CURRICULUM VITAE
Stephanie C. Hicks

Part I

PERSONAL DATA

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health
615 North Wolfe Street
Baltimore, MD 21205-2179 USA
(410) 614-7838

email: shicks19@jhu.edu
twitter: @stephaniehicks
github: stephaniehicks
website: stephaniehicks.com

Research Interests

I develop statistical methods, tools and software for the analysis of genomics data. Specifically, my research addresses statistical challenges in epigenomics, functional genomics and single-cell genomics such as the pre-processing, normalization, analysis of raw, noisy high-throughput data (microarray and next-generation sequencing) leading to an improved quantification and understanding of biological variability.

EDUCATION AND TRAINING

Degrees

○ PhD., Statistics, Rice University, Houston, TX (2013).
Ph.D. Advisors: Marek Kimmel, Ph.D. (Statistics - Rice) and Sharon Plon, M.D., Ph.D. (Baylor College of Medicine, TX Children's)
Dissertation: Probabilistic Models for Genetic and Genomic Data with Missing Information
○ M.A., Statistics, Rice University, Houston, TX (2012).
○ B.S., Mathematics, Louisiana State University, Baton Rouge, LA, (2007).
magna cum laude, Phi Beta Kappa

Postdoctoral Training

○ Postdoctoral Research Fellow (2013 – 2018).
Department of Data Sciences, Dana-Farber Cancer Institute
Department of Biostatistics, Harvard T.H. Chan School of Public Health
Mentor: Rafael Irizarry, Ph.D.
K99/R00 Co-mentors: Franziska Michor, Ph.D. and Bradley Bernstein, M.D, Ph.D

PROFESSIONAL EXPERIENCE

○ Assistant Professor, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health (2018 – present).
○ Member, Johns Hopkins Data Science Lab (2018 – present).
○ Postdoctoral Research Fellow, Department of Data Sciences, Dana-Farber Cancer Institute, Boston, MA (2013 – 2018).
○ Postdoctoral Research Fellow, Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA (2013 – 2018).
○ Research Assistant, Department of Statistics, Rice University (2007 – 2013).
Ph.D. Advisors: Marek Kimmel and Sharon Plon
○ Teaching Assistant, Department of Statistics, Rice University (2007 – 2013).

PROFESSIONAL ACTIVITIES

Professional Memberships

○ Caucus for Women in Statistics (CWS) (2017 – present).
○ American Statistical Association (ASA) (2007 – present).
○ American Mathematical Society (AMS) (2007 – present).
○ American Society of Human Genetics (ASHG) (2007 – 2013).

Participation on Advisory Panels and Committees

○ Chair of ASA Committee on Women in Statistics (2020).
○ Vice Chair of ASA Committee on Women in Statistics (2019).
○ Co-founder of R-Ladies Baltimore: An organization to promote gender diversity in the R community (2018 – present).
○ Member, Emerging Technologies Subcommittee of the Bioconductor Technical Advisory Board (2018 – present).
○ Member of ASA Committee on Women in Statistics (2017 – 2022).
○ Board Member of Cards Against Humanities Science Ambassador Scholarship (2015 – present).

Program Development

○ Program Committee Member of the 2020 Symposium in Data Science and Statistics (2019 – 2020).
○ Program Committee Member of the 2019 Symposium in Data Science and Statistics (2018 – 2019).
○ Session Organizer, Joint Statistical Meetings. (2019). Topic: *Changing the statistics community: effective strategies for promoting an inclusive and equitable culture for women*.
○ Session Organizer, Eastern North Atlantic Region Meeting (2019). Topic: *Teaching Data Science using Case-Studies*.
○ Session Organizer, Joint Statistical Meetings. (2018). Topic: *Selected for a highly competitive Late-Breaking Session on Addressing Sexual Misconduct in the Statistics Community*).
○ Session Chair, Joint Statistical Meetings. (2019). Topic: *Changing the statistics community: effective strategies for promoting an inclusive and equitable culture for women*.
○ Session Chair, Eastern North Atlantic Region Meeting (2019). Topic: *Teaching Data Science using Case-Studies*.
○ Session Chair, Joint Statistical Meetings (2016).

EDITORIAL ACTIVITIES

Editorial Board Membership

○ Associate Editor, *biOverlay* (2018 – present).

