JESSICA R. SHAW

Telephone: 815-975-7515
Email: jessica.rose.shaw@gmail.com
LinkedIn: https://linkedin.com/in/shawjr
GitHub: https://github.com/shawjes/CV_Companion

EXPERTISE

- QC, integration, analysis, interpretation of multi-omics data
- Study design
- Process development
- Project management
- Multivariate statistics
- HPC/Cloud computing
- Data visualization
- Collaboration
- Scientific writing
- Presenting

EDUCATION

8/2014 - 12/2017
University of Colorado – Anschutz
Aurora, CO
Master of Science: Biostatistics
Minor: Statistical Genomics and Genetics

7/2017 – 7/2017
University of Washington
Seattle, WA
Certificate: Summer Institute in Statistical Genetics (SISG)

8/2017 – 8/2017
University of Washington Genetic Analysis Center
Seattle, WA
Certificate: TOPMed Analysis Workshop

8/2009 - 5/2013
Purdue University – West Lafayette
West Lafayette, IN
Bachelor of Science: Management
Concentrations: Marketing, International Business

SUMMARY

Gritty, purpose-driven data scientist with 13 years of experience spanning biostatistics, bioinformatics, and analytics consulting. Possess a M.S. in Biostatistics, Statistical Genomics and Genetics from the University of Colorado.

EXPERIENCE

Illumina, Inc. **Remote**
Senior Clinical Data Scientist 5/2022 – 3/2024

Provided statistical and bioinformatic programming for clinical validation studies of next generation sequencing (NGS) instruments and oncology companion diagnostic assays.

- Produced clinical analysis datasets, tables, listings, and figures (TLFs) using applicable CDISC/ADaM/SDTM standards.
- Supported the successful launches of the NovaSeq 6000Dx high-throughput sequencing instrument and the TruSight Oncology Comprehensive assay.
- Developed a data quality review tool for efficient quality control of raw sequence data.
- Built and maintained computational tools for standardized and reproducible analyses.

Linda Crnic Institute for Down Syndrome at University of Colorado **Aurora, CO**
Bioinformatics Scientist/Omics Data Scientist 9/2019 – 04/2022

Identified single and multi-omics signatures of trisomy 21 (T21) and associated comorbidities through integrative analysis of medical record (MR/EHR), genomic, and multi-omics data (genotype array, WGS, CyTOF single-cell immune profiling, multiplex immunofluorescence cytokine assay, transcriptomic, proteomic, metabolomic).

- Developed and executed integrative analyses of omics and medical record data.
- Demonstrated utility of a polygenic risk score (PRS) for Celiac disease in T21.
- Improved the performance of a PRS with incorporation of blood protein levels.
- Applied robust machine learning methods to predict karyotype and comorbidities.

Colorado Center for Personalized Medicine at University of Colorado **Aurora, CO**
Genetic Biostatistician **4/2017 – 8/2019**

Provided statistical genetics support for several longitudinal, multi-ethnic studies of cardiovascular disease and associated complex traits, with emphasis on analysis of human (epi)genetic data (genotype array, WGS, WES, and methylation array).

- Performed QC, analysis, and interpretation of high-throughput human omics data.
- Planned and executed high-throughput data analyses, including genome-wide association (GWAS), epigenome-wide association (EWAS), methylation quantitative trait loci (mQTL), expression quantitative trait loci (eQTL), gene-by-environment interaction (GxE), Mendelian Randomization (MR), and two-step MR analyses.
- Applied methods for correction of population stratification and relatedness.
- Supported the development, application, and interpretation of polygenic risk scores.
- Contributed to the TOPMed for Precision Medicine initiative and other consortia.
- Performed whole genome imputation using 1000 Genomes reference panels.
- Consulted faculty and PhD students on statistical genetics methods.

PROGRAMMING

- R/RStudio (eg, renv, tidyverse, ggplot2, purrr, furrr, Bioconductor, ChAMP, limma, minfi, DMRcate, DESeq2)
- Python (eg, Numpy, Pandas)
- SAS/SAS Studio
- Bash scripting
- Command line tools (eg, Plink 1.9, Plink 2.0, vcftools, bcftools, bedtools, BWA, Bowtie)
- SQL
- HPC/Cloud/Parallel computing (AWS)
- Version control (eg, Git)

PROFESSIONAL SERVICE

1/2016 – 5/2017
Student Chapter of the American Statistical Association (ASA) at University of Colorado
Founder and President

8/2015-7/2016
University of Colorado at Denver, Department of Biostatistics and Informatics
MS Student-Faculty Representative

HONORS

4/2017
University of Washington Summer Institute in Statistical Genetics Scholarship

8/2009 – 5/2013
Purdue University Honors Program

EXPERIENCE (*continued*)

National Jewish Health — **Denver, CO**
R Programming Consultant — 4/2018 – 1/2020

Developed, documented, and explained code for data wrangling of large geospatial and patient data sets.

