



MASTER MENTION SANTE

Parcours

« Epidémiologie, Données de Santé, Biostatistique »

M1 commun

M2 : 2 Sous parcours

M2 : « Epidémiologie, Données de Santé, Biostatistique »

M2 : « Data Analyst pour les Sciences du Vivant »

Université de Montpellier
2021-2025



Parcours

« Epidémiologie, Données de Santé, Biostatistique »

M1 commun

Responsables : Pascale Fabbro-Peray & Thibault Mura

UFR de Médecine Montpellier-Nîmes

Robert Sabatier & Christelle Reynes

UFR de Sciences Pharmaceutiques

M2 : Sous parcours « Données de Santé »

Tronc commun

1 bloc optionnel à choisir parmi 2 :

- **Méthodologies spécifiques**
- **Radiomics : l'imagerie quantitative et l'IA en radiologie**

Responsables : Pascale Fabbro-Peray & Thibault Mura

UFR de Médecine Montpellier-Nîmes

M2 : Sous parcours « Data Analyst pour les Sciences du Vivant »

Responsables : Robert Sabatier & Christelle Reynes

TABLE DES MATIERES

Présentation du parcours.....	6
Admissions.....	7
Programme	9
Master 1.....	9
Semestre 1	9
UE Mathématiques générales : R Sabatier (2.5 ECTS)	9
UE Introduction à la statistique inférentielle 1: V Ladret (2.5 ECTS).....	9
UE Introduction à la statistique inférentielle 2 : Jn Bacro (2.5 ECTS).....	10
UE Modèle linéaire général :Jn Bacro (2.5 ECTS).....	10
UE Data Mining : R. Sabatier (5 ECTS).....	11
UE Programmation R et SAS niveau 1: C Reynes-(2.5 ECTS).....	12
UE Introduction à la recherche clinique et épidémiologique : P. Fabbro-Peray (2.5 ECTS).....	12
UE Bases de données - Niveau 1 : M. Vivien (2.5 ECTS).....	13
UE Anglais : P Mac Nish (5 ECTS).....	13
UE Grands enjeux en Santé – P Demoly (2.5 ECTS).....	14
UE Aspects technologiques du recueil de données – omiques : C. Reynès (2.5 ECTS).15	
Semestre 2	16
UE Ingénierie de projets & Communication : M HAYOT & C REYNES (2.5 ECTS).....	16
UE Python : C. Reynès & R Sabatier(2.5 ECTS).....	16
UE Programmation R et SAS niveau 2: JN Bacro (2.5 ECTS).....	16
UE Méthodes en épidémiologie quantitative Niveau de base : C Duflos (2.5 ECTS).....	17
UE Big Data & Intelligence Artificielle en Santé : M Hayot (2.5 ECTS).....	18
UE Essais cliniques randomisés: JL. Faillie (2.5 ECTS)	18
UE Analyse statistique de données –omiques : C. Reynès(2.5 ECTS).....	19
UE Etude de cas : C Reynes (2.5 ECTS).....	19
UE Stage : UFR SCIENCES PHARMACEUTIQUES - C Reynes / UFR Médecine - C Serrand (10 ECTS).....	20
Master 2 – sous parcours « Données de Santé ».....	21
Semestre 3	21
UE Analyse des données censurées : T Mura (2.5 ECTS)	21
UE Modèle linéaire généralisé et modèle mixte : I. Carrière (5 ECTS).....	21

UE Machine learning niveau 1 – Application au pronostic : R Sabatier & T Mura (2.5 ECTS).....	22
UE Séminaires de recherche – P Fabbro-Peray (2.5 ECTS).....	23
UE Méthodes en épidémiologie quantitative niveau avancé : C Berr (5 ECTS).....	23
UE Méta-analyses P Fabbro-Peray (2.5 ECTS).....	24
Bloc optionnel 1 : UE Méthodologies spécifiques	25
UE Médico-économie : S Chkair – S Bouvet (2.5 ECTS).....	25
UE Qualité de Vie : C. Gaujoux-Viala (2.5 ECTS)	26
UE Méthodes en Recherche Qualitative : A Oude-Engberink (2.5 ECTS).....	27
UE Machine learning niveau 2 –: R Sabatier (2.5 ECTS).....	28
Bloc optionnel 2 : UE Radiomics : l'imagerie quantitative et l'IA en radiologie.....	29
UE Constitution du signal morpho-physiologique en imagerie médicale : F Pereira (2.5 ECTS).....	29
UE Traitement numérique de l'image médicale : F Pereira (2.5 ECTS).....	30
UE Méthodes en Imagerie quantitative neurologique/oncologique : JP. Bérégi – N Menjot de Champfleury (2.5 ECTS).....	31
UE Méthodes en Imagerie quantitative cardiovasculaire & en radiologie interventionnelle: JP. Bérégi – H Kovacsik (2.5 ECTS).....	32
Semestre 4	33
UE Stage : T Mura (30 ECTS).....	33
Master 2 sous parcours « Data Analyst pour les Sciences du Vivant ».....	34
Semestre 3	34
UE Analyse des données censurées : T Mura (2.5 ECTS)	34
UE Modèle linéaire généralisé et modèle mixte : I. Carrière (5 ECTS).....	34
UE Machine learning niveau 1 – Application au pronostic : R Sabatier & T Mura (2.5 ECTS).....	35
UE Machine learning niveau 2 –: R Sabatier (2.5 ECTS).....	36
UE Séminaires de recherche – P Fabbro-Peray (2.5 ECTS).....	37
UE Base de données – Niveau 2 : M. Vivien (2.5 ECTS).....	37
UE Recueil planifié des données : E Goze (2.5 ECTS)	38
UE Statistique temporelle : JN Bacro (2.5 ECTS)	38
UE Statistique pour l'industrie : ? (2.5 ECTS)	39
UE Etude de cas niveau 2 : R Sabatier (2.5 ECTS)	39
Semestre 4	40
FORMATION INITIALE.....	40
UE Stage : R Sabatier	40