Peer Review Activities

○ **Journal**: Annals of Statistics, Bioinformatics, Biometrics, Biostatistics, BMC Medical Genetics, F1000, Genome Biology, Human Mutation, Nature Communications, Nature Methods, Nature Biotechnology, Nature Structural & Molecular Biology, Nature Scientific Reports, PLoS Computational Biology, PLoS One, R Journal
○ **Book**: Chapman & Hall/CRC

Review of Proposals

○ Review statistical methods proposals for NSF panel (2019)

HONORS AND AWARDS

- Keynote lecture, Bioconductor Asia (2018).
- Keynote lecture, Australian Mathematical Sciences Institute, BioInfoSummer (2018).
- Research Grant, NIH/NHGRI K99/R00 Pathway to Independence Award (2017 – 2021).
- Travel award, Caucus for Women in Statistics (CWS) (2017).
- Travel award, Ascona Workshop on Statistical Challenges in Single-Cell Biology in Ascona, Switzerland (2017).
- Travel award, Women in Statistics and Data Science Conference (2016).
- Stellar Abstract Award, PQG Conference (2015).
- Travel award, Genome Informatics Meeting, Cold Spring Harbor Laboratories (2015).
- Travel award, Women in Statistics Conference (2014).
- Travel and tuition award, 16th Annual Summer Institute in Statistical Genetics at University of Washington (2011).
- Pre-doctoral Fellowship, NIH/NCI T32 Predoctoral Training Program in Biostatistics, Rice University (2009 – 2011).
- Phi Beta Kappa Society (2007).
- LSU Austin Chapter Scholarship Award (2007).
- Undergraduate scholarship, LA-STEM Research Scholars, LSU (2005 – 2007).
- Undergraduate scholarship, HMMI Professors Program, LSU (2004 – 2005).
- TOPS Tuition Award, LSU (2003 – 2007).

PUBLICATIONS *The white numbers indicate first or senior author manuscripts.*

Published Peer-Reviewed Articles

[1] Korthauer K, Kimes PK, Duvallet C, Reyes A, Subramanian A, Teng M, Shukla C, Alm EJ, **Hicks SC**. (2019). A Practical Guide to Methods Controlling False Discoveries in Computational Biology. ▶ *Genome Biol*. **20**:118. PMID: 31164141 PMCID: PMC6547503.

[2] Kumar MS, Slud EV, Okrah K, **Hicks SC**, Hannenhalli S, Corrada Bravo H. (2018). Analysis And Correction Of Compositional Bias In Sparse Sequencing Count Data. ▶ *BMC Genomics*. **19**:799. PMID: 30400812. PMCID: PMC6219007.

[3] **Hicks SC**, Irizarry RA. (2018). A Guide to Teaching Data Science. ▶ *The American Statistician* **72**: 382-391.

[4] **Hicks SC**, Townes FW, Teng M, Irizarry RA. (2018). Missing Data and Technical Variability in Single-Cell RNA-Sequencing Experiments. ▶ *Biostatistics* **19**: 562-578. PMID: 29121214. PMCID: PMC6215955.

[5] **Hicks SC**, Okrah K, Paulson JN, Quackenbush J, Irizarry RA, Corrada Bravo H. (2017). Smooth Quantile Normalization. ▶ *Biostatistics* **19**: 185-198. PMID: 29036413. PMCID: PMC5862355.

[6] Saliba J, Zabriskie R, Ghosh R, Powell BC, **Hicks S**, Kimmel M, Meng Q, Ritter DI, Wheeler DA, Gibbs RA, Tsai FTF, Plon SE. (2016). Pharmacogenetic characterization of naturally occurring germline NT5C1A variants to chemotherapeutic nucleoside analogs. ▶ *Pharmacogenetics and Genomics* **26**: 271-279. PMID: 26906009. PMCID: PMC4853247.

[7] **Hicks SC**[1], Irizarry RA. (2015). quantro: a data-driven approach to guide the choice of an appropriate normalization method. ▶ *Genome Biol* **16**:117. PMID: 26040460. PMCID: PMC4495646.

[8] Osman AA, Neskey DM, Katsonis P, Patel AA, Ward AM, Hsu TK, **Hicks SC**, McDonald TO, Ow TJ, Alves MO, Pickering CR, Skinner HD, Zhao M, Sturgis EM, Kies MS, El-Naggar A, Perrone F, Licitra L, Bossi P, Kimmel M, Frederick MJ, Lichtarge O, Myers JN. (2015). Evolutionary Action Score of TP53 Coding Variants Is Predictive of Platinum Response in Head and Neck Cancer Patients.▶ *Cancer Res* **75**: 1-11. PMID: 25691460.

[9] Neskey DM, Osman AA, Ow TJ, Katsonis P, McDonald T, **Hicks SC**, Hsu TK, Pickering CR, Ward A, Patel A, Yordy JS, Skinner HD, Giri U, Sano D, Story MD, Beadle BM, El-Naggar AK, Kies MS, William WN, Caulin C, Frederick M, Kimmel M, Myers JN, Lichtarge O. (2015). Evolutionary Action score of TP53 (EAp53) identifies high risk mutations associated with decreased survival and increased distant metastases in head and neck cancer. ▶ *Cancer Res* **75**: 1527-1536. PMID: 25634208.

