

Mathilde Sautreuil

Biostatistician & Data Scientist

Work

- MAR 2025 - JUN 2025 **Study Engineer, SOLADIS Inc. (Boston area), Consultant as Data Scientist - Sanofi.**
Development of machine learning (ML) models combined with molecular features to characterize and identify novel lipid nanoparticle (LNP) components:
- Provide an interactive platform (Streamlit application) to explore and visualize lipid data
 - Develop Generative AI algorithms for new lipid compounds
 - Integrate molecular dynamics and machine learning projects
- JAN 2024 - FEB 2025 **Study Engineer, SOLADIS Inc. (Boston area), Consultant as Biomarker statistician - Sanofi.**
In support of the Biomarker Statistics team: Evaluation of potential biomarkers in immunology projects:
- Biomarker statistician lead on a Phase 1 study: Development of the Biomarker strategy
 - Participate in all meetings with the Phase 1 team
 - Define and implement the Biomarker strategy, including cytokine analysis, classical RNA-seq analysis (differential expression and pathway analysis), signature analysis, and deconvolution analysis
 - Collaborate closely with the lead clinical statistician and Biomarker programmers
 - Write documents such as the Statistical Analysis Plan (SAP) and Data Transfer Specifications (DTS)
 - Omics exploratory data analysis to support various studies within the Biomarker Statistics team
- FEB 2022 - DEC 2024 **Study Engineer, SOLADIS (EFOR Group), Consultant as Biomarker Biostatistician - Servier.**
In support of the Biomarker Biostatistics team: Evaluation of potential biomarkers in immunology, oncology, and immuno-oncology projects:
- Benchmarking of deconvolution methods on non-standard data
 - Updating the deconvolution section of the RNA-seq and microarray data analysis pipeline
 - Data analysis on various projects, including (non-exhaustive list):
 - Deconvolution analysis to study immune cell heterogeneity pre- and post-treatment
 - Characterization of signatures and indication selection in oncology and immunology (Analysis of RNA-seq, microarrays, and cytokine datasets)
 - Application of Machine Learning methods on different projects to classify responder and non-responder patients
 - Development of a prediction model to evaluate the toxicity of new compounds and implementation in an RShiny application
- OCT 2020 - OCT 2021 **Research Engineer in Biostatistics, with Dr Guillem Rigau and Dr Etienne Delannoy**
Stat and Genome team, LaMME - Gnet team, IPS2, Paris-Saclay university.
The subject concerns the detection of PPR footprints using machine learning methods at Arabidopsis Thaliana.
- OCT 2017 - OCT 2020 **PhD student in Biostatistics, supervised by Pr Paul-Henry Cournède and Dr Sarah Lemler**
Biomathematics team - MICS Laboratory, CentraleSupélec.
The subject concerns survival analysis and marker detection in oncology and the main objectives are:
- the study of regularization and screening methods to detect the biological markers in the case of renal cancer;
 - the study and the development of neural networks to predict survival duration in high-dimension.
- 2015 - 2017
21 months **Apprentice in Biostatistics, supervised by Dr Caroline Bérard and Dr Nicolas Vergne**
TIBS Team, LITIS EA 4108 - Statistique Team, LMRS UMR 6085 CNRS, Rouen.
Development of statistical methods and tools for the analysis of Omics data.

Education

- 2017-2021 **PhD thesis in Applied Mathematics, MICS, CentraleSupélec, Paris-Saclay University, FRANCE.**
- 2014-2017 **Sandwich Master's Degree in Bioinformatics with distinction, Normandie University, Rouen, FRANCE.**
- 2011-2014 **Bachelor's Degree in Mathematics, Rouen University, FRANCE.**

Tool development for users

- RShiny app **MIDAS: Data visualization and consumption application used at the mRNA Center of Excellence,**
Sanofi internal tool.
- RShiny app **ToxShiny: Predict the toxicity on new compounds,** Servier internal tool.
- R package **survMS: an R package for the simulation of survival data from Cox, AFT and AH models,**
disponible sur le cran et sur <https://github.com/mathildesautreuil/survMS>.
- Python package **NNsurv: Neural Networks based on a discrete-time model to predict the survival duration,**
<https://github.com/mathildesautreuil/nnsurv>.
- R Package **SMM: An R Package for Estimation and Simulation of Discrete-time semi-Markov Models,**
<https://cran.r-project.org/web/packages/SMM/>.
- RShiny application **PROqPCR: PROcessing of qRT-PCR data,**
<https://qpcrapp.shinyapps.io/proqpcr/>.

1 Leighton Street – 02141 Cambridge, MA – United States

☎ +1 617 460 0564 • ✉ mathilde.sautreuil@live.fr

🌐 mathildesautreuil.github.io/persopage