Sophia Yazzourh | CV

Département de biostatistique, Université McGill, Montréal (Canada)
sophiayazzourh@gmail.com • sophiayazzourh.github.io
Française & Marocaine

Chercheuse postdoctorale Depuis janvier 2025
Université McGill – Département d'épidémiologie, biostatistique et santé au travail.
Sous la supervision d' Erica E. M. Moodie.

Chargée de cours Automne 2026
EPIB 507 – *Biostatistics for Health Sciences*, Université McGill.

À l'interface des mathématiques appliquées, de la statistique et de l'apprentissage automatique, je développe des méthodes pour apprendre des stratégies de décision à partir de données, en tenant compte de leur structure séquentielle, de l'incertitude et de la possibilité de décisions alternatives. Je mobilise notamment l'apprentissage par renforcement, la modélisation bayésienne et l'apprentissage statistique. Une part importante de mes recherches est également consacrée à la conception de cadres de simulation complexes et à l'application de ces méthodes à des données réelles, principalement dans le domaine de la santé.

Domaines de recherche : apprentissage par renforcement, décision séquentielle, inférence bayésienne, régimes de traitement dynamiques, évaluation computationnelle des méthodes, données longitudinales, médecine de précision.

Parcours académique

- **Thèse de doctorat en mathématiques appliquées** — **Institut de Mathématiques de Toulouse**
  Algorithmes de décision dans la recherche médicale — *Octobre 2021 - Octobre 2024*
  "Apprentissage par renforcement et outcome-weighted learning bayésien pour la médecine de précision. Intégration de connaissances médicales dans les algorithmes de décision."
  Sous la supervision de Nicolas Savy et Philippe Saint Pierre en collaboration avec Michael Kosorok et Nikki L.B. Freeman de l'*Université de Caroline du Nord* (UNC) à Chapel Hill, Etats-Unis.

- **Diplôme d'ingénieure, grade Master** — **Institut National des Sciences Appliquées Toulouse**
  Spécialité Mathématiques Appliquées & Statistique — *Septembre 2014 - Juin 2021*
  - Machine Learning, High Dimensional and Deep Learning, Artificial Intelligence Framework (NLP, Reinforcement Learning).
  - Projet d'Innovation et Recherche sur le thème de la modélisation d'une variable quantitative dans le cadre de la régression logistique au travers de splines d'interpolation.

- **Programme d'échange universitaire Erasmus** — **National Technical University of Athens, Grèce**
  Départements de mathématiques et d'informatique — *Septembre 2016 – Janvier 2017*
  Séjour d'études international effectué durant la troisième année du premier cycle universitaire.

Industrie

- **Alternance en Big Data & IA pour les activités spatiales** — **Thales Alenia Space**
  Recherche et Développement en Intelligence Artificielle — *Janvier 2020 - Juin 2021*
  - Développement d'une bibliothèque Python d'algorithmes pour la détection d'anomalies, appliquée au domaine spatial.
  - Mise en œuvre et adaptation de méthodes d'Active Anomalies Discovery sur des signaux temporels : détection d'anomalies dans le cadre de l'Active Learning, notamment l'intégration du retour utilisateur dans des algorithmes tels qu'Isolation Forest ou SVM.

  Compétences acquises : ACP, Isolation Forest, SVM, Autoencodeurs, courbes ROC, GIT, Docker, API Flask.

- **Alternance en gestion et processus industriels des segments sols spatiaux** — **Thales Alenia Space**
  Gestion de projets — *Octobre 2019 - Janvier 2020*
  Réflexion sur la transmission d'informations et la cohérence des outils jusqu'en phase de maintenance.
  Connaissances acquises : Cycle de vie des projets segments sols spatiaux et outils associés.

- **Stage de Master 1, GALILEO Mission Segment** — **SOGETI HIGH TECH, Groupe Capgemini**
  Intégration et Validation — *Juillet 2019 - Septembre 2019*
  - Intégration d'un élément prototype.
  - Analyse et modélisation des données et des flux des éléments connectés du composant Mission and Uplink Control Facility (MUCF).
  - Analyse des spécifications de L'Agence Spatiale Européenne, préparation et mise en place de scénarios de tests dans l'équipe d'assemblage.

  Connaissances acquises : Fonctionnement du segment sol de GALILEO

Bourses et financements

- **Bourses obtenues**
  - **2021–2024 — Contrat doctoral du Ministère de l'Enseignement Supérieur et de la Recherche, France.** Bourse doctorale de trois ans obtenue après concours national.
- **Financements obtenus pour l'organisation scientifique**
  - **2026 — Association of African Women Scientists (AFSA) — Scientific Day: New Horizons in Health Statistics.**
    Financements obtenus : **4 700 $ CAD**.
    - School of Population and Global Health (McGill) : 2 400 $;
    - Canadian Statistical Sciences Institute (CANSSI) : 1 700 $;
    - STATQAM (UQAM) : 500 $;
    - AFSA : 100 $.

    Porteuses du projet : Carmelle Chango, Sophia Yazzourh, Erica E. M. Moodie.