---

[1] 'Highly Accessed' on BioMed Central

10 Berger RL, Li LT, **Hicks SC**, Liang MK. (2014). Suture versus preperitoneal polypropylene mesh for elective umbilical hernia repairs. ▶ *J Surg Res* **192**: 426-431. PMID: 24980854.

11 Brahmbhatt R, Carter SA, **Hicks SC**, Berger DH, Liang MK. (2014). Identifying Risk Factors for Surgical Site Complications after Laparoscopic Ventral Hernia Repair: Evaluation of the Ventral Hernia Working Group Grading System. ▶ *Surg Infect (Larchmt)* **15**: 187-193. PMID: 24773169

12 Carter SA, **Hicks SC**, Brahmbhatt R, Liang MK. (2014). Recurrence and Pseudorecurrence after Laparoscopic Ventral Hernia Repair: Predictors and Patient-focused Outcomes. ▶ *Am Surg* **80**: 138-148. PMID: 24480213

13 Li LT, **Hicks SC**, Davila JA, Kao LS, Berger RL, Arita NA, Liang MK. (2014). Circular Closure is Associated with the Lowest Rate of Surgical Site Infection Following Stoma Reversal:A Systematic Review and Multiple Treatment Meta-analysis. ▶ *Colorectal Dis* **16**: 406-416. PMID: 24422861.

14 Li LT, Brahmbhatt R, **Hicks SC**, Davila JA, Berger DH, Liang MK. (2014). Prevalence of Surgical Site Infection at the Stoma Site following Four Skin Closure Techniques: A Retrospective Cohort Study. ▶ *Dig Surg* **31**: 73-78. PMID: 24776653.

15 Li LT, Jafrani RJ, Becker NS, Berger RL, **Hicks SC**, Davila JA, Liang MK. (2014). Outcomes of acute versus elective primary ventral hernia repair. ▶ *J Trauma Acute Care Surg* **76**: 523-528. PMID: 24458061.

16 Liang MK, Li LT, Nguyen MT, Berger RL, **Hicks SC**, Kao LS. (2014). Abdominal reoperation and mesh explantation following open ventral hernia repair with mesh. ▶ *Am J Surg* **204**: 670-676. PMID: 25241955.

17 Liang MK, Berger RL, Nguyen MT, **Hicks SC**, Li LT, Leong M. (2014). Outcomes with Porcine Acellular Dermal Matrix versus Synthetic Mesh and Suture in Complicated Open Ventral Hernia Repair. ▶ *Surg Infect* **15**: 506-512. PMID: 25215466.

18 Nguyen MT, Berger RL, **Hicks SC**, Davila JA, Li LT, Kao LS, Liang MK. (2014). Comparison of Outcomes of Synthetic Mesh vs Suture Repair of Elective Primary Ventral Herniorrhaphy: A Systematic Review and Meta-analysis. ▶ *JAMA Surg* **31**: 73-78. PMID: 24554114.

19 Nguyen MT, Phatak UR, Li LT, **Hicks SC**, Moffett JM, Arita NA, Berger RL, Kao LS, Liang MK. (2014). Review of stoma site and midline incisional hernias after stoma reversal. ▶ *J Surg Res* **190**: 504-509. PMID: 24560428.

20 Berger RL, Li LT, **Hicks SC**, Davila JA, Kao LS, Liang MK. (2013). Development and validation of a risk-stratification score for surgical site occurrence and surgical site infection after open ventral hernia repair. ▶ *J Am Coll Surg.* **217**: 974-982. PMID: 24051068.

21 Clapp ML, **Hicks SC**, Awad SS, Liang MK. (2013). Trans-cutaneous Closure of Central Defects (TCCD) in Laparoscopic Ventral Hernia Repairs (LVHR). ▶ *World J Surg.* **37**: 42-51. PMID: 23052806.

22 Liang MK, Clapp M, Li LT, Berger RL, **Hicks SC**, Awad S. (2013). Patient Satisfaction, Chronic Pain, and Functional Status following Laparoscopic Ventral Hernia Repair. ▶ *World J Surg.* **37**: 530-537. PMID: 23212794.

23 Liang MK, Li LT, Avellaneda A, Moffett JM, **Hicks SC**, Awad SS. (2013). Outcomes and Predictors of Incisional Surgical Site Infection in Stoma Reversal. ▶ *JAMA Surg* **148**: 183-189. PMID: 23426597.

24 Liang MK, Berger RL, Li LT, Davila JA, **Hicks SC**, Kao LS. (2013). Outcomes of laparoscopic vs open repair of primary ventral hernias. ▶ *JAMA Surg* **148**: 1043-1048. PMID: 24005537.