FORMATION EN ALTERNANCE.....	41
UE Alternance.....	41

VERSION DE TRAVAIL

PRESENTATION DU PARCOURS

Le parcours proposé est une **évolution des parcours « Statistiques pour les Sciences de la Vie » & « Statistiques pour les Sciences de la Santé »** de la mention « Mathématiques » de la précédente période. **Le Master 1 devient commun et nous proposons deux sous parcours de M2 : « épidémiologie, données de santé, biostatistiques » et « Data Analyst pour les Sciences du Vivant ».**

Son objectif est de fournir à des étudiants issus principalement du cursus Santé et de licences de biologie, de mathématiques, de Sciences de la Santé et de MIAHS, **d'acquérir une double-compétence en épidémiologie /biostatistiques.** Cette double compétence est particulièrement recherchée sur le marché de l'emploi. Nos étudiants seront de véritables atouts dans une équipe puisqu'ils auront la culture nécessaire en biologie/santé pour maîtriser la problématique d'intérêt et la compétence pour analyser les données de façon adéquate. Cette analyse adéquate des données en biologie/santé est un enjeu majeur pour la recherche des années à venir car les données sont de plus en plus volumineuses et nombreuses et des erreurs dans leur analyse peut conduire (et a déjà conduit par le passé) à des conclusions erronées ou non reproductibles décrédibilisant l'ensemble de la filière recherche. Une expertise réelle en analyse de données est donc aujourd'hui indispensable pour permettre de répondre à des questions épidémiologiques & biologiques complexes.

De plus, nous avons fait évoluer le contenu de la formation pour permettre aux étudiants en santé et en biologie d'acquérir des compétences toujours plus proches du marché de l'emploi en analyse de données : introduction du langage Python, renforcement des enseignements en Machine Learning et intelligence artificielle. Selon le sous parcours M2 et les options choisies, les étudiants pourront acquérir des compétences dans un domaine d'intérêt privilégié: méthodes biostatistiques spécifiques dans le sous parcours « data analyst pour les sciences du vivant », méthodes d'évaluation et d'analyse des données subjectives (Qualité de vie et recherche qualitative) dans le bloc optionnel 1 du sous parcours « épidémiologie, données de santé, biostatistiques » , Méthodes en imagerie quantitative dans le bloc optionnel 2 du sous parcours « épidémiologie, données de santé, biostatistiques ».

Cette évolution est également en cohérence avec l'intégration dans la mention Santé car les applications en santé de toutes ces méthodes sont de plus en plus nombreuses.

En termes d'évolution, il correspond également à un besoin concernant la partie du public cible constituée d'étudiants en Santé et d'étudiants en réorientation issus du Parcours d'Accès Spécifique Santé (PASS) et de la Licence Accès Santé(LAS), en cours de mise en place dans le cadre de la réforme de la PACES.

ADMISSIONS

Cursus Master complet

Niveau de recrutement : Bac + 3

Durée de formation: 2 années d'étude (M1, M2) composées chacune de deux semestres, validant chacun 30 ECTS. Chaque année est composée d'UE d'enseignements et d'1 UE de stage.

Lieu de formation: Université de Montpellier

Accessible en : Formation Initiale ou Formation continue

Conditions d'admission: Les candidatures sont à renseigner sur la plateforme e-candidat de l'Université de Montpellier. Une commission pédagogique examine les candidatures sur dossiers en tenant compte de l'adéquation du profil du candidat, de son parcours et de sa motivation, avec les exigences et la nature de la formation.

Pour les étudiants étrangers, le niveau C1 de français est requis.

Poursuite d'études possible : Doctorat d'Université

Débouchés professionnels :

Biostatisticien en recherche dans une structure publique ou privée, Ingénieur de recherche en biostatistiques, Ingénieur d'étude en biostatistiques, Chercheur dans l'enseignement supérieur, Chercheur dans les organismes de recherche ou en Santé Publique, Chargé d'étude ou chargé de mission dans les institutions & agences sanitaires et sociales, Médecins et praticiens de Santé Publique.

Cursus spécial étudiants en Santé (Médecine, Pharmacie, Odontologie, Maïeutique)

Les étudiants des filières Santé bénéficient d'un crédit de 10 ECTS par leur cursus santé.

Pour valider le master 1 EDSB et accéder à un des sous parcours M2 EDSB, ils doivent valider les 4 UE de 2,5 ECTS suivantes du cursus M1 et réaliser un stage de 10 ECTS.

