PEDRO L. BALDONI

baldoni@email.unc.edu ◇ https://plbaldoni.rbind.io
Department of Biostatistics, University of North Carolina at Chapel Hill
135 Dauer Dr, Chapel Hill, NC 27599

EDUCATION

Ph.D. Candidate, Biostatistics *2014 - present*
Department of Biostatistics
University of North Carolina at Chapel Hill, USA
Advisors: Dr. Naim U. Rashid, Dr. Joseph G. Ibrahim
Dissertation Committee: Dr. Michael I. Love, Dr. Yun Li, Dr. Douglas Phanstiel
Research area: Epigenomics, mixed models, mixture regression, deep learning.
Expected graduation: July 2020

M.S., Statistics *2014*
Department of Statistics
University of Campinas, Brazil
Advisor: Dr. Hildete P. Pinheiro
International exchange program: Aarhus University, Denmark (August/2013 - February/2014)
Link to thesis: http://repositorio.unicamp.br/jspui/handle/REPOSIP/307180

B.S., Statistics *2011*
Department of Statistics
University of Campinas, Brazil

PROFESSIONAL POSITIONS

Graduate Research Assistant *2015 - 2019*
University of North Carolina at Chapel Hill

Collaborative Studies Coordinating Center
Hispanic Community Health Study/Study of Latinos (HCHS/SOL)
Supervisors: Dr. Jianwen Cai, Dr. Daniela Sotres-Alvarez

Center for AIDS Research (CFAR)
Supervisor: Dr. Michael G. Hudgens

Statistician *2012*
CPqD Foundation, Campinas, Brazil

PROFESSIONAL ACTIVITIES

Professional Memberships

· American Statistical Association (ASA), 2018 - present

· Eastern North American Region (ENAR), 2017 - present

Program Development

· Session Chair, Eastern North American Region (ENAR), 2019. Topic: Replicability in Big Data Precision Medicine.

HONORS, AWARDS, SCHOLARSHIPS

· University Cancer Research Fund Award - Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill (2019-2020)

· Max Halperin Award (Excellence in graduate studies) - Department of Biostatistics, University of North Carolina at Chapel Hill (2016)

· Science Without Borders (Full Ph.D. program) - CAPES, Brazil (2014-2018)

· International Exchange Program (Aarhus University, Denmark) - Santander Bank (2013-2014)

· CAPES Scholarship (Full MS program) - University of Campinas, Brazil (2012-2014)

PEER-REVIEWED PUBLICATIONS

1. Mollan, K. R., Trumble, I. M., Reifeis, S. A., Ferrer, O., Bay, C. P., **Baldoni, P. L.**, & Hudgens, M. G. (2020). Precise and accurate power of the rank-sum test for a continuous outcome. *Journal of Biopharmaceutical Statistics*. https://doi.org/10.1080/10543406.2020.1730866

2. **Baldoni, P. L.**, Rashid, N. U., & Ibrahim, J. G. (2019). Improved Detection of Epigenomic Marks With Mixed Effects Hidden Markov Models. *Biometrics*, 75 (4): 140113. https://doi.org/10.1111/biom.13083

3. Trumble, I. M., Allmon, A. G., Archin, N. M., Rigdon, J., Francis, O., **Baldoni, P. L.**, & Hudgens, M. G. (2017). SLDAssay: a software package and web tool for analyzing limiting dilution assays. *Journal of immunological methods*, 450, 10-16., 450, pp.10-16. https://doi.org/10.1016/j.jim.2017.07.004

4. Lee, S. K., Zhou, S., **Baldoni, P. L.**, Spielvogel, E., Archin, N. M., Hudgens, M. G., ... & Swanstrom, R. (2017). Quantification of the latent HIV-1 reservoir using ultra deep sequencing and primer ID in a viral outgrowth assay. *Journal of acquired immune deficiency syndromes* (1999), 74(2), 221. https://doi.org/10.1097/QAI.0000000000001187

5. Clutton, G., Xu, Y., **Baldoni, P. L.**, Mollan, K. R., Kirchherr, J., Newhard, W., ... & Archin, N. (2016). The differential short-and long-term effects of HIV-1 latency-reversing agents on T cell function. *Scientific reports*, 6, 30749., 6. https://doi.org/10.1038/srep30749

SUBMITTED MANUSCRIPTS

1. **Baldoni, P. L.**, Rashid, N.U., & Ibrahim, J.G.. Efficient Detection and Classification of Epigenomic Changes Under Multiple Conditions. *Journal of the American Statistical Association (under review)*. https://www.biorxiv.org/content/10.1101/864124v1

2. Elfassy T., Zeki Al Hazzouri A., Cai J., **Baldoni P. L.**, Llabre M., Rundek T., Raij L., Lash J., Talavera G. A., Wasserthiel-Smoller S., Daviglus M., Booth J. N. III, Castaneda

S., Garcia M., & Schneiderman N.. Incidence of hypertension among US Hispanic/Latinos: the Hispanic Community Health Study/Study of Latinos, 2008-2017. 2019. *Journal of the American Heart Association (revise and resubmit)*

