

## **Proposition de stage : analyse des multi-résistances chez *E. coli* à partir d'approches de fouille de données**

**Durée du stage** : 5 à 6 mois

**Laboratoire** : Laboratoire Modélisation, Epidémiologie et Surveillance des Risques Sanitaires (MESuRS), Conservatoire national des Arts et Métiers (Cnam)

**Encadrement** : Laura Temime (Cnam), en collaboration avec Gabriel Birgand et Olivier Lemenand

**Contexte** : Les bactéries multirésistantes aux antibiotiques, qui ne sont plus sensibles qu'à un nombre restreint d'antibiotiques, sont de plus en plus répandues, avec d'importantes conséquences potentielles en termes de morbidité et de mortalité. Mieux connaître les profils de multirésistance peut aider à limiter la sélection de ces bactéries et à optimiser la prise en charge des infections associées.

Les approches de fouille de données basées sur la recherche de règles d'association (« association rule mining ») peuvent être utilisées pour identifier les principaux profils de multirésistance ainsi que leur fréquence, comme l'a démontré une étude récente sur *S. aureus* s'appuyant sur des données internationales [1].

**Objectifs du stage** : Le stage visera à analyser des données issues d'un large système français de surveillance de l'antibiorésistance en ville et en Ehpad (<https://medqualville.antibioresistance.fr/>), afin de mieux comprendre les phénomènes de multirésistance chez *E. coli*.

Une première analyse consistera à décrire les résistances observées dans la base de données sur la période 2018-21, en fonction de l'âge des individus concernés, de leur origine géographique, du site de prélèvement etc.

Dans un deuxième temps, l'ensemble des profils de multirésistance présents dans les données seront identifiés à l'aide d'un outil d'*association mining* (algorithme Apriori). Différentes métriques seront calculées afin de décrire la fréquence de ces profils et d'évaluer la probabilité qu'ils reflètent de véritables résistances croisées (support, indice de confiance, lift).

Enfin, des « réseaux de résistance » seront reconstruits, permettant une visualisation graphique des profils de multirésistance les plus significatifs, tels qu'identifiés par des tests statistiques.

Dans la mesure du possible, toutes ces analyses seront réalisées par année d'étude et par région, afin d'être en mesure d'étudier les dynamiques temporelles et géographiques.

**Outils** : programmation en R

**Référence** :

[1] Cazer C., et al. Analysis of Multidrug Resistance in *Staphylococcus aureus* with a Machine Learning-Generated Antibigram. *Antimicrobial Agents and Chemotherapy*, 2021; 65(4): e02132-20.

**Informations pratiques** :

Le stage se déroulera au printemps 2023, dans les locaux du laboratoire MESuRS du Cnam, situés dans le 3<sup>ème</sup> arrondissement de Paris. Il fera l'objet d'une gratification.

Pour candidater, envoyer CV et lettre de motivation à Laura Temime ([laura.temime@lecnam.net](mailto:laura.temime@lecnam.net)).