A**, Elhadad N, Blei D. Deep Survival Analysis. To appear in the 1st Conference on Machine Learning and Health Care (MLHC), Aug. 2016.

Ranganath R, **Perotte A**, Elhadad N, Blei D. The Survival Filter: Joint Survival Analysis with a Latent Time Series. *Uncertainty in Artificial Intelligence*, Amsterdam, Netherlands, 2015.

Overby C, Weng C, Haerian K, **Perotte A**, Hripcsak G. Evaluation Considerations for EHR-based Phenotyping Algorithms: A Case Study for Drug Induced Liver Injury, *AMIA Summit on Translational Bioinformatics*, 2013.

Hripcsak G, Albers D, **Perotte A**. Interpreting Lagged Linear Correlation and Using Range to Prioritize. *AMIA Summit on Translational Bioinformatics*, 2012.

Perotte A, Bartlett N, Elhadad N, Wood F. Hierarchically Supervised Latent Dirichlet Allocation. *Advances in Neural Information Processing Systems 24 (NIPS 2011)* 2609–2617.

Hripcsak G, Albers D, **Perotte A**. Using lagged linear correlation to find relationships between laboratory values and clinician concepts. *AMIA Summit on Translational Bioinformatics*, 2011.

Socher R, Gershman S, **Perotte A**, Sederberg P, Norman K, Blei D. A Bayesian Analysis of Dynamics in Free Recall. *Advances in Neural Information Processing Systems 22 (NIPS 2009)* 1714–1722.

Posters and Presentations

Perotte A, Elhadad N. A probabilistic model for learning relationships between diagnosis codes and clinical free text. *AMIA Annual Symposium*, 2016.

Perotte A. Scientific Entrepreneurship. Multidisciplinary Patient Oriented Research (MPOR) Colloquium, Irving Institute for Clinical and Translational Research, November 22, 2016.

Perotte A. El futuro de la innovacion en Chile y los equipos interdisciplinarios. Columbia University Global Center. University of Chile. Santiago, Chile. October 27, 2016.

Perotte A. Predicting Kidney Disease Progression with Large-Scale Patient Data. Data Science Day. Columbia University Data Science Institute. New York. 2016.

Ranganath R, **Perotte A**, Elhadad N, Blei D. The Survival Filter. NSF Workshop: Data Science, Learning, and Applications to Biomedical & Health Sciences. New York. 2016.

Perotte A, Ranganath R, Hirsch JS, Blei D, Elhadad N. Combining traditional statistical methods and machine learning methods for risk prediction in chronic kidney disease progression using electronic health record data. AMIA Journal Club, 2015.

Schmidt M, Claassen J, **Perotte A**, Albers D, Hripcsak G. Understanding and Visualizing Heterogeneous High Frequency Data in the Neurological ICU. Inaugural Symposium: From Big Data to Big Ideas. Institute for Data Science and Engineering, Columbia University, 2013.

Claassen J, **Perotte A**, Albers D, Schmidt J, Tu B, Badjatia N, Lee K, Mayer S, Connolly E, Hirsch L, Hripcsak G. Electrographic seizures after subarachnoid hemorrhage lead to derangement of brain homeostasis in humans. *Critical Care 2011*, 15(Suppl 1):P331 (doi: 10.1186/cc9751)

Maurer MS, Albers D, **Perotte A**, Chen C, Hripcsak G. Hemoconcentration is Associated with Lower Mortality Post Hospitalization for Heart Failure. *American College of Cardiology Annual Scientific Session & Expo*, 2012.

Perotte A, Bartlett N, Elhadad N, Wood F. Hierarchically Supervised Latent Dirichlet Allocation. *New York Academy of Sciences Sixth Annual Machine Learning Symposium*, 2011.

Perotte A, Hripcsak G. Using Density Estimates to Aggregate Patients and Summarize Disease Evolution. *AMIA Summit on Translational Bioinformatics*, 2011.

Perotte A. Characterization of Disease Time Course using ICD9 codes. *National Library of Medicine Informatics Training Conference*, 2011.

Perotte A, Hripcsak G. Using the Entropy of ICD9 Documentation Across Patients to Characterize Disease Chronicity. *AMIA Annual Symposium Proceedings*. 2010.

Perotte A. Characterization of Disease Time Course Using ICD9 Codes. National Library of Medicine Training Conference, 2011.

Book Chapters

Wood F, **Perotte A**. Mixed Membership Classification for Documents with Hierarchically Structured Labels. *Handbook on Mixed Membership Models*. Ed. Airoldi EM, Blei D, Erosheva EA, and Fienberg E. Chapman and Hall/CRC 2014. 305âĂŞ323. Print.

Last updated: February 1, 2017