

UE 1.3. Disciplinaire thématique

6 ECTS

## EC3 : Méthodes statistiques avancées et modélisation

Équipe pédagogique : Alain Duhamel, Michaël Génin, Mohamed Lemdani, Guillemette Marot, Claire Pinçon.

Contact : [alain.duhamel@univ-lille.fr](mailto:alain.duhamel@univ-lille.fr)

### Résumé

Cet enseignement constitutif (EC) apporte des connaissances approfondies sur les outils statistiques à utiliser en recherche médicale avec l'apprentissage d'un logiciel approprié (R) pour répondre aux besoins de ces analyses. Les outils abordés vont de l'analyse de survie à la régression (linéaire et logistique) en passant par l'ensemble des méthodes d'analyse de données (analyse en composantes principales, analyse discriminante, ...) qui préfigurent les outils de fouilles de données à la base des sciences de données (data science)

### Objectifs pédagogiques :

Le suivi de cet EC a pour objectif d'apporter une maîtrise théorique et pratique des procédures statistiques avancées nécessaires pour analyser efficacement les données biologiques et médicales. Les méthodes statistiques abordées d'analyser finement les données cliniques (analyse de survie, méthodes d'analyse des données, régression linéaire, régression logistique). La maîtrise de l'outil logiciel en lien avec l'emploi de ces méthodes fait également partie des objectifs attendus.

### Bloc de Compétences et de Connaissances-BCC 1 : Connaitre les concepts de base en Biologie Santé

#### Compétences acquises (directes/indirectes) :

Cet enseignement contribue à fournir les compétences pour contribuer à une activité de recherche en Biologie-Santé (BC1), en préparant les étudiants à :

- cerner les enjeux de la recherche en Biologie-Santé ;
- organiser une veille bibliographique de la littérature scientifique internationale ;
- fonder des hypothèses sur les concepts les plus récents en recherche Biologie-Santé.

#### Prérequis :

- Avoir suivi au 1er semestre l'UE1.1-EC4 : Méthodes Statistiques de Base.

#### Contenu :

L'EC comprend des enseignements magistraux, des travaux dirigés et des travaux pratiques sur ordinateur (avec apprentissage d'un logiciel statistique).

- Introduction du logiciel statistique (4 heures).
- Analyse de survie (10 heures).
- Rappels d'algèbre linéaire (4 heures).
- Régression linéaire multiple (8 heures).
- Analyse Factorielle Discriminante (6 heures).
- Régression logistique (4 heures).
- Méthodes de Classification (4 heures).
- Analyse en Composantes Principales (4 heures)