25 Neskey DM, Klein JD, **Hicks S**, Garden AS, Bell DM, El-Naggar Ak, Kies MS, Weber RS, Kupferman ME. (2013). Prognostic Factors Associated With Decreased Survival in Patients With Acinic Cell Carcinoma. ▶ *JAMA Otolaryngol Head Neck Surg* **139**: 1195-1202. PMID: 24076756.

26 Nguyen MT, Li LT, **Hicks SC**, Davila JA, Suliburk JW, Leong M, Kao LS, Berger DH, Liang MK. (2013). Readmission following open ventral hernia repair: incidence, indications, and predictors.▶ *Am J Surg* **206**: 942-948. PMID: 24296099.

27 Subramanian A, Clapp ML, **Hicks SC**, Awad SS, Liang MK. (2013). Laparoscopic ventral hernia repair: Primary versus secondary hernias. ▶ *J Surg Res* **181**: e1-5. PMID: 22795342.

28 Cheung HC[2], San Lucas FA, **Hicks S**, Chang K, Bertuch AA, Ribes-Zamora A. (2012). An S/T-Q cluster domain census unveils new putative targets under Tel1/Mec1 control. ▶ *BMC Genomics* **13**: 664. PMID: 23176708.

---

[2] 'Highly Accessed' on BioMed Central

[29] **Hicks S**[3], Wheeler DE, Plon SE, Kimmel M. (2011). Prediction of Missense Mutation Functionality Depends on both the Algorithm and Sequence Alignment Employed. ▶ *Human Mutation*. **32**: 661-668. PMID: 21480434.

Open Case Studies Data Analyses

[30] Stephens A, Jager L, Taub M, **Hicks S**. (2019, February 14). opencasestudies/ocs-police-shootings-firearm-legislation: Firearm Legislation and Fatal Police Shootings in the United States (Version v1.0.0). ▶ *Zenodo*. doi: `http://doi.org/10.5281/zenodo.2565249`.

[31] Kuo P-L, Jager L, Taub M, **Hicks S**. (2019, February 14). opencasestudies/ocs-healthexpenditure: Exploring Health Expenditure using State-level data in the United States (Version v1.0.0). ▶ *Zenodo*. doi: `http://doi.org/10.5281/zenodo.2565307`.

Refereed Letters, Communications, Book Chapters, Proceedings, Technical Reports, Other

[32] **Hicks SC**. (2014). When Women in Statistics Come to Know Their Power. ▶ *Chance* **27**.

[33] **Hicks S**, Plon SE, Kimmel M. (2013). Statistical Analysis of Missense Mutation Classifiers. ▶ *Hum Mut* **34**: 405-406. PMID: 23086893.

[34] **Hicks SC**. Probabilistic Models for Genetic and Genomic Data with Missing Information, Ph.D. Thesis, Rice University, 2013.

Under Review / Pre-prints

[35] Lahnemann D, Koester J, Szczurek E, McCarthy D, **Hicks S**, Robinson MD, Vallejos C, Beerenwinkel N, Campbell K, Mahfouz A, Pinello L, Skums P, Stamatakis A, Stephan-Otto Attolini C, Aparicio S, Baaijens J, Balvert M, de Barbanson B, Cappuccio A, Corleone G, Dutilh B, Florescu M, Guryev V, Holmer R, Jahn K, Lobo TJ, Keizer E, Khatri I, KieÅbasa S, Korbel J, Kozlov A, Kuo T-H, Lelieveldt B, Mandoiu I, Marioni J, Marschall T, Molder F, Niknejad A, Rczkowski L, Reinders M, de Ridder J, Saliba A-E, Somarakis A, Stegle O, Theis F, Yang H, Zelikovsky A, McHardy A, Raphael B, Shah S, Schonhuth S. (2019). 12 Grand Challenges in Single-Cell Data Science. ▶ *PeerJ Preprints* doi: `https://doi.org/10.7287/peerj.preprints.27885v1`.

[36] **Hicks SC**, Peng RD. (2019). Elements and Principles for Characterizing Variation between Data Analyses. ▶ *arXiv*. doi: `https://arxiv.org/abs/1903.07639`.

[37] **Hicks SC**, Peng RD. (2019). Evaluating the Success of a Data Analysis. ▶ *arXiv*. doi: `https://arxiv.org/abs/1904.11907`.

[38] Amezquita RA, Carey VJ, Carpp LN, Geistlinger L, Lun ATL, Marini F, Rue-Albrecht K, Risso D, Soneson C, Waldron L, Pages H, Smith M, Huber W, Morgan M, Gottardo R, **Hicks SC**. (2019). Orchestrating Single-Cell Analysis with Bioconductor. ▶ *bioRxiv*. doi: `https://doi.org/10.1101/590562`.