- UE « Statistiques inférentielles niveau 1 »
- UE « Modèle linéaire général »
- UE « R-SAS niveau 1 »
- UE « Méthodes en épidémiologie quantitative de base »

Niveau de recrutement

- Etudiants en Médecine de DFGSM3, DFASM1 à 3, internes en Médecine
- Etudiants en Pharmacie de
- Etudiants en Odontologie de
- Etudiants en Maïeutique de DFGSMa3, DFASMa1 & 2

2 UE par an pour les étudiants en 2^{ème} cycle, 4UE en 1 an pour les internes.

Conditions d'admission: Les inscriptions seront réalisées selon un processus spécifique géré par la faculté de Médecine ou de Pharmacie, sans passer par la plateforme e-candidat, sur dossier constitué d'un CV et d'une lettre de motivation.

Il n'y a pas d'équivalence de crédits pour le M2, quelle que soit la formation initiale. En Master 2, les étudiants doivent valider 30 crédits d'enseignement au premier semestre et 30 crédits de stage au second semestre. La préinscription se fait via la plateforme e-candidat.

VERSION DE TRAVAIL

PROGRAMME

MASTER 1

SEMESTRE 1

UE MATHÉMATIQUES GÉNÉRALES : R SABATIER (2.5 ECTS)

- **Description :**

L'UE « Mathématiques générales » est une UE de remise à niveau et d'homogénéisation du bagage mathématique nécessaire à la bonne compréhension des méthodes statistiques présentées dans les différentes UE du master EDSB.

Le cours se focalise sur les fondamentaux suivants :

- Algèbre linéaire : matrices et calcul matriciel ; matrices inversibles et singulières ; valeurs propres et vecteurs propres d'une matrice ; diagonalisation d'une matrice ; matrice inverse généralisée ; résolution de système linéaire d'équations.
- Analyse : dérivée de fonctions réelles et dérivées partielles de fonctions multivariées ; recherche d'extrema de fonctions différentiables ; développements limités de fonctions différentiables.

- **Objectifs :**

- maîtriser les aspects mathématiques nécessaires pour d'autres UEs du master
- comprendre les implications pratiques de ces méthodes
- s'initier au maniement de logiciels de biostatistiques (R)

- **Pré-requis nécessaires :**

- outils mathématiques, niveau Terminale scientifique

- **Responsable de l'UE : Robert SABATIER**

- **Intervenants:**

Robert SABATIER, UFR des Sciences Pharmaceutiques

UE INTRODUCTION A LA STATISTIQUE INFÉRENTIELLE 1: V LADRET (2.5 ECTS)

- **Description**

La formation présente une introduction aux aspects théoriques et pratiques des analyses statistiques des données de santé.

Les méthodes statistiques seront présentées et des applications seront systématiquement faites à l'aide du logiciel R, via l'interface graphique Rcommander.

Seront notamment abordés : les Variables aléatoires, l'estimation et les intervalles de confiance les tests statistiques paramétriques et non paramétriques

Chaque méthode sera présentée de manière théorique et une place aux applications sera faite pour chacune. Afin de faciliter l'assimilation des connaissances, cette UE alterne cours théoriques et applications pratiques / cas concrets / travaux sur ordinateur.

- **Objectifs**

Donner les outils théoriques et apprendre à utiliser un logiciel de premier niveau utile pour l'analyse statistique des données de santé (médecine, pharmacie, ...).

- **Pré-requis nécessaires :**

Aucune notion de statistique n'est nécessaire. Les notions de base seront reprises lors

- **Responsable de l'UE : Véronique LADRET**

- **Intervenants:**

Véronique LADRET, UM, UFR Médecine

Nicolas MOLINARI, UM, UFR Sciences Pharmaceutiques

UE INTRODUCTION A LA STATISTIQUE INFÉRENTIELLE 2 : JN BACRO (2.5 ECTS)

- **Description :**

Le cours approfondit les premières notions d'estimation et de tests statistiques vues au niveau 1 du cours de statistique inférentielle.

Pour l'estimation, l'accent sera mis sur l'estimation par maximum de vraisemblance et les propriétés statistiques qui s'y rattachent. Les 3 tests fondamentaux de la statistique, à savoir test du score, test de Wald et test du rapport de vraisemblance seront détaillés.

Les principes généraux des tests statistiques et la démarche statistique à suivre seront rappelés et approfondis (puissance d'un test statistique) et les contextes à deux populations ou plus seront abordés.

L'interface R-Studio du logiciel statistique R sera présentée et utilisée pour les travaux dirigés.

- **Objectifs :**

- Comprendre de manière approfondie le principe général des tests statistiques.
- Comprendre les approches d'inférence statistique fondées sur le maximum de vraisemblance.
- Être en mesure de mettre en œuvre correctement le(s) test(s) statistique(s) adapté(s) à la question traitée.
- Être capable d'argumenter une décision prise quant au rejet ou non de l'hypothèse nulle et d'en apprécier les conséquences.

- **Pré-requis nécessaires :**

Cours de statistique inférentielle niveau 1

- **Pré-requis recommandés :**

Notions de variables aléatoires et de lois de probabilité.

- **Responsable de l'UE : Jean-Noel Bacro**

- **Intervenants**

Jean-Noel Bacro, Faculté des Sciences

UE MODELE LINEAIRE GENERAL : JN BACRO (2.5 ECTS)

- **Description :**

Il s'agit de donner un aperçu du modèle linéaire général en l'introduisant avec son écriture matricielle qui permet une illustration, au travers de quelques exemples simples, des notions générales de régression, simple ou multiple, d'analyse de variance et des modèles combinant variables continues et variables discrètes. Le modèle linéaire général permet d'avoir une présentation (et une écriture) unifiée de ces différents cas.

