Sophia Yazzourh | CV

Department of Biostatistics, McGill University, Montreal (Canada)
sophiayazzourh@gmail.com • sophiayazzourh.github.io
French & Moroccan

Postdoctoral Researcher Since January 2025
McGill University – Department of Epidemiology, Biostatistics and Occupational Health.
Under the supervision of Erica E. M. Moodie.

Course Lecturer Fall 2026
EPIB 507 – *Biostatistics for Health Sciences*, McGill University.

At the intersection of applied mathematics, statistics, and machine learning, I develop methods for learning decision-making strategies from data, accounting for their sequential structure, uncertainty, and the possibility of alternative decisions. My research draws on reinforcement learning, Bayesian modeling, and statistical learning. A substantial part of my work also focuses on designing complex simulation frameworks and applying these methods to real-world data, primarily in healthcare.

Research Interests: Reinforcement Learning, Sequential Decision-Making, Bayesian Inference, Dynamic Treatment Regimes, Computational Evaluation of Statistical Methods, Longitudinal Data Analysis, Precision Medicine.

Education

- **PhD in Applied Mathematics** — **Institute of Mathematics of Toulouse**
  Decision Algorithms in Medical Research — *October 2021 - October 2024*
  "Reinforcement Learning and Bayesian Outcome-Weighted Learning for Precision Medicine. Integration of medical knowledge into decision algorithms."
  Supervised by Nicolas Savy and Philippe Saint Pierre in collaboration with Michael Kosorok and Nikki L.B. Freeman from the *University of North Carolina* (UNC) at Chapel Hill, USA.

- **Engineering Degree (Master's Level)** — **National Institute of Applied Sciences Toulouse**
  Specialization in Applied Mathematics & Statistics — *September 2014 - June 2021*
  - Machine Learning, High-Dimensional and Deep Learning, Artificial Intelligence Frameworks (NLP, Reinforcement Learning).
  - Innovation and Research Project on modeling a quantitative variable in the context of logistic regression using interpolation splines.

- **Erasmus** — **NTUA, Greece**
  Hosted by the Departments of Mathematics and Computer Science — *September 2016 - January 2017*
  European exchange program undertaken during the first semester of my third year in the Bachelor's program at National Technical University of Athens.

Industry

- **Apprenticeship in Big Data & AI for Space Activities** — **Thales Alenia Space**
  Research and Development in Artificial Intelligence — *January 2020 - June 2021*
  - Development of a Python library of algorithms for anomaly detection, applied to the space domain.
  - Implementation and adaptation of Active Anomalies Discovery methods on time-series signals: anomaly detection within the framework of Active Learning, particularly the integration of user feedback into algorithms such as Isolation Forest or SVM.

  Technical expertise: Python, PCA, Isolation Forest, SVM, Autoencoders, ROC Curves, Git, Docker, Flask API.

- **Apprenticeship in management and industrial processes** — **Thales Alenia Space**
  Project Management — *October 2019 - January 2020*
  Focus on information transmission and tool coherence through to the maintenance phase.
  Technical expertise: Life cycle of space ground segment projects and associated tools.

- **Master's Internship, GALILEO Mission Segment** — **SOGETI HIGH TECH, Capgemini Group**
  Integration and Validation — *July 2019 - September 2019*
  - Integration of a prototype element.
  - Data and flow analysis of connected components within the Mission and Uplink Control Facility (MUCF).
  - Analysis of the European Space Agency's specifications, preparation, and implementation of test scenarios within the assembly team.

  Technical expertise: Knowledge of the GALILEO Ground Segment and its Global Operation.

Grants and Funding

- **Grants Awarded**
  - **2021–2024 – Doctoral Grant from the French Ministry of Higher Education and Research, France.**
    Three-year doctoral funding awarded through a national competitive selection process.
- **Funding Secured for Scientific Event Organization**
  - **2026 – Association of African Women Scientists (AFSA) – Scientific Day: New Horizons in Health Statistics.**
    Funding secured: **CAD $4,700**.
    - School of Population and Global Health (McGill): CAD $2,400;
    - Canadian Statistical Sciences Institute (CANSSI): CAD $1,700;
    - STATQAM (UQAM): CAD $500;
    - AFSA: CAD $100.

    Project leads: Carmelle Chango, Sophia Yazzourh, Erica E. M. Moodie.