Lifecourse Epidemiology of Adiposity & Diabetes (LEAD) Center at University of Colorado Anschutz Medical Campus — **Aurora, CO**
Data Manager and Analyst — 3/2016 – 3/2017

Managed and analyzed clinical, genomic, and epigenetic data for two longitudinal epidemiological studies of adiposity-related outcomes in children and adolescents exposed to gestational diabetes.

Department of Biostatistics & Informatics at University of Colorado Anschutz Medical Campus — **Aurora, CO**
Course Developer — 1/2015 – 2/2016

Authored, promoted, and co-taught an R25-funded course on power and sample size estimation for longitudinal and multilevel study designs.

Colorado Summer Institute in Biostatistics (COSIBS) at University of Colorado Anschutz Medical Campus — **Aurora, CO**
Instructor, Statistical Estimation Theory Module — 7/2016 – 7/2016

Independently developed lesson plans and delivered lectures for a cohort of undergraduates considering graduate education in biostatistics.

Department of Biostatistics & Informatics at University of Colorado Anschutz Medical Campus — **Aurora, CO**
Head Graduate Teaching Assistant — 6/2015– 8/2015
Graduate Teaching Assistant — 8/2014 – 12/2014

Independently taught Applied Biostatistics I, SAS, R and SPSS courses. Developed new syllabi and course materials. Trained incoming Graduate Teaching Assistants.

University of California MIND Institute — **Davis, CA**
Junior Research Specialist — 11/2013 – 8/2014

Assisted with study recruitment, data management and analysis.

Purdue University Oncological Sciences Center — **West Lafayette, IN**
Epigenetics Database Intern — 9/2012 – 7/2013

Helped develop a database of environmental factors associated with epigenetic changes in breast cancer.

The Hershey Company — **Hershey, PA**
Category Management Co-Op — 1/2012 – 5/2012

Leveraged Nielsen hierarchical databases to consult on pricing strategy and market dynamics for new product innovations.

The Nielsen Company — **Multiple**
Pharmaceutical Market Research Analyst — 7/2013 – 11/2013
Professional Services Analytics Intern (Client: Kraft) — 6/2012 – 8/2012
Professional Services Analytics Intern (Clients: ConAgra, Kellogg) — 5/2011 – 8/2011

Leveraged Nielsen massively hierarchical databases and proprietary analytics tools to provide marketing analytics consulting services.

PUBLICATIONS

Reference	Data type(s)	Language(s) used
Araya, P, **Shaw, JR**, …, Espinosa, JM. "Single cell immune profiling identifies rare cell types distinguishing patients with and without Trisomy 21." Manuscript in preparation.	• Single cell time of flight (CyTOF) • Medical record (MR) • Participant questionnaire	• Python • R/RStudio
Glueck, DH, Li, Q, Macleod, AJ, Litkowski, EM, Zamperlini, M, Yang, X, Bia, Jiang, Ritzhaupt, A, Sommer, M, Lourinho Moura do Valle, N, **Shaw, JR** , Muller, KE. Power and Sample Size for Balanced Linear Mixed Models with Clustering and Longitudinality: GLIMMPSE 3.0. *Journal of Statistical Software*. Submitted.	• Epidemiological • Simulated	• SAS • R/RStudio
Stahl, MG, **Shaw, JR**, Neetha Paul Eduthan, Angela L Rachubinski, Keith P Smith, Belinda Enriquez Estrada, Matthew D Galbraith, Ronald J Sokol, Sameer Chavan, Laura Ann Leaton, Katherine M Kichula, Paul J Norman, Jill M Norris, Edwin Liu, & Espinosa, JM. (2022). Multi-omics assessment of genetic risk for celiac disease in down syndrome. *medRxiv*. https://doi.org/10.1101/2022.09.27.22280436	• Genotype array • BeadArray (*.idat) • WGS • Imputed genomes • Imputed 5-digit HLA genotypes • Proteomic • Transcriptomic • Multiplex cytokine • Metabolomic • Medical record (MR) • Participant questionnaire	• R/RStudio • Shell scripting • Command line tools
Araya, P, **Shaw, JR**, Smith, KP, Waugh, KA, Granrath, RE, Enrique-Estrada, B, Worek, K, Kinning, KT, Sullivan, KD, Jordan, KR, Galbraith, MD, Rachubinski, AL, Espinosa, JM. Decoding trisomy 21 as a risk factor for severe COVID19: immune impacts of karyotype, premature senescence and comorbidities. Submitted.	• Genotype array • WGS • Proteomic • Transcriptomic • Multiplex cytokine • Metabolomic	• R/RStudio
Lee, M, Huan, T, McCartney, DL, Chittoor, G, de Vries, M, Lahousse, L, Nguyen, JN, Brody, JA, Castillo-Fernandez, J, Terzikhan, N, Qi, C, Joehanes, R, Min, JL, Smilnak, GJ, **Shaw, JR**, Yang, CX, Colicino, E, Hoang, TT, Bermingham, ML, Xu, H, Justice, AE, Xu, CJ, Rich, SS, Cox, SR, Vonk, JM, Prokić, I, Sotoodehnia, N, Tsai, PC, Schwartz, JD, Leung, JM, Sikdar, S, Walker, RM, Harris, SE, van der Plaat, DA, Van Den Berg, DJ, Bartz, TM, Spector, TD, Vokonas, PS, Marioni, RE, Taylor, AM, Liu, Y, Barr, RG, Lange, LA, Baccarelli, AA, Obeidat, M, Fornage, M, Wang, T, Ward, JM, Motsinger-Reif, AA, Hemani, G, Koppelman, GH, Bell, JT, Gharib, SA, Brusselle, G, Boezen, HM, North, KE, Levy, D, Evans, KL, Dupuis, J, Breeze, CE, Manichaikul, A, London, SJ. (2022). Pulmonary function and blood DNA methylation: A multiancestry epigenome-wide association meta-analysis. *American Journal of Respiratory and Critical Care Medicine, 206*(3), 321–336. https://doi.org/10.1164/rccm.202108-1907OC	• Methylation microarray • Medical record (MR) • Participant questionnaire	• R/RStudio • Shell scripting Command line tools

