

# Biostatistiques intermédiaires: Stratégies de modélisation de la régression multivariée

Partie 1

Plan de cours

2024-2025

## COURS EN LIGNE

### MODULE 5

Méthodologie

### DATE/HEURE

Jeudi 20 mars 2025

13h30 à 16h30 HAE

### LANGUE

English

### INSCRIPTION GRATUITE


[www.camccol.com](http://www.camccol.com)

## PRÉSENTÉ PAR



**Dr Steven Hawken**

Professeur agrégé, Université d'Ottawa

 [shawken@ohri.ca](mailto:shawken@ohri.ca)

## DESCRIPTION

Ce cours s'adresse aux statisticien-ne-s, épidémiologistes, scientifiques des données et autres chercheur-e-s ou étudiant-e-s ayant une connaissance de base de la modélisation par régression. Le cours couvrira les stratégies générales d'ajustement des modèles de prédiction pour les variables continues, catégoriques et temporelles, y compris : l'analyse exploratoire/la visualisation des données; l'imputation des données manquantes; la sélection des covariables; la spécification du modèle; la validation/calibration du modèle; la gestion de la non-linéarité; et le choix entre les modèles statistiques conventionnels et les modèles d'apprentissage automatique (et les différences entre ces types de modèles). Le matériel de cours et les études de cas/exemples feront un usage intensif de R, de RStudio et des logiciels `r Hmisc` et `rms` de Frank Harrell. Le cours suivra la philosophie générale du manuel *Regression Modelling Strategies*, 2<sup>ème</sup> édition par Frank Harrell.

## OBJECTIFS D'APPRENTISSAGE

- Méthodes d'exploration, de description et de compréhension de vos données en vue de la modélisation par régression
- Ajustement des modèles de régression multivariés appropriés pour les variables continues, catégorielles et de taux instantanés
- Aborder les questions de taille d'échantillon et de surajustement
- Approches pour traiter les données manquantes
- Traitement des relations complexes non linéaires ou non additives
- Tester/quantifier les associations entre un ou plusieurs prédicteurs et la réponse, et interpréter le modèle ajusté
- Validation et calibrage du modèle pour évaluer la précision de la prédiction et identifier les surajustements
- Apprendre les différences entre l'apprentissage automatique et les modèles statistiques, et comment choisir la meilleure approche pour un problème donné