

# Silvia Diz de Almeida

silviadiz03@gmail.com | +34627056438 | Santiago de Compostela (Spain)

Personal site: <https://silviaadiz.github.io/>

## BIOSTATISTICIAN

Biostatistician with 4+ years in genomics research and 5+ years of expertise in R programming. Experienced in statistical analysis, genetic risk prediction, and biomarker identification for large-scale biomedical studies. Proven ability to mentor junior researchers and contribute to high-impact publications. Seeking to apply statistical expertise to drug development and clinical trials in the pharmaceutical industry.

---

## PROFESSIONAL EXPERIENCE

Center for Research in Molecular Medicine and Chronic Diseases - Genomic Medicine Group

### Biostatistician | Research technician

Jun 2020-Jul 2024

- **Led** statistical analyses to identify **genetic biomarkers** with potential for clinical and drug development applications. **Discovered five novel genes** linked to COVID-19 severity, small-cell lung cancer, and aggressive periodontal disease, contributing to three first-author publications in high-impact journals
- Executed and designed analyses for the COVID-19 host genetics SCOURGE consortium, involving **multi-country clinical and genetic data from +10,000 patients**, to identify genetic determinants of disease severity.
- Developed **risk prediction models** on a large patient cohort. Assessed the **incremental value of genetic biomarkers in clinical risk stratification**, supporting their potential integration into precision medicine applications.
- **Applied multiple statistical methods** such as hypothesis testing, classification and clustering, fixed-effects meta-analysis, PCA, logistic mixed models and regularized regression, ANOVA, univariate analysis and descriptive statistics.
- Collaborated with **interdisciplinary teams**, including clinicians, geneticists, odontologists and psychology researchers, across national and international projects. Contributed with both design and execution of studies, resulting in successful project completions and co-authorship on +10 publications.
- **Mentored junior researchers** in R programming, statistical methods, and bioinformatics, including the co-supervision of a MSc thesis awarded a grade of 10/10.
- Built and **automated R pipelines** to handle statistical analyses more efficiently, creating **reusable scripts** to eliminate redundancy and allowing the team to run analyses across multiple projects simultaneously.
- **Harmonized** and cleaned a **multi-center dataset involving 30+ research units** for large-scale analysis.

### Intern

Feb 2020-Jun 2020

- Assisted in database wrangling and data preparation using R, supporting data analysis workflows.
  - Acquired and applied research methodologies in bioinformatics and genomics.
- 

## EDUCATION

**PhD** in Molecular Medicine - University of Santiago de Compostela - 2024

- Statistical Genetics; Bioinformatics (*Summa Cum Laude*)

**MSc** in Statistical Techniques - University of Santiago de Compostela - 2021

- Relevant courses: multivariate analysis, spatial analysis, probability models, mixed models, parametric and non parametric regression, survival analysis, time series

**BSc** in Biology - University of Santiago de Compostela

---

## SKILLS

**Statistics and programming:** Advanced R, R-markdown, exploratory data analysis, frequentist statistics, regression modelling (GLMs, mixed models), classification, parametric and non parametric statistics, data visualization, analysis of variance, survival analysis, multivariate statistics, statistical learning, bash, beginner python.

**Soft skills:** Time management, adaptability, self-learning, resourcefulness.

## COURSES AND CERTIFICATIONS

- 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
  - Python Programming (Udemy) - In progress
  - Certificate of Proficiency in English - Cambridge C2 (2021)
- 

## PROJECTS

### Genetic determinants of COVID-19 (SCOURGE consortium):

Role: Co-lead biostatistician.

- Worked with genetic and clinical data from +10,000 patients recruited in Spain and 6 countries from Latin-America to identify genetic variants associated with the disease and integrated expression data to prioritize potential therapeutic targets such as the gene *AQP3*.
- Designed and executed computational experiments to evaluate Polygenic Risk Scores (PRS) as potential biomarkers for COVID-19, assessing their clinical utility for personalized risk stratification; identified limitations and modest predictive improvements (2%) over baseline models.
- Tested an alternative approach for training PRS models in genetically diverse cohorts based on genetic similarity between individuals.
- Explored the behaviour of the genetic score in relation to key clinical variables and related subgroups, such as age, sex, and relevant comorbidities, and its association with other COVID-19 complications.
- Actively participated in subprojects derived from the study.

### Genetics of aggressive periodontitis (PERIOGEN study):

- Role: Assistant biostatistician.
    - Performed quality control of genetic data.
    - Contributed with R workflows for polygenic risk score development and analysis.
    - Assisted in drafting the grant proposal that secured funding for the Phase II study, contributing to the design and analysis plan for the continuation of the project.
  - Role: Co-lead biostatistician.
    - Coordinated a remote collaboration, managing task distribution and timelines for analysis, report generation, and manuscript preparation.
    - Meta-analyzed cohort data from large genetic association studies to identify novel biomarkers for the disease.
- 

## SELECTED PUBLICATIONS

ORCID: 0000-0003-2813-8928

- **Diz-de Almeida, S.**, Cruz, R., ..., Carracedo, A. (2024) Novel risk loci for COVID-19 hospitalization among admixed American populations. *eLife* 13:RP93666.
- **Diz-de Almeida, S.**, Richter, G., ... ,Cruz, R., Schaefer, A. S. (2023). A genome-wide association study meta-analysis in a European sample of stage III/IV grade C periodontitis patients 35 years of age identifies new risk loci. *Journal of clinical periodontology*. <https://doi.org/10.1111/jcpe.13922>
- **Diz-de Almeida, S.**, Cruz, R., ... , Flores, C., Carracedo, A. (2022). Novel genes and sex differences in COVID-19 severity. *Human molecular genetics*, 31(22), 3789–3806. <https://doi.org/10.1093/hmg/ddac132>
- Alemany-Navarro, M., **Diz-de Almeida, S.**, ... ,Carracedo, A. (2023). Psychiatric polygenic risk as a predictor of COVID-19 risk and severity: insight into the genetic overlap between schizophrenia and COVID-19. *Translational psychiatry*, 13(1), 189. <https://doi.org/10.1038/s41398-023-02482-7>