A partir de cette écriture les étapes essentielles sont présentées, estimation, intervalles de confiance, tests d'hypothèses multiples ...Une place particulière est accordée à l'analyse de variance « classique » avec les possibilités de « séparer » les effets de plusieurs sources de variation et avec une introduction aux modèles comportant des effets aléatoires.

Toutes ces étapes sont illustrées par des exemples spécifiques, permettant par ailleurs de bien identifier les informations fournies par les logiciels statistiques « usuels » et de clairement identifier les modèles utilisés (et leurs écritures). Certaines extensions sont évoquées (régression lasso, modèles non linéaires, non paramétriques).

- **Objectifs :**
 - Connaître les conditions d'utilisation des différentes formes du modèle linéaire général
 - Savoir mettre en œuvre ces méthodologies (logiciels R et SAS)
 - Avoir un regard critique sur les résultats
- **Pré-requis nécessaires :**
 - outils vus dans l'UE Mathématiques Générales et Statistiques Inférentielles 1
- **Responsable de l'UE : Jean-Noël BACRO**
- **Intervenants**
Jean-Noël BACRO, Faculté des Sciences

UE DATA MINING : R. SABATIER (5 ECTS)

- **Description :**

L'objectif de ce premier enseignement d'Analyse de Données est l'acquisition des méthodes multivariées les plus classiques pour la fouille de données, l'apprentissage supervisé et non supervisé. L'enseignement sera centré sur les contraintes d'utilisation et l'interprétation des résultats afin de rendre les étudiants autonomes dans le choix et l'utilisation de ces méthodes qui constituent des outils élémentaires et indispensables dans l'analyse de tout fichier de données multivarié. Les notions théoriques fondamentales (s'appuyant sur les enseignements de l'UE Maths Générales) seront données aux étudiants afin qu'ils saisissent le fonctionnement global des méthodes afin d'optimiser leur maîtrise des outils mais sans volonté d'entrer dans un détail qui n'est pas nécessaire à leur bonne compréhension. Ainsi, l'enseignement sera en permanence guidé par des exemples concrets d'application appartenant à divers domaines d'application.

Chacune des méthodes enseignées en cours sera appliquée grâce au logiciel R au cours de séances de TDs. Ces séances permettront l'approfondissement de la compréhension des différents outils proposés et fourniront également aux étudiants des scripts types qu'ils pourront réutiliser par la suite.

- **Objectifs :**
 - savoir choisir une méthode d'analyse multivariée adaptée aux problématiques classiques
 - savoir mettre en œuvre ces méthodes dans le logiciel R
 - mettre en œuvre les bonnes pratiques de programmation
 - être autonome dans la manipulation et l'interprétation des résultats

- **Pré-requis nécessaires :**
 - outils vus dans l'UE Mathématiques Générales et R/SAS1
- **Responsable de l'UE : Robert SABATIER**
- **Intervenants:**
 - Christelle REYNES, UFR Sciences pharmaceutiques
 - Robert SABATIER, UFR Sciences pharmaceutiques

UE PROGRAMMATION R ET SAS NIVEAU 1: C REYNES-(2.5 ECTS)

- **Description :**

L'UE a pour but de familiariser les étudiants à la bonne utilisation des logiciels de statistique R (logiciel libre) et SAS. Dans une première partie du cours, les aspects théoriques des deux langages seront acquis par l'intermédiaire d'un MOOC. La seconde partie de l'UE consistera à mettre en pratique, lors de séances en présentiel, les notions vues pendant les MOOC.
- **Objectifs :**
 - s'initier à deux langages de programmation très utilisés en biostatistique
 - connaître les spécificités d'usage de chaque langage
 - apprendre les bonnes pratiques de programmation
 - acquérir les réflexes pour gagner en autonomie sur des fonctions nouvelles
 - mise en pratique des méthodes vues dans d'autres UEs
- **Pré-requis nécessaires :**
 - aucun
- **Responsable de l'UE : Christelle REYNES**
- **Intervenants:**
 - Christelle REYNES, UFR Sciences pharmaceutiques

UE INTRODUCTION A LA RECHERCHE CLINIQUE ET EPIDEMIOLOGIQUE : P. FABBRO-PERAY (2.5 ECTS)

- **Description :**

Les objectifs pédagogiques concernent l'appropriation des grands concepts qui sont déclinés dans le domaine de la recherche clinique et épidémiologique. Ces notions seront approfondies en M1 dans l'UE « Méthodes en épidémiologie niveau de base » et en M2 dans l'UE « Méthodes en épidémiologie quantitative niveau avancé ».

