Ryan Sun

Contact Information

Department of Biostatistics
University of Texas M.D. Anderson Cancer Center
1400 Pressler St, FCT4.5078
Houston, Texas 77030

Cell: (646) 753-2975
E-mail: rsun3@mdanderson.org
Website: https://ryanrsun.github.io/
Citizenship: United States, Canada

Education

Harvard University, Boston, Massachusetts

Ph.D., Biostatistics, May 2017

- Dissertation Title: "Methods for High-Dimensional Inference in Genetic Association Studies"
- Advisor: Professor Xihong Lin

M.A., Biostatistics, May 2014

Columbia University, New York City, New York

B.S., Applied Mathematics, May 2012

Professional Experience

Empiric Asset Management, New York, New York

Analyst **May 2011 - May 2012**

Quantitative analyst for hedge fund with equity market neutral strategy.

Honors and Awards

Travel Award for IMS New Researchers Conference, 2018

Harvard Program in Quantitative Genomics Postdoctoral Travel Award, 2017

ENAR Distinguished Student Paper Award, 2017

Harvard University Distinction in Teaching Award, 2014-2015

Cyprus National Government Environmental Health Travel Scholarship, 2013

Columbia University Lauren P. Breakiron Scholarship, 2008-2012

Academic Experience

University of Texas MD Anderson Cancer Center, Houston, Texas

Assistant Professor, Department of Biostatistics **July 2019 - Present**

Harvard School of Public Health, Boston, Massachusetts

Postdoctoral Research Fellow **May 2017 - June 2019**

Develop statistical methods for the analysis of high-dimensional biomedical data. Collaborations include the EPA Superfund project and the International Lung Cancer Consortium (ILCCO).

Harvard School of Public Health, Boston, Massachusetts

Instructor **July 2016 - August 2016**

Redesigned and delivered real analysis course for incoming Biostatistics Ph.D. students. Planned lectures, created class materials, and taught all sessions.

Teaching Assistant **September 2012 - June 2016**

Assigned to three doctoral-level biostatistics courses and one introductory statistics course.

- BIST245 Analysis of Multivariate and Longitudinal Data
- BIST232 Statistical Methods II (Distinction in Teaching award)
- BIST230 Probability Theory and Applications I (Distinction in Teaching award)
- BIO210 Rates and Proportions

PAPERS RECEIVING PEER-REVIEW

Sun, R.*, Bouchard, M.B.*, and Hillman, E.M.C. SPLASSH: Open source software for camera-based high-speed, multispectral in-vivo optical image acquisition. *Biomedical Optics Express* 2010; 1(2): 385-397.

Wang, Z., Claus-Henn, B., Wang, C., Wei, Y., Su, L., **Sun, R.,** Chen, H., Wagner, P.J., Lu, Q., Lin, X., Wright, R., Bellinger, D., Kile, M., Mazumdar, M., Tellez-Rojo, M.M., Schnaas, L., and Christiani, D.C. Genome-wide gene by Pb exposure interaction analysis identifies UNC5D as a candidate gene for neurodevelopment. *Environmental Health* 2017; 16(1): 81.

Sun, R., Carroll, R.J., Christiani, D.C., and Lin, X. Testing for gene-environment interaction under exposure misspecification. *Biometrics* 2018; 74(2): 653-662.

Orkaby, A., Rich, M.W., **Sun, R.,** Lux, E., Wei, L.J., and Kim, D.H. Pravastatin for primary prevention in older adults: restricted mean survival time analysis. *Journal of the American Geriatrics Society* 2018; 66(10): 1987-1991.

Sun, R., Wang, Z., Claus Henn, B., Su, L., Lu, Q., Lin, X., Wright, R., Bellinger, D., Kile, M., Mazumdar, M., Tellez-Rojo, M.M., Schnaas, L., and Christiani, D.C. Identification of novel loci associated with infant cognitive ability. *Molecular Psychiatry* 2018+ (Epub available ahead of print).

Sun, R., Hui, S., Bader, G., Lin, X., and Kraft, P. Powerful gene set analysis in GWAS with the Generalized Berk-Jones statistic. *PLOS Genetics* 2019; 15(3): e1007530.

Gaynor S.*, **Sun, R.*,** Lin, X., and Quackenbush, J. Identification of differentially expressed gene sets using the Generalized Berk-Jones statistic. *Bioinformatics* 2019+ (Epub available ahead of print).

Sun, R. and Lin, X. Set-based tests for genetic association using the Generalized Berk-Jones statistic. *Journal of the American Statistical Association* 2019+ (accepted).