[39] Townes FW, **Hicks SC**, Ayree MJ Irizarry RA. (2019). Feature Selection and Dimension Reduction for Single Cell RNA-Seq based on a Multinomial Model. ▶ *bioRxiv*. doi: `https://doi.org/10.1101/574574`.

[40] **Hicks SC**, Irizarry RA. (2017). Technology-independent estimation of cell type composition using differentially methylated regions. ▶ *bioRxiv*. doi: `https://doi.org/10.1101/213769`.

PRACTICE ACTIVITIES

Software and data packages

Bioconductor software packages

[1] **mbkmeans** ▶ Methodology and R/Bioconductor software package implementing the mini-batch optimization for *k*-means clustering proposed in Sculley (2010) for large datasets including single-cell RNA-sequencing data [Available on Bioconductor]. The mini batch *k*-means algorithm can be run with data stored in memory or on disk (e.g. HDF5 file format).

---

[3]Cited and Discussed in Nature **482**: 257-262. 09 Feb 2012. PMID: 22318607

2 **qsmooth** ▶ Methodology and R/Bioconductor software package that implements a generalization of quantile normalization, referred to as smooth quantile normalization (**qsmooth**), which is based on the assumption that the statistical distribution of each sample should be the same (or have the same distributional shape) within biological groups or conditions [Available on Bioconductor].

3 **quantro** ▶ Methodology and R software package that can be used to test for differences between groups of distributions to guide the choice if quantile normalization should be used [Available on Bioconductor]

Bioconductor data packages

1 **benchmarkfdrData2019** ▶ R/Bioconductor data package that contains benchmarking results for experimental and simulated data sets used in Korthauer and Kimes et al. (2019) to compare methods for controlling the false discovery rate. A shiny app to interactively explore the data is available on the kdkorthauer/benchmark-shiny repository on GitHub [Available on Bioconductor]

2 **bodymapRat** ▶ R/Bioconductor data package that contains an `SummarizedExperiment` from the Yu et al. (2013) paper that performed the rat BodyMap across 11 organs and 4 developmental stage [Available on Bioconductor]

3 **TENxPBMCData** ▶ R/Bioconductor data package that contains a set of `SingleCellExperiment` objects with single-cell RNA-sequencing data from peripheral blood mononuclear cells generated by 10X Genomics. [Available on Bioconductor]

GitHub software packages

1 **methylCC** ▶ Methodology and R software package to estimate the cell composition of whole blood in DNA methylation samples in microarray or sequencing platforms [Available on GitHub]. The methodology and software was developed in collaboration with Rafael Irizarry at Dana-Farber Cancer Institute.

2 **quantroSim** ▶ A supporting data simulation R-package for the *quantro* R-package to simulate gene expression and DNA methylation data [Available on GitHub]

3 **explainr** ▶ Software R package to translate S3 objects into text using standard templates in a simple and convenient way. Developed at ROpenSci Unconference with Hilary Parker, David Robinson and Roger Peng. [Available on GitHub]

4 **postMUT** ▶ A tool implemented in Perl and R to predict the functionality of missense mutations [Available on GitHub]

GitHub data packages

1 **trapnell2014myoblasthuman** ▶ R data package that contains an ExpressionSet object from Trapnell et al. (2014) that performed a time-series experiment of bulk and single cell RNA-Seq at four time points in differentiated primary human myoblasts. [Available on GitHub]

2 **patel2014gliohuman** ▶ R data package that contains a SummarizedExpression object from Patel et al. (2014) with single cell and bulk RNA-Seq data on five human glioblastoma tumors. [Available on GitHub]

3 **colonCancerWGBS** ▶ Cov files produced from [Bismark](http://www.bioinformatics.babraham.ac.uk/projects/bismark/) after mapping six paired tumor-normal WGBS samples from Ziller et al. (2013) PMID: 23925113. Only chr22. [Available on GitHub]

4 **myAffyData** ▶ AffyBatch object from an experiment using P493-6 cells expressing low or high levels of c-Myc. Data from Loven et al. (2012) Cell 151: 476-482. [Available on GitHub]

5 **BackgroundExperimentYeast** ▶ AffyBatch object from an experiment to measure NSB and optical noise in yeast. [Available on GitHub]

TEACHING

Academic Advisees

1 Dyjack, Nathan. Doctor of Philosophy, Biostatistics (2019 – Present)
2 Zhang, Jingning. Doctor of Philosophy, Biostatistics (2018 – 2019).
3 Sawyer, Holly, Masters of Public Health, Biostatistics and Epidemology (2018 – present).

T32 Fellowship Committee

1 Hui, Ken. Johns Hopkins School of Medicine (2019 – Present).

Thesis Committees / Thesis Reader

1 Nabi, Razieh. Doctor of Philosophy, Computer Science (2019 – Present).
2 Fang, Weixiang. Doctor of Philosophy, Biostatistics (2019 – Present).
3 Ji, Zhicheng. Doctor of Philosophy, Biostatistics (2019 – Present).