PUBLICATIONS (*continued*)

Reference	Data type(s)	Language(s) used
Galbraith, MD, Kinning, KT, Sullivan, KD, Araya, P, Smith, KP, Granrath, RE, **Shaw, JR**, Baxter, R, Jordan, K, Russell, S, Dzieciatkowska, M, Reisz, J, Gamboni, F, Cendali, F, Monte, A, Bennett, T, Hansen, K, Hsieh, EWY, D'Alessandro, A, Espinosa, JM. (2022). Specialized interferon action in COVID-19. *Proceedings of the National Academy of Sciences, 119*(11), e2116730119. https://doi.org/10.1073/pnas.2116730119	• Antibody titers • Proteomic • Transcriptomic • Multiplex cytokine • Metabolomic	• R/RStudio
Sullivan, KD, Galbraith, MD, Kinning, KT, Bartsch, KW, Levinsky, NC, Araya, P, Smith, KP, Granrath, RE, **Shaw, JR**, Baxter, RM Jordan, KR, Russell, SA, Dzieciatkowska, ME, Reisz, JA, Gamboni, F, Cendali, FI, Ghosh, T, Monte, AA, Bennett, TD, Miller, MG, Hsieh, EW, D'Alessandro, A, Hansen, KC, Espinosa, JM. (2021). The COVIDome Explorer researcher portal. *Cell Reports, 36*(7), 109527. https://doi.org/10.1016/j.celrep.2021.109527	• Antibody titers • Proteomic • Transcriptomic • Multiplex cytokine • Metabolomic	• R/RStudio
Lee, MK, Stanislawski, MA, Litkowski, E, Raghavan, Sridharan, Harrall, KK, **Shaw, JR**, Glueck, DH, Lange, EM, Dabelea, D, Lange, LA. (2021). Genetic risk score for type 2 diabetes and traits related to glucose-insulin homeostasis in youth: The exploring perinatal outcomes among children (EPOCH) study. *Diabetes Care, 44*(9), 2018–2024. https://doi.org/10.2337/dc21-0464	• Genotype array • Imputed genomes • Medical record (MR) • Participant questionnaire	• R/RStudio • Shell scripting • Command line tools
Liggett, LA, Galbraith, MD, Smith, KP, Sullivan, KD, Granrath, RE, Enriquez-Estrada, B, Kinning, KT, **Shaw, JR**, Rachubinski, AL, Espinosa, JM, DeGregori, J. (2021). Precocious clonal hematopoiesis in Down syndrome is accompanied by immune dysregulation. *Blood Advances, 5*(7), 1791–1796. https://doi.org/10.1182/bloodadvances.2020003858	• Genotype array • Imputed genomes	• R/RStudio • Shell scripting • Command line tools
Galbraith, MD, Kinning, KT, Sullivan, KD, Baxter, R, Araya, P, Jordan, KR, Russell, S, Smith, KP, Granrath, RE, **Shaw, JR**, Dzieciatkowska, Monika, Ghosh, T, Monte, AA, D'Alessandro, A, Hansen, KC, Bennett, TD, Hsieh, EWY, Espinosa, JM. Seroconversion Stages COVID19 into Distinct Pathophysiological States. *ELife, vol. 10, Mar. 2021, p. e65508*. PubMed, doi:10.7554/eLife.65508	• Antibody titers Proteomic • Transcriptomic • Multiplex cytokine • Metabolomic	• R/RStudio
Raffield, LM, Lu, AT, Szeto, MD, Little, A, Grinde, KE, **Shaw, J**, Auer, PL, Cushman, M, Horvath, S, Irvin, MR, Lange, EM, Lange, LA, Nickerson, DA, Thornton, TA, Wilson, JG, Wheeler, MM, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, TOPMed Hematology & Hemostasis Working Group, Zakai, N. A., & Reiner, A. P. (2020). Coagulation factor VIII: Relationship to cardiovascular disease risk and whole genome sequence and epigenome-wide analysis in African Americans. *Journal of Thrombosis and Haemostasis: JTH, 18*(6), 1335–1347. https://doi.org/10.1111/jth.14741	• Methylation microarray • Medical record (MR) • Participant questionnaire	• R/RStudio • Shell scripting • Command line tools
Stanislawski, MA, **Shaw, JR**, Pern, W, Lange, EM, Dabelea, D, & Lange, LA. (2020). Genetic risk for hepatic fat among an ethnically diverse cohort of youth: The exploring perinatal outcomes among children study. *The Journal of Pediatrics, 220*, 146-153.e2. https://doi.org/10.1016/j.jpeds.2020.01.031	• Genotype array Imputed genomes • Medical record (MR) • Participant questionnaire	• R/RStudio • Shell scripting • Command line tools