Publications

- **Article publié :**
  - **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.
- **Articles soumis ou en préparation :**
  - **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*. Manuscrit en préparation.
  - **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*. Manuscrit en préparation.
  - Gopinauth, T., **Yazzourh, S.**, Moodie, E. E. M. *Shared-Parameter Dynamic Weighted Ordinary Least Squares*. Manuscrit en préparation.
  - **Yazzourh, S.**, Moodie, E. E. M. *Bayesian Near-Equivalent Q-learning for Dynamic Treatment Regimes*. Manuscrit en préparation.
- **Manuscrit de doctorat :**
  - **Yazzourh, S.** (2024). *Apprentissage par renforcement et outcome-weighted learning bayésien pour la médecine de précision. Intégration de connaissances médicales dans les algorithmes de décision*. Université de Toulouse. Manuscrit de doctorat.

Présentations orales

Présentations à venir

- **Séminaire de l'International Laboratory on Learning Systems (ILLS) – École de technologie supérieure (ÉTS)** (*Montréal, 5 novembre 2026*) : « *From Q-learning to Near-Equivalent Policies: Expanding Treatment Options in Precision Medicine* ».
  Présentation invitée.
- **20th International Joint Conference CFE-CMStatistics 2026** (*Berlin, Allemagne, 13 décembre 2026*) : « *Bayesian outcome-weighted learning* ».
  Présentation dans la session thématique *Estimation of heterogeneous treatment effects*, organisée par Denis Talbot.
- **Séminaire de l'équipe Biostatistics – Bordeaux Population Health (BPH)** (*Bordeaux, France, 7 janvier 2027*).
  Présentation invitée.
- **Congrès annuel de la Société canadienne de statistique (SSC)** (*Québec, 2027*) : « *Bayesian Near-Equivalent Policies* ».
  Conférencière invitée dans la session *Modern Methods for Heterogeneous and Individualized Treatment Effects*, organisée par Guanbo Wang.

Présentations invitées

- **Congrès annuel de la Société canadienne de statistique** (*Hamilton, 3 juin 2026*) : « *Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes* ».
  Conférencière invitée dans la session *Recent Advances in Dynamic Treatment Regime Inference*.
- **Séminaire de biostatistiques de McGill** (*Montreal, 18 février 2026*) *"Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes"*
- **Séminaire de statistiques – Université de Glasgow** (*Glasgow, 7 janvier 2026*) *"Reinforcement learning for precision medicine: Integrating medical knowledge into decision-making algorithms"*
- **Conférence annuelle CSEB 2025** (*Montréal, 13 août 2025*) *"Construction de multiples recommandations*

epsilon-équivalentes dans le cadre des stratégies dynamiques de traitement."

- **Séminaire du Centre de Rercheches Mathématiques – Université de Montréal** (*Montréal, 25 avril 2025*) *"Apprentissage par renforcement pour la médecine de précision : intégrer les connaissances médicales dans les algorithmes de décision."*
- **Séminaire de statistique – Université du Québec à Montréal** (*Montréal, 3 avril 2025*) *"Integrating Medical Knowledge into Reinforcement Learning for Dynamic Treatment Regimes."*
- **Présentation à l'équipe du PhairLab (Michael Kosorok)** (*UNC Chapel Hill, 21 avril 2023*) *"Integration of Medical Knowledge into Reinforcement Learning for Dynamic Treatment Regimes."*

Contributions

- **Journées Scientifiques du Réseau Santé Numérique du Québec** (*Université de Laval, mars 2026*) *"Near-Equivalent Q-learning Policies for Dynamic Treatment Regimes"*
- **Colloque Francophone Interfactultaire de Recherche en Biostatistiques (CFIRB)** (*CHUM, Université de Montréal, Montréal, 10 août 2025*) *"Construction de multiples recommandations epsilon-équivalentes dans le cadre des stratégies dynamiques de traitement."* **Prix de la meilleure présentation.**
- **Congrès annuel de la Société Canadienne de Statistique** (*Saskatoon, 27–30 mai 2025*) *"Rewards Construction Based on Preference Learning for Dynamic Treatment Regimes."*
- **8e Journée annuelle de recherche en (Bio)Statistiques** (*Université McGill, Montréal, 9 mai 2025*) *"Bayesian Outcome-Weighted Learning."* **Prix de la meilleure présentation.**
- **10e Journées Statistiques du Sud** (*Université Toulouse III – Paul Sabatier, Toulouse, 18 juin 2024*) *"Bayesian Outcome-Weighted Learning."*
- **Séminaire Étudiant – Institut de Mathématiques de Toulouse** (*Université Toulouse III – Paul Sabatier, Toulouse, 6 juin 2024*) *"Introduction to Reinforcement Learning."*
- **55e Journées de Statistique (SFdS)** (*Université de Bordeaux, Bordeaux, 27 mai 2024*) *"Construction de récompenses par apprentissage par préférences pour les modèles d'apprentissage par renforcement appliqués aux stratégies de traitements adaptatifs."*
- **Journée Toulousaine « Statistique pour la Biologie » – INRAE** (*INRAE, Toulouse, 15 novembre 2022*) *"Apprentissage par renforcement : application aux stratégies de traitements adaptatifs."*