Présentation :

 - des principaux champs de l'épidémiologie et de la recherche clinique,
 - des grands types d'enquêtes observationnelles et expérimentales ;
 - des points clefs méthodologiques des essais cliniques randomisés (randomisation, aveugle, intention de traiter) ;
 - des concepts d'association statistique et de causalité,
 - des différentes mesures de fréquence d'une maladie ou d'un phénomène de santé (prévalence, taux d'incidence, risque) ;
 - des mesures d'association entre une maladie et un facteur de risque (différence de risque, risque relatif, odds ratio) et d'impact (fraction de risque attribuable),
 - des paramètres d'évaluation de la valeur diagnostique d'un test (sensibilité, spécificité, valeurs prédictives)
 - de la notion de confusion

- de la notion de biais : principaux types de biais, (biais de sélection, biais de classification et biais de confusion)
- **Objectifs :**
Les étudiants devront connaître, à l'issue de cette UE, les définitions de tous les concepts évoqués, être capables d'associer les différents champs de l'épidémiologie aux types d'études et aux indicateurs appropriés, de reconnaître les biais potentiels en fonction du type d'étude.
- **Pré-requis nécessaires :**
Aucun
- **Responsable de l'UE : Pascale Fabbro-Peray**
- **Intervenants:**
Claudine Berr, INSERM
Pascale Fabbro-Peray, UFR Médecine
Isabelle Jaussent, INSERM
Joanna Norton, INSERM
Fabienne Séguret, CHU Montpellier

UE BASES DE DONNEES - NIVEAU 1 : M. VIVIEN (2.5 ECTS)

- **Description :**
 - Le data-management peut être vu comme l'art de ranger ou de stocker des données non forcément liées en vue d'en assurer un traitement ultérieur efficace. Il existe des modèles normalisés pour stocker ces données (méthode merise, modèle relationnel) et un langage (langage SQL) pour traiter et gérer ces données. La cohérence des données est assurée en empêchant la saisie de « n'importe quoi » et celle-ci permet une exploitation rationnelle de ces dernières (typiquement traitements statistiques).
 - Pour ce cours de première année, le logiciel Access sera utilisé pour s'initier au data-management : méthode Merise, langage SQL, import/ export de données. Des compléments en informatique seront aussi présentés : processeur, un disque dur : à quoi ça sert, comment ça marche ; compression, cryptage et réseau ; introduction à la conception de site internet : centralisation des données pour gérer des recueils décentralisés de données.
- **Objectifs :**
 - connaître et savoir utiliser les modèles normalisés d'organisation des données
 - savoir implémenter et utiliser ces modèles sous Access
- **Pré-requis nécessaires :**
aucun
- **Responsable de l'UE : Myrtille VIVIEN**
- **Intervenants:**
Myrtille VIVIEN, UFR Sciences Pharmaceutiques

UE ANGLAIS : P MAC NISH (5 ECTS)

- **Description* :**
Cours-TD d'anglais de spécialité et anglais de communication qui vise l'autonomie professionnelle en langue anglaise.

20 heures en face à face et 20 heures de travail personnel individuel sur moodle

- Compréhension orale avec supports vidéo et échanges en groupe
- Compréhension écrite à partir d'articles de la presse scientifique
- Expression orale en interaction – travaux en groupes, jeux de rôle
- Expression orale en présentations individuelles et / ou en équipes
- Expression écrite – en lien avec le monde de l'entreprise

- **Objectifs**

Les étudiants travailleront les cinq compétences langagières décrites par le Cadre Européen Commun de Référence pour les Langues (CECRL) afin d'accéder à une aisance orale et écrite compatible avec le monde du travail et avec des interlocuteurs anglophones.

- **Prérequis nécessaires**

Notions de grammaire et des compétences d'expression écrite et orale élémentaires (A2)

- **Prérequis recommandés**

Le niveau B2 du CECRL à l'oral comme à l'écrit.

- **Responsable de l'UE : Pauline Mc NISH**

- **Intervenants**

Pauline Mc Nish, UM, Faculté des Sciences

UE à CHOIX (1 parmi les 2)

UE GRANDS ENJEUX EN SANTE – P DEMOLY (2.5 ECTS)

- **Description :**

Nous vivons tous plus longtemps, mais vivre en bonne santé est le souhait de tous. A partir du catalogue des maladies les plus fréquentes, seront abordées les maladies les plus coûteuses, celles émergentes, celle liées au vieillissement, celles liées au changement climatique, celles évitables, donc le rôle de l'environnement et des facteurs nutritionnels, de l'organisation des soins, de la prise en charge des groupes à risque.

- **Objectifs :**

L'étudiant devra connaître, à l'issue de cette UE, les grands indicateurs de santé dans les pays en fonction du niveau de revenu avec un focus particulier sur la France. Il connaîtra les enjeux liés à l'organisation des soins et les actions de soins et de prévention menées ou à mener pour répondre aux besoins. Il aura abordé les enjeux de santé liés à notre environnement. Il saura les situer dans le contexte du développement durable de l'ONU.

- **Pré-requis nécessaires :**

- L'étudiant doit connaître les définitions des grands indicateurs de santé (morbidity, mortalité, espérance de vie).
- Il devra avoir lu le rapport le plus récent de l'Organisation Mondiale de la Santé concernant les *Statistiques sanitaires mondiales*.
- Il devra avoir des notions d'organisation du système de santé français.