PUBLICATIONS (*continued*)

Reference	Data type(s)	Language(s) used
Shaw, JR (2017). *Double generalized linear model and nonparametric alternatives for the detection of variability quantitative trait loci, The*. CU Anschutz Digital Collections. https://doi.org/10.25677/sfbh-tz79	• Genotype array	• R/RStudio • Shell scripting • Command line tools
Shaw, JR, Muller, KE, Glueck, DH. (2015). Selecting a valid sample size for longitudinal and multilevel study designs- short course. Informatics Institute. https://informatics.research.ufl.edu/selecting-a-valid-sample-size-for-longitudinal-and-multilevel-study-designs-short-course.html	• Epidemiological • Simulated	• SAS • R/RStudio

PRESENTATIONS

Reference	Audience composition
Shaw, JR. Finding 'good enough': Statistical model building for real-world data. VitriVax Biosciences. March 2024.	• Engineers • Wet lab scientists • Executives • Operations support staff
Shaw, JR, Guan, Li, Hsu, Chien. Computational efficiency in R: Tips, tricks, and strategies. Programming Forum, Illumina Division of Biostatistics and Clinical Data Science. March 2023.	• Biostatisticians • Bioinformaticians • Executives
Shaw, JR. Quantile estimation in SAS, SAS/IML, R, JMP. Programming Forum, Illumina Division of Biostatistics and Clinical Data Science. December 2022.	• Biostatisticians • Clinical data scientists
Shaw, JR. Computational workflow for manual gating of immune cell populations in CyTOF data. University of Colorado Anschutz Anschutz Code Review. July 2020.	• Bioinformaticians • Research data scientists
Shaw, JR. Methods and considerations for integrative multi-omics analysis. Thrombosis Genetics in African Americans (TGEN) working group (5R01HL132947-03). October 2018.	• Biostatisticians • Bioinformaticians • Epidemiologists • Clinicians
Shaw, JR, Raffield, LM, Lange LA, Valdar W, Lange, EM. Deconstructing a variant by sex interaction: Sex as a collection of environmental factors. Colorado Center for Personalized Medicine Annual Retreat. September 2018.	• Biostatisticians • Bioinformaticians • Epidemiologists • Clinicians • Wet lab scientists
Shaw, JR, Corty RW, Raffield, LM, Reiner, AP, Wilson, JG, Lange, LA, Valdar, W, Lange, EM. Evaluating the role of genetic variants on white blood cell count variability in the Jackson Heart Study and the Atherosclerosis Risk in Communities Study. American Society for Human Genetics Annual Meeting, Orlando, FL, October 2017.	• Biostatisticians • Bioinformaticians • Epidemiologists • Clinicians • Wet lab scientists