Posters

- **Séminaire de Statistique et Optimisation – Institut de Mathématiques de Toulouse** (*Institut de Mathématiques de Toulouse, Toulouse, 6 décembre 2022*) *"Introduction d'Expertise Médicale dans les Modèles d'Apprentissage par Renforcement appliqués aux Stratégies de Traitements Adaptatifs."*

Enseignements et projets pédagogiques

- **Chargée de cours – EPIB 507 : Biostatistics for Health Sciences** (*Automne 2026*)
  Département d'épidémiologie, biostatistique et santé au travail, Université McGill, Canada.
  - Responsable d'un enseignement universitaire de 3 crédits destiné aux étudiants en sciences de la santé.
  - Enseignements théoriques et pratiques couvrant :
    - Probabilités, distributions et statistiques descriptives ;
    - Estimation, intervalles de confiance et tests d'hypothèses ;
    - Inférence sur les moyennes et les proportions, ANOVA et tests du Chi-deux ;
    - Corrélation, régression linéaire simple et multiple ;
    - Programmation statistique et analyse de données avec R.
  - Évolution des supports pédagogiques, des examens écrits et pratiques sur R.
  - Encadrement de projets de groupe consacrés à l'analyse critique et à la présentation d'articles scientifiques ou de jeux de données publiques.
- **Doctorante chargée d'enseignement en mathématiques – Institut Universitaire de Technologie, Génie Mécanique et Productique** (*2021–2024*)
  - Enseignement en mathématiques aux trois premières années du premier cycle universitaire, sous forme de cours, travaux dirigés et travaux pratiques.

- Disciplines enseignées :
  - Analyse, calcul différentiel et intégral ;
  - Algèbre linéaire appliquée ;
  - Probabilités et statistiques ;
  - Méthodes numériques et optimisation ;
  - Algorithmique et modélisation numérique.
- Enseignement en français et en anglais.
- Conception, adaptation et traduction de supports pédagogiques (cours, exercices et travaux pratiques).

Engagements et services à la communauté universitaire

Organisation de manifestations scientifiques

- **Organisatrice d'une session thématique – Congrès annuel de la Société canadienne de statistique (SSC)** (*2027, à venir*) : Session intitulée « *Évaluer et guider les stratégies thérapeutiques* ».
- **Co-organisatrice de l'AFSA Scientific Day: New Horizons in Health Statistics** (*2026*) : Journée scientifique consacrée à la statistique de la santé, incluant une table ronde sur les carrières scientifiques.
- **Responsable du séminaire étudiant de l'Institut de Mathématiques de Toulouse** (*2022–2024*) : Organisation bimensuelle de séminaires scientifiques pour doctorants et postdoctorants.
- **Membre des comités scientifique et d'organisation des 10e Journées Statistiques du Sud** (*2024*) : Participation à la programmation scientifique et à l'organisation de la manifestation.

Évaluation scientifique et jurys

- **Évaluatrice d'articles scientifiques (*Peer Reviewer*)** (*depuis 2026*) : *Biometrics* et *The Annals of Applied Statistics*.
- **Membre du jury – Congrès canadien des étudiants en statistique (CESS)** (*2026, 2027 à venir*) : Société canadienne de statistique (SSC).
- **Membre du jury – Journée de recherche en biostatistique (*Biostatistics Research Day*)** (*2026, 2027 à venir*) : Université McGill.

Engagements dans la communauté scientifique

- **Représentante à McGill – Association des Femmes Scientifiques Africaines (AFSA)** (*depuis 2025*) : Co-organisation de séminaires et d'activités visant à promouvoir les carrières scientifiques des femmes africaines.
- **Présidente de l'association étudiante du département de mathématiques de l'INSA Toulouse** (*2019–2020*) : Représentation étudiante, organisation d'événements et coordination de projets associatifs.
- **Membre de sociétés savantes** : Société française de statistique (SFdS) et Société canadienne de statistique (SSC).

Vulgarisation scientifique et engagement communautaire

- **Intervenante – Programme DECLICS** (*2022–2024*) : Présentation de parcours scientifiques et de métiers en mathématiques et en ingénierie auprès d'élèves du secondaire (lycées français).
- **Bénévole – Association Ouverture Rencontres Évolution (ORE)** (*2014–2021*) : Soutien scolaire en mathématiques, sciences et français et conception d'une formation d'initiation à Python et à l'algorithmique.

Centre d'intérêts

- **Joueuse de handball** :
  - **Actuel (Canada)** — Joueuse des Celtiques de Montréal en Division 1.
  - **Carrière en France** — Joueuse en club au niveau National 3 : entraînements 3–4 fois par semaine et matchs chaque week-end.
  - **Universitaire** — Participation aux Championnats de France Universitaires (2015, 2016, 2017, 2018, 2019). Responsable et capitaine de l'équipe féminine de l'INSA Toulouse (2016–2019)
- **Outils de développement et IA :** Git, Docker, Google Cloud, Obsidian, Mistral, Ollama.
- **Langues étrangères :** Allemand , Chinois et Grec.