- **Pré-requis recommandés :**

aucun

- **Responsable de l'UE : Pascal DEMOLY**

- **Intervenants :**

ANNESI-MAESANO Isabella, INSERM
AVIGNON Antoine, UM, UFR Médecine
CAIMMI Davide, CHU de Montpellier
DEMOLY Pascal, UM, UFR Médecine
NINOT Gregory, UM, UFR Staps
KASE-TANNO Luciana, CHU de Montpellier

UE ASPECTS TECHNOLOGIQUES DU RECUEIL DE DONNEES – OMiques : C. REYNES (2.5 ECTS)

- **Description :**

De par les cursus antérieurs des étudiants ciblés par cette formation, la notion des données –omiques (génomique, protéomique, transcriptomique, métabolomique,...) ne devrait pas leur être totalement étrangère. L'objectif de ce premier module autour des données –omiques est d'approfondir leur connaissance de la structure de ces données en apprenant leurs modes de production de la paillasse jusqu'au bureau du biostatisticien. Cette structure est le résultat des technologies, la plupart du temps très innovantes, qui conduisent au recueil de ce type de données. Les aspects fondamentaux des technologies à l'œuvre seront donc enseignés uniquement dans le but de permettre une bonne maîtrise des caractéristiques intrinsèques des données obtenues. Les notions de grande dimension, de discrétisation, de normalisation,... seront donc abordées. Les premières étapes de pré-traitement des données devront être acquises par les étudiants.

- **Objectifs :**

- connaître les caractéristiques principales des données de transcriptomique et de protéomique
- connaître l'importance des pré-traitements pour ces données
- savoir mettre en œuvre les pré-traitements de ces données dans le logiciel R

- **Pré-requis nécessaires :**

Aucun

- **Pré-requis recommandés :**

notions élémentaires de biologie moléculaire

- **Responsable de l'UE : Christelle REYNES**

- **Intervenants :**

Christelle REYNES, UFR Sciences pharmaceutiques
Jacques COLINGE, UFR Médecine
Stéphanie RIALLE

SEMESTRE 2

UE INGENIERIE DE PROJETS & COMMUNICATION : M HAYOT & C REYNES (2.5 ECTS)

Initiation à la publication scientifique et à la communication scientifique. Bases de données bibliographiques. Modalités de réalisation d'un recueil d'articles. Règles d'interrogation. Initiation à Zotero.

UE PYTHON : C. REYNES & R SABATIER(2.5 ECTS)

- **Description :**
Le logiciel Python devient l'un des standards pour le traitement de données biologiques en parallèle du logiciel R. Cette UE a pour objectif d'initier les étudiants à ce langage de programmation au travers du suivi d'un MOOC et de séances de mise en pratique.
- **Objectifs :**
 - connaître les bases du langage Python
 - mettre en œuvre les bonnes pratiques de programmation
 - savoir mettre en œuvre sous Python les méthodes vues dans l'UE Data Mining
- **Pré-requis nécessaires :**
 - bonnes pratiques de programmation vues au S1
 - conditions d'utilisation des méthodes vues dans l'UE Data Mining
- **Responsables de l'UE : Christelle REYNES & Robert SABATIER**
- **Intervenants :**
Christelle REYNES, UFR Sciences pharmaceutiques
Robert SABATIER, UFR Sciences pharmaceutiques

UE PROGRAMMATION R ET SAS NIVEAU 2: JN BACRO (2.5 ECTS)

- **Description :**
L'UE R & SAS – Niveau 2 fait suite à l'UE « **Programmation R et SAS niveau 1** » du S1 et a pour objectif d'approfondir la maîtrise des logiciels R et SAS.
- **Objectifs :**
 - être plus autonomes dans l'automatisation des procédures sous R et SAS
- **Pré-requis nécessaires :**
bonnes pratiques de programmation vues au S1
acquis de l'UE R/SAS niveau 1
- **Responsable de l'UE : Jean-Noël BACRO**
- **Intervenants :**
Jean-Noël BACRO, Faculté des Sciences
Christelle REYNES, UFR Sciences pharmaceutiques

UE METHODES EN EPIDEMIOLOGIE QUANTITATIVE NIVEAU DE BASE : C DUFLOS (2.5 ECTS)

- **Description :**

Cette UE participe à l'apprentissage de l'épidémiologie observationnelle dans le Master EDSB. Elle fait directement suite à l'UE « Introduction à l'épidémiologie et à la recherche clinique » du S1, en approfondissant la méthodologie générale des études observationnelles ; les connaissances abordées feront également appel aux connaissances acquises en statistiques inférentielles lors du S1.

Cette UE s'attachera d'abord à fournir, pour les principaux éléments de méthodologie en recherche observationnelle, une définition approfondie des méthodes et des critères permettant de s'assurer de la qualité de la recherche au moment de la formulation d'un protocole de recherche. Cette UE s'attachera ensuite à donner les principes de rédaction et de lecture d'un protocole de recherche et d'un article original en recherche observationnelle. Les connaissances acquises dans cette UE seront appliquées à des situations présentant des sources de complexités particulières dans l'UE « Méthodes en épidémiologie quantitative Niveau avancé » du S4.

- **Objectifs :**

Être capable de mobiliser les connaissances acquises sur ...

- Les prérequis de cette UE
- La planification des enquêtes cas-témoins et des cohortes
- Les analyses de survie
- Les méthodes de sondage
- La mesure de l'exposition par questionnaires
- Les phénomènes de confusion et d'interaction, et leur exploration par les analyses stratifiées et multivariées
- La standardisation des indicateurs de santé
- Les enquêtes d'intervention primaire et secondaire
- La structure d'un protocole de recherche et d'un article selon STROBE

... pour construire un protocole de recherche et critiquer un article original en recherche observationnelle.

