

# Samir Rachid Zaim, PhD

samir.rachid.zaim@gmail.com · (520) 719-9224 · [LinkedIn](#) · [GitHub](#)

Applied researcher and technical leader with 10+ years of experience with > 10 completed projects & publications, and individual contributor with management experience. Excellent team player with a proven track record of delivering projects in high-pace biotech research environments. Seeking to contribute to scientific drug discovery, RWE, technology development, and improved patient care.

## Professional Experience

### **Supervisor, Biostatistics**

*Jan 2025 - Present*

*Allen Institute, Immunology. Seattle WA*

- Lead a team of 3 junior bioinformaticians supporting 5 cross-functional clinical projects
- Owned statistical strategy and SAP development for 10+ studies, tracking costs and milestones
- Reviewed novel AI & statistical approaches and onboarded advanced methods across studies
- Directed analysis of a complex oncology study; presented findings to executive leadership and advisory board and delivered a high-priority manuscript

#### **Biostatistician, II**

*10/2023-12/2024*

- Served as senior IC providing biostatistical leadership across 10+ in-house clinical studies
- Co-led onboarding of novel proteomics technologies to support a biomarker patent application
- Published a study and patent application predicting arthritis onset with the Dept. of Defense
- Lead a survey study assessing impact of Long COVID on quality of life in Latino communities

#### **Biostatistician, I**

*3/2021-09/2023*

- Awarded 1 of 6 Internal Allen Institute awards ('23) to expand research in auto-immune diseases
- Led development of a novel statistical R package for functional genomic inference
- Designed statistical study for an NIH grant to molecularly profile post-clinical trial participants

### **Center for Biomedical Informatics & Biostatistics Researcher**

*8/2016 - 1/2021*

*Univ. of Arizona, Center for Biomedical Informatics and Biostatistics, Tucson AZ*

- Led and contributed to 8 novel methods in statistical machine learning approaches in medicine
- Recognized by the [IMIA yearbook on Bioinformatics](#) as a top research contributor in 2018
- Developed scientific communication skills presenting 10+ conference talks & posters, and two career columns in *Nature Magazine*

### **Collegiate Fellow & Data Scientist**

*7/2014 - 5/2016*

*Parkland Center for Clinical Innovations & Pivotal Technologies, Dallas TX*

- Designed a risk-surveillance dashboard to identify at-risk asthmatic children
- Co-developed a sepsis monitoring surveillance algorithm to triage patients and improve care
- Deployed a predictive model to identify 30-day readmissions to optimize resource allocation

## Core Skills

### **Biostatistics and Machine Learning Experience:**

- Cross-sectional & longitudinal data clinical studies, study design and power analysis, EHR data mining, statistical modeling, classifier development, mixed effect modeling, GLMMs, survival modeling, regression, survey design, survey data analysis, tensor decomposition,

supervised & unsupervised learning, high-dimensional data clustering, high-dimensional feature selection, high-dimensional data visualization, statistical & ml software development

#### **Programming Experience:**

- Language: R (10+ yrs), Python (2+ yrs), SAS & STATA (1-2 yrs), SQL (1-2 yrs),
- Environment: Linux (> 5 yrs), HPC (5 yrs), Google Cloud Platform (3+ yrs)

#### **Technical Leadership Experience:**

- Statistical consulting, Statistical Analysis Plans, Software benchmarking, project management, cross-functional statistical collaborations

#### **Technical Leadership Management:**

- Requirements gathering, milestone tracking, resource & budget management
- Cross-functional leadership, technical project planning, technical
- Jira, Trello, Git/GitHub, Tableau, Slack, Confluence, Jupyter Notebooks, R Markdown

### Patents

**BIOMARKERS TO PREDICT ONSET OF CLINICAL RHEUMATOID ARTHRITIS, U.S. PROVISIONAL PATENT**  
Application No. 63/597,631

### Awards, Honors, and Recognition

<b>Editor's Pick Articles:</b> Arthritis & Rheumatology, September 2025 Edition	07/2025
<b>Commissioned Editorial,</b> Arthritis & Rheumatology	09/2025
<b>Post-baccalaureate Research Grant (&gt; \$57,500),</b> Allen Institute, WA	06/2023
<b>Data Science Ambassador: (\$1000),</b> Univ. Arizona College of Medicine, AZ	09/2019
<b>Top Contributor in Bioinformatics,</b> International Medical Informatics Association, France	2018
<b>Center for Biomedical Informatics and Biostats Fellowship: (~\$113,336),</b> Univ. Arizona.	06/2016
<b>Fellowship in Public Interest Communications: (\$3,000),</b> NonProfit Roundtable, DC	06/2013
<b>Kellogg's International Scholar: (~\$180,000),</b> Carleton College, MN	09/2010

### Open-source Statistical and Machine Learning Software Packages

**MOCHA:** a statistical algorithm for functional genomic inference: [CRAN](#)

**binomialRF:** a machine learning tool to identify genomic biomarker and interactions: [CRAN](#), [GitHub](#)

### Selected Publications

- "Serum proteomic signatures before the diagnosis of rheumatoid arthritis..." *Arthritis & Rheumatology* 77.9 (2025). **Led biomarker study in collaboration with the Dept. of Defense**
- "MOCHA's advanced statistical modeling of scATAC-seq data enables functional genomic inference in large human cohorts." *Nature Communications*(2024). **Novel statistical software**
- "binomialRF: interpretable combinatoric efficiency of random forests to identify biomarker interactions." *BMC bioinformatics*(2020). **Novel machine learning software development.**

### Education

**2021 Ph.D.:** Statistics, *Univ. of Arizona, Tucson, Arizona*

**2014 BA:** Mathematics/Statistics, *Carleton College, Northfield, MN*