3. Kaplan, R., **Baldoni, P.L.**, Strizich, G.M., Perez-Stable, E.J., Saccone, N.L., Peralta, C.A., Perreira, K.M., Gellman, M.D., Williams-Nguyen, J., Rodriguez, C., Lee, D.J., Daviglus, M., Talavera, G.A., Lash, J.P., Cai, J. & Franceschini, N.. Current smoking but not former smoking is a risk factor for incident hypertension. Six year follow up of the Hispanic Community Health Study Study of Latinos. *Hypertension (under review)*

WORKING MANUSCRIPTS

1. **Baldoni, P. L.**, Rashid, N.U., & Ibrahim, J.G.. Single-cell ChIP- and ATAC-seq Modeling via Deep Learning Variational Autoencoder.

2. **Baldoni, P. L.**, Sotres-Alvarez, D., Lumley, T., & Shaw, P.A.. On the use of Regression Calibration in a Complex Survey Design with Application to the Hispanic Community Health Study/Study of Latinos.

3. Cai, J., Zeng, D., Butera, N.M., **Baldoni, P. L.**, Maitra, P., & Dong, L.. Comparisons of Statistical Methods for Handling Attrition in Longitudinal Studies with Complex Survey Sampling.

4. Liese, A.D., Kaplan, R., Qi, Q., Thrasher, J.F., Sutherland, M.W., Lee, D.J., Thyagarajan, B., Talavera, G.A., **Baldoni, P. L.**, & Cai, J.. Effects of smoking and smoking cessation on incidence of diabetes in Hispanic/Latino populations in the US: Results from the Hispanic Community Health Study/Study of Latinos(HCHS/SOL).

PUBLICLY AVAILABLE SOFTWARE

1. **mixNBHMM**: detection and classification of differential enrichment regions from ChIP-seq experiments under multiple conditions https://github.com/plbaldoni/mixNBHMM

2. **ZIMHMM**: detection of broad enrichment regions from multiple ChIP-seq experimental replicates via a zero inflated mixed effects hidden Markov model https://github.com/plbaldoni/ZIMHMM

PRESENTATIONS

Scientific Meetings (Contributed)

· Efficient Detection and Classification of Epigenomic Changes Under Multiple Conditions, *Eastern North American Region (ENAR)*, March, 2020.

· Efficient Detection and Classification of Epigenomic Changes Under Multiple Conditions, *Joint Statistical Meetings (JSM)*, August, 2020.

· Detection and Classification of Changes in Protein-DNA Binding Activity With Applications in Diffuse ChIP-seq Data, *Joint Statistical Meetings (JSM)*, July, 2019.

· Integrative HMM With Mixture Model for Differential Pattern Detection of Broad Epigenomic Marks, *Eastern North American Region (ENAR)*, March, 2019.

· Improved Detection of Epigenomic Marks With Mixed Effects Hidden Markov Models, *Eastern North American Region (ENAR)*, March, 2018.

· A Statistical Method for the Analysis of Multiple ChIP-seq Datasets, *Joint Statistical Meetings (JSM)*, July, 2017.

· A Statistical Method for the Analysis of Multiple ChIP-seq Datasets, *Eastern North American Region (ENAR)*, March, 2017.

Invited Talks

· Laboratory for Statistical and Translational Genomics, *University of Pennsylvania*, January, 2020.

Other Meetings and Events

· On the use of Regression Calibration in a Complex Survey Design with Application to the Hispanic Community Health Study/Study of Latinos, *Collaborative Studies Coordinating Center, University of North Carolina at Chapel Hill*, November, 2019.

· Statistical Strategies for the Analysis of Diet-Disease Models that Correct for Error-Prone Exposures, *Collaborative Studies Coordinating Center, University of North Carolina at Chapel Hill*, May, 2010.

· Calculating and Comparing Age Standardized Cumulative Incidence of Hypertension across Hispanic/Latino Background Groups, *Collaborative Studies Coordinating Center, University of North Carolina at Chapel Hill*, April, 2019.

· Statistical methods for HIV-1 reservoir estimation in viral outgrowth assays, *Center for AIDS Research (CFAR), UNC Gillings School of Public Health*, June, 2016.

TEACHING EXPERIENCE

Recitation Lecturer
Department of Biostatistics, The University of North Carolina at Chapel Hill
BIOS 545 - Principles of Experimental Analysis (2016, Undergraduate level class)

Teaching Assistant
Department of Biostatistics, The University of North Carolina at Chapel Hill
BIOS 680 - Introductory Survivorship Analysis (2018, MS level class)
BIOS 735 - Introduction to Statistical Computing (2019, 2020, PhD level class)

Department of Statistics, University of Campinas, Brazil
ME607 - Time Series (2013, Undergraduate level class)

SKILLS AND SERVICE

· Computing skills: proficient in R, C++, and SAS.

· Languages: Portuguese (native speaker), English, and Italian.

· Student body representative in the Department of Statistics, University of Campinas, Brazil, 2013