- **Pré-requis nécessaires :**

Connaissance :

- des principaux champs de l'épidémiologie et de la recherche clinique,
- des grands types d'enquêtes observationnelles et expérimentales ;
- des concepts d'association statistique et de causalité,
- des différentes mesures de fréquence d'une maladie ou d'un phénomène de santé (prévalence, taux d'incidence, risque) ;
- des mesures d'association entre une maladie et un facteur de risque (différence de risque, risque relatif, odds ratio) et d'impact (fraction de risque attribuable),
- de la notion de confusion
- de la notion de biais : principaux types de biais, (biais de sélection, biais de classification et biais de confusio)
- de la notion d'aveugle
- de la définition et de l'interprétation d'un intervalle de confiance
- de la théorie des tests statistiques

- **Pré-requis recommandés :**

Connaissance :

- des points clefs méthodologiques des essais cliniques randomisés (randomisation, intention de traiter) ;

- des paramètres d'évaluation de la valeur diagnostique d'un test (sensibilité, spécificité, valeurs prédictives)

Ces pré-requis sont recommandés car il peut y être fait référence à titre d'exemple ou de mise en perspective des notions vues dans cette UE.

- **Responsable de l'UE : Claire Duflos**

- **Intervenants :**

Claudine Berr, INSERM
 Claire Duflos, UM, UFR médecine
 Pascale Fabbro-Peray, UM, UFR médecine
 Isabelle Jaussent, INSERM
 Thibault Mura, UM, UFR médecine
 Joanna Norton, INSERM
 Fabienne Séguret, UM, UFR médecine

UE BIG DATA & INTELLIGENCE ARTIFICIELLE EN SANTE : M HAYOT (2.5 ECTS)

UE ESSAIS CLINIQUES RANDOMISES: JL. FAILLIE (2.5 ECTS)

- **Description :**

Cette UE aborde les objectifs et principes généraux méthodologiques, réglementaires et éthiques des essais cliniques

Notions abordées :

- Principaux aspects éthiques et réglementaires
- Notions générales sur les essais de phases 1 et 2
- Méthodologie de l'essai de phase 3 :
 - Randomisation, aveugle, intention de traiter
 - Critères de jugement, placebo, NSN
 - Expression des résultats ; analyses intermédiaires, analyses de sous-groupes
 - Design particuliers : cross-over, plan factoriel 2x2, essais de non infériorité
 - Validité, biais
- Essais non médicamenteux, essais en cluster, méthodes adaptatives
- Protocole d'essai clinique, réalisation d'un essai en pratique, SPIRIT, eCRF
- Lecture critique d'article sur un essai clinique LCA - CONSORT

- **Objectifs :**

- Les étudiants devront connaître, à l'issue de cette UE, les définitions de tous les concepts évoqués, être capables de juger la méthodologie d'un essai clinique, connaître les étapes de réalisation d'un essai clinique en pratique.

- **Pré-requis nécessaires : Aucun**

- **Pré-requis recommandés : Aucun**

- **Responsable de l'UE : Jean-Luc FAILLIE**

- **Intervenants :**

Jean-Luc FAILLIE UM, UFR Médecine

UE ANALYSE STATISTIQUE DE DONNEES –OMIQUES : C. REYNES(2.5 ECTS)

- **Description :**
Cette UE a pour objectif de prendre les données telles qu'obtenues à la fin de l'UE « Aspects technologiques du recueil de données –omics » et de mener l'analyse jusqu'à répondre à la (ou les) question(s) applicative(s) initiale(s). Les problématiques abordées s'articuleront essentiellement autour de la détection d'entités différentiellement exprimées entre plusieurs conditions. Cette question générale peut se décliner en différentes approches auxquelles on doit répondre par des méthodes statistiques spécifiques.
- **Objectifs :**
 - connaître et savoir implémenter les méthodes les plus classiques du traitement de données transcriptomiques et protéomiques
- **Pré-requis nécessaires :**
acquis de l'UE Aspects Technologiques du Recueil de Données -omiques
- **Responsable de l'UE : Christelle REYNES**
- **Intervenants:**
Christelle REYNES, UFR Sciences Pharmaceutiques
Jacques COLINGE, UFR Médecine

UE ETUDE DE CAS : C REYNES (2.5 ECTS)

- **Description :**
L'étude de cas se place dans le cadre de l'apprentissage par projet et a pour objectif de mettre l'étudiant dans une situation professionnalisante lui permettant de suivre un projet d'analyse de données de la planification de l'expérience au rendu des résultats interprétés.
La première partie de l'étude de cas a deux objectifs : faire tout d'abord émerger, sur une thématique fournie, une question précise et ensuite, en fonction de la question, avoir prévu le plan de recueil des données en anticipant le futur traitement statistique et éventuellement le recueil lui-même.
- **Objectifs :**
 - s'approprier une problématique biologique
 - apprendre à formuler précisément une ou plusieurs questions biologiques nécessitant un traitement statistique
 - savoir planifier le recueil de données
 - savoir anticiper le plan d'analyse statistique
 - savoir communiquer dans un contexte pluri-disciplinaire
- **Pré-requis nécessaires :**
acquis des UEs du M1
- **Responsable de l'UE : Christelle REYNES**
- **Intervenants :**
Christelle REYNES, UFR Sciences Pharmaceutiques
Robert SABATIER, UFR Sciences Pharmaceutiques

**UE STAGE : UFR SCIENCES PHARMACEUTIQUES - C REYNES / UFR MEDECINE - C SERRAND
(10 ECTS)**

- **Description :**

Le stage de Master 1 a pour objectif de s'immerger dans une structure qui utilise les statistiques pour l'épidémiologie ou la biologie (public/privé) pendant une durée minimale de 8 semaines. Au cours de ce stage l'étudiant doit prendre en main une problématique épidémiologique ou biologique, réaliser une étude bibliographique autour du sujet, proposer et mettre en œuvre un plan d'analyse statistique pour répondre à la question de départ, dans le respect des pratiques sur le lieu de stage.