Classroom Instruction - Co-Instructor

Johns Hopkins Bloomberg School of Public Health

○ 140.712 Advanced Data Science II (2019).
○ 140.711 Advanced Data Science I (2019).
○ 140.712 Advanced Data Science II (2018).
○ 140.711 Advanced Data Science I (2018).

Harvard T.H. Chan School of Public Health

○ BIO 260 Introduction to Data Science (2016).

Classroom Instruction - Teaching Fellow/Assistant

Cold Spring Harbor Laboratories

○ Statistical Methods for Functional Genomics (2015).

Harvard edX

○ Statistics and R for the Life Sciences (Harvard edX - PH525.1x) (2015).
○ Case study: DNA methylation data analysis (Harvard edX - PH525.8x) (2015).
○ Data Analysis for Genomics (Harvard edX - PH525x) (2014).

Harvard University

○ CS 109 Introduction to Data Science, *Lead Teaching Fellow of 25 Teaching Fellows* (2014).

Rice University

- STAT 423/623 Probability in Bioinformatics and Genetics (2010, 2011).
- STAT 552 Applied Stochastic Processes (2010).
- STAT 305 Introduction to Statistics for the Biosciences (2008, 2009).
- STAT 310 Probability and Statistics (2007, 2008).

Classroom Instruction - Workshops

- Instructor, workshop on Statistical Analysis and Comprehension of Single Cell RNA-Sequencing Data in R / Bioconductor. Australian Mathematical Sciences Institute (AMSI) Bioinfosummer. 2018 Dec 3-7, University of Western Australia, Perth, Australia.
- Instructor, Introduction to R, RStudio and the "Tidyverse". How It's Done Seminar in the Biostatistics and Computational Biology Department, DFCI. 2017 Sept 20. Boston, MA, USA.
- Instructor, Introduction to R and R Package Demonstration. Department of Biostatistics and Computational Biology Training Sessions. DFCI. 2016 March 9. Boston, MA, USA.

Educational resources

- **opencasestudies**: An educational resource containing a repository of case studies to teach how to analyze data (Website: https://opencasestudies.github.io)

RESEARCH GRANT PARTICIPATION *The solid dots • indicate grants to SCH.*

Ongoing Research Support

- *Enabling Analysis and Comprehension of Human Cell Atlas through R / Bioconductor (Chan Zuckerberg Initiative).*
  Dates: 07/01/2019 – 06/30/2022. Principal Investigator: Stephanie Hicks, PhD.
  Responsibility: Co-PI (MPI).

- *Hands-On Data Science Education for Thousands of Students in Tackling Public Health Challenges (Bloomberg Philanthropies).*
  Dates: 09/01/2019 - 08/31/2020. Principal Investigator: Stephanie C. Hicks, PhD.
  Responsibility: Principal Investigator.

- *Characterization of high-grade serous ovarian cancer subtypes via single-cell profiling (NIH/NCI R01).*
  Dates: 05/01/2019 - 04/30/2023. Principal Investigator: Casey Greene, PhD. Jen Doherty, PhD.
  Responsibility: Co-Investigator.

- *Practical, reproducible, and statistically-rigorous workflows for single-nucleus analysis of childhood cancer data (Alexs Lemonade Stand Foundation).*
  Dates: 03/01/2019 - 02/29/2020. Principal Investigator: Stephanie C. Hicks, PhD.
  Responsibility: Principal Investigator.

- *Statistical Analysis and Comprehension of the Human Cell Atlas in R / Bioconductor: Access and Scalable Infrastructure (Chan Zuckerberg Initiative, Computational Tools for the Human Cell Atlas).*
  Dates: 03/01/2018 – 08/31/2019. Principal Investigator: Davide Risso, PhD.
  Responsibility: Co-Investigator.

- *Statistical Methods for the Normalization and Quantification of Single-Cell RNA-Sequencing Data (NIH/NHGRI K99/R00).*
  Dates: 12/23/2016 – 02/28/2021. Principal Investigator: Stephanie C. Hicks, PhD.
  Responsibility: Principal Investigator.

Completed

- *Single-Cell Dynamics for Precision Medicine in Cancer (The Jayne Koskinas Ted Giovanis Foundation).*
  Dates: 08/10/2018 – 08/09/2019. Principal Investigators: Sahand Hormoz, PhD. Yuchao Jiang, PhD. Stephanie C. Hicks, PhD.
  Responsibility: Co-PI (MPI).

