



**École d'été interdisciplinaire en numérique de la santé  
du 26 au 30 mai 2025**

**Les grandes banques de données médicales et administratives du Québec  
pour la recherche en santé : un cas d'étude**

**Plan d'activité pédagogique**

---

Type de l'activité : ☒ Atelier ☐ Présentation

Objectif : Le présent document décrit le contenu scientifique et l'organisation de l'activité pédagogique « Les grandes banques de données médicales et administratives du Québec pour la recherche en santé : un cas d'étude » présentée lors de l'école d'été.

---

## **1 Introduction**

L'activité pédagogique « Les grandes banques de données médicales et administratives du Québec pour la recherche en santé : un cas d'étude » se déroulera dans le cadre de l'école d'été interdisciplinaire en numérique de la santé (EINS) et s'inscrit dans le thème « Analyse de données en santé ».

### **1.1 Renseignements sur le présentateur**

Prénom, nom : Yohann Chiu

Affiliation principale :

Professeur adjoint au Département de médecine de famille et de médecine d'urgence,  
Faculté de médecine et des sciences de la santé, Université de Sherbrooke

Courriel : [Yohann.Chiu@USherbrooke.ca](mailto:Yohann.Chiu@USherbrooke.ca)

Site Web : <https://www.usherbrooke.ca/recherche/fr/specialistes/details/yohann.chiu>

### **Biographie**

Yohann Chiu est professeur adjoint à la Faculté de médecine et des sciences de la santé ainsi que chercheur régulier au Centre de recherche du CHUS dans l'axe « Santé : populations, organisation, pratiques ». Biostatisticien de formation, il se spécialise dans l'analyse des grandes banques de données médicoadministratives tant à l'aide d'outils statistiques que d'apprentissage automatique (modèles de régression, de classification ou de survie). Ses applications portent principalement sur les enjeux de santé reliés aux maladies chroniques et à la santé mentale. Ses projets actuels incluent la surveillance de la polypharmacie et les trajectoires de santé du trouble du spectre de l'autisme.

## 2 Description

Cette section présente le contenu de l'activité et les principaux objectifs.

### 2.1 Contenu

Les grandes banques de données médicales et administratives du Québec sont une mine d'or de renseignements pour l'amélioration des soins de santé. Comme elles sont avant tout créées dans un objectif de gestion et d'administration des soins, elles peuvent parfois être inexploitées en recherche par manque de connaissances sur leur manipulation ou sur les méthodes d'analyses adéquates. Cette présentation portera sur les différentes banques de données qui existent au Québec, particulièrement celles de la Régie de l'assurance maladie du Québec (RAMQ). La présentation explorera différentes façons dont ces données peuvent être exploitées pour répondre à des questions de recherche menant à des analyses descriptives et à des modèles statistiques. L'utilisation fréquente des services d'urgence et la surveillance de la polypharmacie serviront d'applications, avec des méthodes abordées à la frontière de la biostatistique, de l'épidémiologie et de l'apprentissage automatique, sous l'angle commun de l'optimisation des soins de santé. Les applications se feront avec le logiciel R. Les avantages et les limites des grandes banques de données médicales et administratives seront aussi abordés.

### 2.2 Objectifs de formation

Cette activité permettra à une personne étudiante :

- O1. d'identifier les grandes banques de données médicales et administratives qui existent au Québec ;
- O2. de connaître les opportunités et les limites de ces banques pour la recherche en santé ;
- O3. d'appliquer des modèles d'apprentissage automatique de base pour la prédiction d'issues de santé.

## 3 Références

Cette section présente les principales références documentaires utilisées pour construire l'activité et les références pour approfondir des concepts présentés.

### 3.1 Références essentielles

Chiu, Y. M., Dufour, I., Courteau, J., Vanasse, A., Chouinard, M. C., Dubois, M. F., ... & Hudon, C. (2022). Profiles of frequent emergency department users with chronic conditions: a latent class analysis. *BMJ open*, 12(9), e055297.

Chiu, Y. M., Courteau, J., Dufour, I., Vanasse, A., & Hudon, C. (2023). Machine learning to improve frequent emergency department use prediction: a retrospective cohort study. *Scientific Reports*, 13(1), 1981.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer.

### 3.2 Références complémentaires

Blais, C., Jean, S., Sirois, C., Rochette, L., Plante, C., Larocque, I., ... & Émond, V. (2014). Quebec integrated chronic disease surveillance system (QICDSS), an innovative approach. *Chronic diseases and injuries in Canada*, 34(4).

Sirois, C., Khoury, R., Durand, A., Deziel, P. L., Bukhtiyarova, O., Chiu, Y. M., ... & Simard, M. (2021). Exploring polypharmacy with artificial intelligence: data analysis protocol. *BMC Medical Informatics and Decision Making*, 21(1), 1-8.

Sirois, C., Boiteau, V., Chiu, Y. M., Gilca, R., & Simard, M. (2022). Exploring the associations between polypharmacy and COVID-19-related hospitalisations and deaths: A population-based cohort study among older adults in Quebec, Canada. *BMJ open*, 12(3), e060295.