Silvia Diz de Almeida

BIOSTATISTICIAN, PHD IN HUMAN GENETICS

Biostatistician with 4+ years in genomics research and 5+ years of expertise in R programming. Experienced in regression modelling, risk prediction, and biomarker identification for large-scale biomedical studies. Seeking to apply statistical expertise to clinical trials and RWE studies.

silviadiz03@gmail.com
+3462705648
Santiago de Compostela (Spain)
https://silviaadiz.github.io/
in/silvia-diz/

PROFESSIONAL EXPERIENCE

Biostatistician - Researcher

CiMUS (University of Santiago de Compostela), Genomic Medicine Group

Jun 2020-Jul 2024
Santiago de Compostela

- Conducted statistical analysis of genomic and biomedical data, identifying and prioritizing candidate genes associated with various diseases.
- Executed and designed analyses for the SCOURGE consortium, involving multi-country clinical and genetic data from +10,000 patients, to identify genetic determinants underlying COVID-19 severity.
- Implemented and validated regression models for disease risk prediction and stratification and assessed Polygenic Risk Scores (PRS) as potential personalized medicine tools.
- Applied multiple statistical methods, including hypothesis testing, descriptive statistics, fixed-effects meta-analysis, PCA, mixed models, GLMs and regularized regression, model validation.
- Collaborated in interdisciplinary teams across national and international projects, resulting in co-authorship of +10 publications, 3 of them as first author.
- Built and automated R and bioinformatic pipelines for large-scale genomic studies in complex diseases.
- Mentored junior researchers in R programming, statistical methods, and bionformatics. Co-supervised a MSc thesis.
- Conducted data harmonization and pre-processing of large databases using R.
- Communicated to non-specialized audiences and participated in the dissemination and reporting of results.

Research intern

CiMUS (University of Santiago de Compostela), Genomic Medicine Group

Feb 2020-Jun 2020
Santiago de Compostela

- Assisted in database wrangling and data preparation using R.
- Conducted quality control of genomic data and researched statistical genetic methods.

EDUCATION

- PhD in Molecular Medicine (Statistical Genetics; Bioinformatics) - University of Santiago de Compostela, 2024
- MSc in Statistical Techniques - University of Santiago de Compostela, 2021
- BSc in Biology - University of Santiago de Compostela, 2015

SKILLS

Exploratory data analysis	Regression modeling	Case-control studies
Univariate and multivariate statistics	Statistical learning	Data cleansing and visualization

Programming

Advanced R | Basic python | Intermediate bash | Bioinformatics software

Other skills

Adaptability | Problem-solving | Autonomy | Proactivity | Spanish (Native) | English (Cambridge C2 certificate)

COURSES

- Hands on clinical reporting using R (Genentech - Coursera) - November 2024
  - ADAM datasets, statistical programming, admiral, TLFs, NEST, teal
- A crash course in causality (University of Pennsylvania - Coursera) - December 2024
  - Causality, propensity score matching, IPTW, IVs

PROJECTS AND RESEARCH

- **Main projects**
  - *Genetic determinants of COVID-19 (SCOURGE consortium)* - 2020-2024
    Role: Co-lead biostatistician.
    Summary: International project aimed at studying the genetic basis of COVID-19 and the development of genetic tools for personalized medicine.
  - *Genetics of aggresive periodontitis (PERIOGEN study) - 2022-2024*
    Role: Assistant biostatistician.
    Summary: National project focused on identifying genetic factors contributing to aggressive periodontitis.
- **Teaching and mentorship**
  - *STEMbach program* - 2023-2024
    Role: Project co-supervisor.
    Summary: Excellence program in STEM where high school students develop a research project in collaboration with university researchers.
  - Supervision of a *Bioinformatics for health sciences* MSc Thesis - 2023
    Role: Thesis co-director.
    Summary: "*Comparison of genetic imputation in individual cohorts versus joint imputation: Involved factors and proposed actions*."
- **Seminars and talks**
  - *MSc in Biomedical Science*: Seminars in R programming, pharmacogenetics and polygenic risk scores.
  - *Basic statistics and R programming workshop:* session for high school students.
  - Seminar: "*Polygenic Risk Scores for COVID-19 Hospitalization in Spanish and Latin-American Populations*" at the Foundation Health Research Institute, Santiago de Compostela.
  - Talk: "*The Jessica Jones of DNA: How statistics help uncover the mysteries of complex phenotypes*" at a high school.
- **Publications**
  - ORCID: 0000-0003-2813-8928