- *Predoctoral Training Program in Biostatistics, Rice University (NIH T32 CA096520).*
  Dates: 08/01/2009 – 12/31/2011. Principal Investigator: Marina Vannucci, PhD.
  Responsibility: Predoctoral Trainee.

ACADEMIC SERVICE

Department of Biostatistics

- Member, Committee for Emerging Opportunities for Department Retreat (2018).

Johns Hopkins University

- Participant, Panel on how JHU can help junior faculty succeed (follow up to COACHE faculty satisfaction survey), Peabody Library (2018).

PRESENTATIONS

Scientific Meetings (Invited)

1 Data Science Leadership Summit. 2019 Nov 7-9. Santa Fe, NM. USA.

2 *Addressing Public Health Challenges through Data Science Education.*
   American Public Health Association Annual Meeting. 2019 Nov 2-6. Philadelphia, PA. USA.

3 *Addressing Open Challenges in Data Science Education.*
   Mathematics Departments and the Explosive Growth of Computational and Quantitative Offerings in Higher Education. 2019 Oct 25. Washington, D.C. USA

4 *Teaching data science through case studies in Public Health.*
   Joint Statistical Meeings. 2019 Jul 27-Aug 1. Denver, CO. USA.

5 *Useful Tools for Teaching And Outreach In Data Science.*
   Symposium on Data Science and Statistics (SDSS). 2019 May 29-Jun 1. Bellevue, WA. USA.

6 *Missing Data and Technical Variability in Single-Cell RNA-Sequencing Experiments.*
   ENAR Conference. 2019 Mar 24-27. Philadelphia, PA. USA.

7 *Orchestrating Single-Cell RNA-sequencing Analysis with Bioconductor.*
   Keynote presentation at Bioconductor Asia on 2018 Nov 29 at the University of Melbourne, Melbourne, Australia.
   Invited presentation at Australian Mathematical Sciences Institute (AMSI) Bioinfosummer on 2018 Dec 3-7 at the University of Western Australia, Perth, Australia.

8 *Analyzing Genomics Data in R with Bioconductor.*
   DC R Conference. 2018 Nov 8. Washington D.C., USA.

9 *Estimating cell type composition in whole blood using differentially methylated regions.*
   12th Annual Symposium and Poster Session on Genomics and Bioinformatics at Johns Hopkins University. 2018 Oct 25. Baltimore, MD. USA.

10 *Open Challenges in Single-Cell RNA-Sequencing.*
    Single Cell Data Science: Making Sense of Data from Billions of Single Cells. 2018 Jun 4-8. Lorentz Center, Leiden University, Netherlands.

11 *Estimating cell type composition in whole blood using differentially methylated regions.*
   Bioconductor Conference. 2017 Jul 26-28. Boston, MA, USA.

12 *Batch effects and technical biases in scRNA-Seq data.*
   HSCI Single-Cell Workshop. 2016 Nov 29-30. Harvard Medical School, Boston, MA, USA.

13 *Towards progress in batch effects and biases in single-cell RNA-seq data.*
   Single-Cell Genomics Conference. 2016 Sept 14-16. Wellcome Genome Campus, Hinxton, Cambridge, UK.

14 *On the widespread and critical impact of systemic bias and batch effects in single-cell RNA-seq data.*
   Joint Statistical Meetings. 2016 Jul 31-Aug 4. Chicago, IL, USA.

15 *On the widespread and critical impact of systemic bias and batch effects in single-cell RNA-seq data.*
   Boston Single-Cell Network Meeting. 2016 Mar 15. Boston, MA, USA.

16 *Normalization of DNA methylation and Gene Expression Data in the Context of Global Variation.*
   Bioinformatics Meeting, Division of Immunology, Harvard Medical School. 2014 Sept 18. Boston, MA, USA.

Scientific Meetings (Contributed)

17 `mbkmeans`: *fast clustering for single cell data using mini-batch k-means.*
   Bioconductor Conference. 2019 Jul 24-27. New York City, NY, USA.

18 *Open challenges in Single-Cell RNA-Sequencing.*
   Joint Statistical Meeings. 2018 Jul 29-Aug 2. Vancouver, CA.

19 *Missing Data and Technical Variability in Single-Cell RNA-Sequencing Experiments.*
   Joint Statistical Meetings. 2017 Jul 29-Aug 3. Baltimore, MD, USA.

20 *Missing Data and Technical Variability in Single Cell RNA-Sequencing Experiments.*
   Ascona Workshop 2017: Statistical Challenges in Single Cell Biology. 2017 Apr 30-May 5. Ascona, Switzerland.

21 *Transforming the Classroom to Teach Statistics and Data Science with Active Learning.*
   Women in Statistics and Data Science Conference. 2016 Oct 20-22. Charlotte, NC, USA.