PUBLISHED CORRESPONDENCE

Sun, R., Horiguchi, M., and Wei, L.J. Interpreting the benefit of trifluridine/tipiracil in metastatic colorectal cancer with respect to progression-free survival and overall survival. *Journal of Clinical Oncology* 2018; 36(13): 1378.

Sun, R., Rich, M.W., and Wei, L.J. Pembrolizumab plus chemotherapy in lung cancer. *New England Journal of Medicine* 2018; 379(11): e18.

Sun, R., Nie, L., Huang, B., Kim, D.H., and Wei, L.J. Quantifying immunoscore performance. *The Lancet* 2018; 392(10158): 1624.

Sun, R. and Wei, L.J. Regional Hyperthermia With Neoadjuvant Chemotherapy for Treatment of Soft Tissue Sarcoma. *JAMA Oncology* 2019; 5(1): 112-113.

Sun, R., Zhu, H., and Wei, L.J. Assessing the prognostic value of automated bone scan index in prostate cancer. *JAMA Oncology* 2019; 5(2): 270.

Sun, R., Orkaby, A.O., Kim, D.H., Zhu, H., and Wei, L.J. Interpreting non-inferiority of biodegradable-polymer stents to durable-polymer stents. *The Lancet* 2019; 393(10184): 1932-1933.

Sun, R., Lee, H., and Wei, L.J. Interpreting the long-term prognostic value of total mesorectal excision plane quality in rectal adenocarcinoma. *JAMA Surgery* 2019+.

Working Papers

Sun, R.*, Xu, M.*, Li, X., Gaynor, S., Zhou, H., Bossé, Y., Lam, S., Tsao, M., Tardon, A., Chen, C., Doherty, J., Goodman, G., Egil Bojesen, S., Teresa, M.T., Johansson, M., Field, J.K., Bickeböller, H, Wichmann, H., Risch, A., Rennert, G., Arnold, S., Wu, X., Melander, O., Brunnström, H., Marchand, L.L., Zong, X., Liu, G., Andrew, A., Duell, E., Kiemeney, L.A., Shen, H., Haugen, A., Johansson, M., Grankvist, K., Caporaso, N., Woll, P., Teare, M.D., Scelo, G., Hong, Y., Yuan, J., Lazarus, P., Schabath, M.B., Aldrich, M.C., Albanes, D., Brennan, P., Barbie, D., Mak, R., Hung, R.J., Amos, C.I., Christiani, D.C., Lin, X. Identification of inflammation and immune-related risk variants associated with squamous cell lung cancer.

Sun, R.*, Claggett, B.L.*, Tian, L., Solomon, S.D., Szczech, L., Pfeffer, M.A., and Wei, L.J. On-treatment analysis for assessing the between-group difference in a comparative randomized clinical study.

Park, H-R, Panganibana, R.A., **Sun, R.**, Shumyatcher, M., Himes, B.E., Christiani, D.C., and Lu, Q. Protective effects of microRNA-124 on arsenic-induced endoplasmic reticulum stress and cytotoxicity in human neural cells and associations with neurodevelopmental outcomes in children.

*Indicates equal contribution as first authors.

All working papers are currently submitted or have been submitted, and preprints are available upon request.

Invited Talks

Identification of inflammation and immune-related risk variants associated with squamous cell lung cancer. International Lung Cancer Consortium Annual Meeting. Toronto, Ontario, Canada. September 2018.

Set-based tests using the Generalized Berk-Jones statistic in genetic association studies. Jilin University School of Mathematics Statistics Colloquium. Changchun, Jilin, China. May 2018.

A unified framework for composite null inference in mediation, pleiotropy, and replicability analyses related to genetic association studies. Harvard School of Public Health Program in Genetic Epidemiology and Statistical Genetics. Boston, Massachusetts. February 2018.

A unified framework for composite null inference in mediation, pleiotropy, and replicability analyses related to genetic association studies. Harvard T.H. Chan School of Public Health P01 Environmental Statistics Retreat. Boston, Massachusetts. October 2017.

The role of inflammation pathways in lung cancer and coronary artery disease. Broad Institute Statistical Genetics Meeting. Boston, Massachusetts. October 2017.

Methods for high-dimensional inference in genetic association studies. University of New Hampshire Statistics Graduate Seminar. Durham, New Hampshire. May 2017.