Publications

- **Published Article:**
  - **Yazzourh, S.**, Savy, N., Saint Pierre, P., Kosorok, M. (2024). *Medical Knowledge Integration into Reinforcement Learning Algorithms for Dynamic Treatment Regimes*. *International Statistical Review*. DOI: 10.1111/insr.12617.
- **Submitted and Ongoing Manuscripts:**
  - **Yazzourh, S.**, Freeman, N. L. B. *Bayesian Outcome Weighted Learning*. Preprint: arXiv:2406.11573.
  - **Yazzourh, S.**, Savy, N., Saint Pierre, P., Kosorok, M. *Rewards Construction Based on Preference Learning for Dynamic Treatment Regimes*.
  - **Yazzourh, S.**, Moodie, E. E. M. *Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes*. Preprint: arXiv:2603.19440. Code: GitHub.
  - **Yazzourh, S.**, Coulombe, J., Moodie, E. E. M., El-Riachi, D., Du, F., Jiao, T. *SPRINTing ahead: Practical Guidance for Readying Public Data for Longitudinal Analyses*. Manuscript in preparation.
  - **Yazzourh, S.**, Moodie, E. E. M., Shortreed, S. M. *The KPsynth: a synthetic reproducible benchmark for comparing adaptive treatment strategy methodology drawn from electronic health records on patients with depression*. Manuscript in preparation.
  - Gopinauth, T., **Yazzourh, S.**, Moodie, E. E. M. *Shared-Parameter Dynamic Weighted Ordinary Least Squares*. Manuscript in preparation.
  - **Yazzourh, S.**, Moodie, E. E. M. *Bayesian Near-Equivalent Q-learning for Dynamic Treatment Regimes*. Manuscript in preparation.
- **Doctoral Thesis:**
  - **Yazzourh, S.** (2024). *Reinforcement Learning and Bayesian Outcome-Weighted Learning for Precision Medicine. Integration of Medical Knowledge into Decision Algorithms*. University of Toulouse. Doctoral Thesis.

Oral Presentations

Upcoming Presentations

- **International Laboratory on Learning Systems (ILLS) Seminar – École de technologie supérieure (ÉTS)** (*Montreal, November 5, 2026*): *"From Q-learning to Near-Equivalent Policies: Expanding Treatment Options in Precision Medicine"*.
  Invited presentation.
- **20th International Joint Conference CFE-CMStatistics 2026** (*Berlin, Germany, December 13, 2026*): *"Bayesian outcome-weighted learning"*.
  Presentation in the thematic session *Estimation of heterogeneous treatment effects*, organized by Denis Talbot.
- **Biostatistics Team Seminar – Bordeaux Population Health (BPH)** (*Bordeaux, France, January 7, 2027*).
  Invited presentation.
- **Annual Meeting of the Statistical Society of Canada (SSC)** (*Quebec City, 2027*): *"Bayesian Near-Equivalent Policies"*.
  Invited speaker in the session *Modern Methods for Heterogeneous and Individualized Treatment Effects*, organized by Guanbo Wang.

Invited Presentations

- **Annual Meeting of the Statistical Society of Canada** (*Hamilton, June 3, 2026*): *"Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes"*.
  Invited speaker in the session *Recent Advances in Dynamic Treatment Regime Inference*.
- **McGill Biostatistics Seminar** (*Montreal, February 18, 2026*) *"Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes"*
- **Statistics Seminar – University of Glasgow** (*Glasgow, January 7, 2026*) *"Reinforcement learning for precision medicine: Integrating medical knowledge into decision-making algorithms"*
- **Annual Conference of the Canadian Society for Epidemiology and Biostatistics (CSEB)** (*Montreal, August 13, 2025*) *"Construction of Multiple ϵ-Equivalent Recommendations for Dynamic Treatment Regimes"*

- **Centre de Recherches Mathématiques Seminar – Université de Montréal** (*Montreal, April 25, 2025*) *"Reinforcement Learning for Precision Medicine: Integrating Medical Knowledge into Decision Algorithms"*
- **Statistics Seminar – Université du Québec à Montréal** (*Montreal, April 3, 2025*) *"Integrating Medical Knowledge into Reinforcement Learning for Dynamic Treatment Regimes"*
- **Presentation to the PhairLab (Michael Kosorok)** (*UNC Chapel Hill, April 21, 2023*) *"Integration of Medical Knowledge into Reinforcement Learning for Dynamic Treatment Regimes"*

Contributed Presentations

- **Scientific Days of the Réseau Santé Numérique du Québec** (*Université Laval, March 2026*) *"Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes"*
- **Colloque Francophone Interfacultaire de Recherche en Biostatistiques (CFIRB)** (*CHUM, Université de Montréal, Montreal, August 10, 2025*) *"Construction of Multiple ϵ-Equivalent Recommendations for Dynamic Treatment Regimes"*
  Best Presentation Award.
- **Annual Meeting of the Statistical Society of Canada** (*Saskatoon, May 27–30, 2025*) *"Rewards Construction Based on Preference Learning for Dynamic Treatment Regimes"*
- **8th Annual (Bio)Statistics Research Day** (*McGill University, Montreal, May 9, 2025*) *"Bayesian Outcome-Weighted Learning"*
  Best Presentation Award.
- **10th Journées Statistiques du Sud** (*Université Toulouse III – Paul Sabatier, Toulouse, June 18, 2024*) *"Bayesian Outcome-Weighted Learning"*
- **Student Seminar – Institute of Mathematics of Toulouse** (*Université Toulouse III – Paul Sabatier, Toulouse, June 6, 2024*) *"Introduction to Reinforcement Learning"*
- **55th French Statistical Days (SFdS)** (*Université de Bordeaux, Bordeaux, May 27, 2024*) *"Reward Construction via Preference Learning for Reinforcement Learning Models Applied to Adaptive Treatment Strategies"*
- **Toulouse Seminar "Statistics for Biology" – INRAE** (*INRAE, Toulouse, November 15, 2022*) *"Reinforcement Learning: Application to Adaptive Treatment Strategies"*