22 *On the widespread and critical impact of batch effects and systematic bias in single-cell RNA-Seq data*[4].
   PQG Conference: Single-cell Genomics: Technology, Analysis, and Applications. 2015 Nov 5-6. Boston, MA, USA.

23 *quantro: When should you use quantile normalization?.*
   Flashlight talk at Bioconductor Conference. 2014 Jul 30-Aug 1. Boston, MA, USA.

24 *Modeling Discovery Of Functional SNPs From Genome Scale Data.*
   Joint Statistical Meetings. 2011 Aug 5. Miami, FL, USA.

25 *Prediction of Missense Mutation Functionality Depends on both the Algorithm and Sequence Alignment Employed.*
   Human Genome Variation Society's Exploring the Functional Consequences of Genomic Variation Meeting. 2010 Nov 1. Washington, D.C., USA.

Scientific Meetings (Conference Proceedings)

26 *Functional analysis of genomic variants identified through whole exome sequencing of pediatric lymphocytic leukemia kindreds.*
   AACR Special Conference: Pediatric Cancer at the Crossroads: Translating Discovery into Improved Outcomes. 2013 Nov 3-6. San Diego, CA.

27 *Functional analysis of genomic variants identified through whole exome sequencing for susceptibility to lymphocytic leukemia.*
   AACR 103rd Annual Meeting. 2012 Mar 31-Apr 4. Chicago, IL.

---

[4]Selected for a Stellar Abstract Award

Invited Seminars

28 Human Medical Genetics and Genomics Program, University of Colorado (Anschutz Medical Campus). 2020 Feb 27. Denver, CO. USA.

29 Department of Biostatistics and Informatics, Colorado School of Public Health. 2020 Feb 26. Denver, CO. USA.

30 Carnegie Institution for Science. 2020 Jan 6. Baltimore, MD. USA.

31 *Addressing Open Challenges in Genomics and Data Science Education.*
    Department of Statistics and Data Sciences, Smith College. 2019 Apr 18. Northampton, MA. USA.

32 *Making data science accessible world-wide in the Johns Hopkins Data Science Lab.*
    Department of Statistics, University of Connecticut. 2019 Feb 27. Storrs, CT. USA.

33 *Applications of Latent Variables in Identifying Systematic Errors in Genomics.*
    Department of Statistics, Rice Unviersity. 2019 Feb 11. Houston, TX. USA.

34 *Estimating cell type composition in whole blood using differentially methylated regions.*
    Department of Statistics, Oregon State University. 2019 Jan 14. Corvallis, OR. USA.

35 *Estimating cell type composition in whole blood using differentially methylated regions.*
    Department of Biostatistics, The University of Alabama at Birmingham. 2018 Nov 2. Birmingham, AL. USA.

36 *Technology-independent estimation of cell type composition using differentially methylated regions.*
    Office of Biostatistics Research at NIH/NHLBI. 2018 Apr 24. Bethesda, MD. USA.

37 *Setting the Stage for Reproducibility and Replicability in Science.*
    Department of Biology, Brandeis University. 2017 Mar 22. Waltham, MA, USA.

38 *Why Statistics Matters in the Analysis of Genomics Data.*
    Department of Biology, Louisiana State University. 2015 Feb 11. Baton Rouge, LA, USA.

Other Meetings and Events

39 *Practical Genomics Workshop* (website). Presented a talk on Why Statistics Matters: Analysis of Genomics Data. 2019. June 4-7. Baltimore, MD. USA.

40 *National Human Genome Research Institutes Genome to Phenotype Strategic Planning Meeting* (website). 2019 Jan 22-24. Rockville, MD, USA. Three-day workshop led by the NHGRI to discuss research opportunities and barriers in these scientific areas and to consider options for future NHGRI programs to identify paradigm-shifting areas of genomics that will expand the field into new frontiers and enable novel applications to human health and disease.

41 *2018 Women in Statistics and Data Science Conference* (website). 2018 Oct. Cincinnati, OH, USA.
    Lead a project on building a children's book to highlight women in Statistics and Data Science.

42 *2018 Data Science Innovation Lab: Mathematical Challenges of Single Cell Dynamics* (website). 2018 Jun. Bend, OR. USA.
    Invited to participate in a five-day workshop with early-career biomedical and quantitative investigators to create collaborations, answering important questions in single cell data.

43 *Mozilla Open Science Leadership Summit* (website). 2015 Sept. Mozilla, Toronto, Canada.
    Invited to work with community leaders furthering open practice and open science through creating community events, tools for collaboration and learning resources.

44 *ROpenSci Unconference* (website). 2015 Mar. GitHub, San Francisco, CA, USA.
    Invited to work with over 40 R enthusiasts from industry, academia, non-profits and government on projects supporting open data, open science and data visualization in R.
    Contributed to R-packages including explainr and catsplainr.