Set-based tests using the Generalized Berk-Jones statistic in genetic association studies. Channing Division of Network Medicine Statistical Genetics and Networks Science Meeting. Boston, Massachusetts. January 2017.

Set-based tests using the Generalized Berk-Jones statistic in genetic association studies. Broad Institute Statistical Genetics Meeting. Boston, Massachusetts. January 2017.

Pathway analysis and gene-based inference in genomic data with the Generalized Berk-Jones statistic. Harvard T.H. Chan School of Public Health P01 Environmental Statistics Retreat. Wellesley, Massachusetts. October 2016.

Testing for gene-environment interaction under misspecification of the environment. Harvard School of Public Health Program in Genetic Epidemiology and Statistical Genetics. Boston, Massachusetts. April 2016.

Contributed Talks and Posters

A unified framework for inference in mediation, pleiotropy, and replicability analyses related to genetic association studies. Joint Statistical Meetings. Vancouver, British Columbia, Canada. August 2018.

Set-based tests using the Generalized Berk-Jones statistic in genetic association studies. Joint Statistical Meetings. Baltimore, Maryland. August 2017.

Set-based tests using the Generalized Berk-Jones statistic in genetic association studies. Eastern North Atlantic Region Spring Meeting. Washington, District of Columbia. March 2017.

Pathway analysis and gene-based inference in genomic data with the Generalized Berk-Jones statistic. Harvard T.H. Chan School of Public Health Program in Quantitative Genomics Conference. Boston, Massachusetts. November 2016.

The Generalized Berk-Jones statistic for SNP-set tests in genetic association studies. Joint Statistical Meetings. Chicago, Illinois. August 2016.

The Generalized Berk-Jones statistic for SNP-set tests in genetic association studies. Eastern North Atlantic Region Spring Meeting. Austin, Texas. March 2016.

Testing for gene-environment interaction under misspecification of the environment. Superfund Research Program Annual Meeting. San Juan, Puerto Rico. November 2015.

Testing for gene-environment interaction under misspecification of the environment. International Chinese Statistical Association - Canada Chapter Statistics and Data Science Symposium. Calgary, Alberta. August 2015.

Testing for gene-environment interaction under misspecification of the environment. Joint Statistical Meetings. Seattle, Washington. August 2015.

Professional Service

Section Chair, Methods for Single Cell Genomic Analysis, Joint Statistical Meetings 2017

Reviewer: Journal of the American Statistical Association, PLOS Genetics, npj Schizophrenia, Frontiers in Genetics, Scientific Reports, Bioinformatics

Departmental Service

Co-organizer, Program in Quantitative Genomics Seminar, 2018-2019

Owner and Administrator, Harvard Biostatistics Department Slack Messaging Space, 2017-2019

Mentor, Harvard StatStart and Harvard Summer Program in Biostatistics and Computational Biology, 2015-2018

Software Packages

GBJ: An R package for calculating the Generalized Berk-Jones statistic and its p-value. Also provides test statistic and corrected p-value for Higher Criticism, Generalized Higher Criticism, and Berk-Jones statistics when factors in a set are correlated.

GEint: An R package to calculate the exact bias of interaction coefficients in misspecified gene-environment interaction models. Also implements the Bootstrap Inference with Corrected Sandwich (BICS) procedure for testing of gene-environment interaction terms in generalized linear models.

reconstructKM: An R package to reconstruct individual-level patient data from Kaplan-Meier curves published in academic journals.

GOFexactPvalue: A C++ binary to calculate the exact p-value of Goodness-of-Fit statistics (Higher Criticism, Generalized Berk-Jones, etc.) when there are only a small number (less than 10) of correlated factors in a set.

PATENTS

SPLASSH: Open source software for camera-based high-speed, multispectral in-vivo optical image acquisition.

ADVISEES

Mengting Li, Biostatistics Master's Thesis Committee (2018)

REFERENCES

Xihong Lin, Professor
Department of Biostatistics and Department of Statistics
Harvard T.H. Chan School of Public Health
email: xlin@hsph.harvard.edu

Lee-Jen Wei, Professor
Department of Biostatistics
Harvard T.H. Chan School of Public Health
email: wei@hsph.harvard.edu

Peter Kraft, Professor
Department of Epidemiology and Department of Biostatistics
Harvard T.H. Chan School of Public Health
email: pkraft@hsph.harvard.edu

David Christiani, Professor
Department of Environmental Health and Department of Epidemiology
Harvard T.H. Chan School of Public Health
email: dchris@hsph.harvard.edu

Teaching references available upon request.

Prepared: August 20, 2019