Posters

- **Statistics and Optimization Seminar – Institute of Mathematics of Toulouse** (*Toulouse, December 6, 2022*) *"Incorporating Medical Expertise into Reinforcement Learning Models Applied to Adaptive Treatment Strategies"*

Teaching and Pedagogical Projects

- **Course Lecturer – EPIB 507: Biostatistics for Health Sciences** (*Fall 2026*)
  Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Canada.
  - Instructor responsible for a 3-credit university course for students in health sciences.
  - Theoretical and practical instruction covering:
    - Probability, distributions, and descriptive statistics;
    - Estimation, confidence intervals, and hypothesis testing;
    - Inference on means and proportions, ANOVA, and chi-square tests;
    - Correlation, simple and multiple linear regression;
    - Statistical programming and data analysis using `R`.
  - Development and revision of teaching materials, written examinations, and practical `R` examinations.
  - Supervision of group projects focused on critical analysis and presentation of scientific articles or public datasets.

- **Doctoral Teaching Instructor in Mathematics – University Institute of Technology, Mechanical Engineering and Manufacturing** (*2021–2024*)
  - Taught mathematics to first-, second-, and third-year undergraduate students through lectures, tutorials, and practical sessions.
  - Subjects taught:
    - Calculus and differential and integral analysis;
    - Applied linear algebra;
    - Probability and statistics;
    - Numerical methods and optimization;
    - Algorithms and numerical modeling.
  - Teaching in French and English.
  - Designed, adapted, and translated teaching materials (lectures, exercises, and practical sessions).

Academic Service and Community Engagement

Organization of Scientific Events

- **Organizer of a Thematic Session – Annual Meeting of the Statistical Society of Canada (SSC)** (*2027, upcoming*): Session entitled *"Evaluating and Guiding Therapeutic Strategies"*.
- **Co-organizer of the AFSA Scientific Day: New Horizons in Health Statistics** (*2026*): Scientific event dedicated to health statistics, including a panel discussion on scientific careers.
- **Organizer of the Student Seminar at the Institute of Mathematics of Toulouse** (*2022–2024*): Bi-monthly organization of scientific seminars for doctoral and postdoctoral researchers.
- **Member of the Scientific and Organizing Committees of the 10th Journées Statistiques du Sud** (*2024*): Contribution to the scientific program and event organization.

Peer Review and Evaluation Committees

- **Peer Reviewer** (*since 2026*): *Biometrics* and *The Annals of Applied Statistics*.
- **Judge – Canadian Statistics Student Conference (CSSC)** (*2026, 2027 upcoming*): Statistical Society of Canada (SSC).
- **Judge – Biostatistics Research Day** (*2026, 2027 upcoming*): McGill University.

Scientific Community Engagement

- **McGill Representative – Association of African Women Scientists (AFSA)** (*since 2025*): Co-organizing seminars and activities promoting scientific careers for African women.
- **President of the Student Association of the Mathematics Department, INSA Toulouse** (*2019–2020*): Student representation, event organization, and coordination of student initiatives.
- **Member of Scientific Societies**: French Statistical Society (SFdS) and Statistical Society of Canada (SSC).

Scientific Outreach and Community Involvement

- **Speaker – DECLICS Program** (*2022–2024*): Presentations on scientific career paths and careers in mathematics and engineering to secondary school students in France.
- **Volunteer – Ouverture Rencontres Évolution (ORE) Association** (*2014–2021*): Tutoring in mathematics, science, and French, and development of an introductory course in Python and algorithms.

Interests

- **Handball Player**:
  - **Current (Canada)** – Player for the Celtiques de Montréal in Division 1.
  - **Career in France** – National 3 level club player: training 3–4 times per week and weekly matches.
  - **University** – Participated in the French University Championships (2015, 2016, 2017, 2018, 2019). Head and captain of the INSA Toulouse women's team (2016–2019).
- **Development and AI Tools:** Git, Docker, Google Cloud, Obsidian, Mistral, Ollama.
- **Foreign Languages:** German, Chinese, and Greek.