- **Objectifs :**

- s'approprier une problématique épidémiologique ou biologique
- s'initier au travail et à la communication dans une équipe pluri-disciplinaire
- savoir réaliser et communiquer les résultats d'une étude bibliographique
- savoir argumenter le choix d'une méthode d'analyse
- savoir mettre en œuvre un plan d'analyse préalablement validé
- savoir présenter et interpréter les résultats d'une étude statistique

- **Pré-requis nécessaires :**

acquis des UEs du M1

- **Responsables de l'UE : Christelle REYNES & Chris SERRAND**

SEMESTRE 3

UE ANALYSE DES DONNEES CENSUREES : T MURA (2.5 ECTS)

- **Description**

L'analyse des données censurées a pris une place majeure en santé publique et en médecine. L'objectif de ce cours est de présenter les notions de base sur l'analyse de données censurées (analyse de survie) et de détailler les modèles les plus communément utilisés pour leur analyse. Les notions de censure à droite, à gauche et par intervalle seront présentées, puis les notions de censure non informative et de censure informative. L'analyse de la survie proprement dite sera abordée, à l'aide de méthode non paramétrique (estimateur de Nelson Aalen, de Kaplan Meier et actuarielle) et survie paramétrique (modèle exponentiel, Weibull, Weibull Généralisé), survie semi paramétrique (modèle de Cox) et comparaison de survies. L'introduction de covariables dans les cas paramétriques et semi paramétrique, en particulier le modèle de Cox, sera ensuite abordée. Les cas particuliers des modèles à entrée retardée, l'utilisation de variables dépendantes du temps et les modèles de survie à risques compétitifs seront également traités. Le lien entre l'approche de survie avec le processus Markovien ou non Markovien sera introduit. Le calcul du nombre de sujets à inclure dans le cas de la comparaison de courbes de survie sera également présenté.

- **Objectifs**

Les étudiants devront être capable, à l'issue de cette UE de manipuler les principales méthodes d'analyses de données censurées, en les appliquant à des données réelles :

- en choisissant le modèle le plus adapté aux données et à la question posée,
- en calculant à la main les estimateurs de bases (Nelson Aalen, de Kaplan Meier et actuarielle)
- en calculant à la main un test du log-Rank
- en programmant les analyses basées sur modèles non paramétrique, paramétriques et semi paramétriques sur les logiciels R ou SAS,
- en étant capable de vérifier les conditions de validité des modèles
- et étant capable d'interpréter les résultats qui en sont issus

- **Pré-requis nécessaires :**

Notion de base (import de données, gestion et manipulation de données, statistiques de base) sur les logiciel SAS et R.

- **Responsable de l'UE : Thibault MURA-TODESCO**

- **Intervenants:**

Thibault MURA-TODESCO, UFR Médecine
Laure-Anne Gutierrez, INSERM

UE MODELE LINEAIRE GENERALISE ET MODELE MIXTE : I. CARRIERE (5 ECTS)

- **Description :**

L'objectif du cours est d'étendre les acquis du cours modèle linéaire général à deux classes de modèles : les modèles linéaires généralisés et les modèles linéaires mixtes.

Les modèles linéaires généralisés permettent de modéliser une variable réponse dont la distribution autour de son espérance n'est plus contrainte à être gaussienne mais peut appartenir à la famille exponentielle. Celle-ci comprend en sus de la distribution normale, un certain nombre de lois classiques notamment pour l'étude des données catégorielles (loi binomiale, multinomiale), de comptage (loi de poisson). La théorie des modèles tels que la régression logistique, la régression multinomiale, le modèle log-binomial et le modèle de poisson est présentée ainsi que des exemples d'applications dans le domaine de l'épidémiologie. On détaille les interprétations des différentes fonctions de lien et des paramètres des modèles. La stratégie de construction d'un modèle multivarié est présentée ainsi que l'évaluation de l'adéquation du modèle sélectionné, dans le cas de la régression logistique. La méthode des scores de propension qui s'appuie sur le modèle logistique est aussi présentée et appliquée à des exemples.

Les modèles linéaires mixtes sont une extension du modèle linéaire général à une variable réponse continue de distribution approximativement gaussienne avec des observations non indépendantes. Le cours développe plus particulièrement l'application aux données longitudinales répétées dans le temps. En introduction, les biais de confusion spécifiques aux données longitudinales sont abordés. La catégorisation des données manquantes est présentée et les différentes méthodes d'imputations sont développées. La théorie des modèles linéaires à structure de covariance paramétrée est ensuite détaillée puis celle des modèles linéaires mixtes. Une large part du cours concerne l'interprétation des paramètres des modèles et les critères de choix de modèles illustrés par des exemples dans le domaine clinique et biologique.

La mise en œuvre avec SAS et R est abordée de façon détaillée sous forme de programmes et sorties commentées.

- **Objectifs :**

Le cours permettra aux étudiants de construire, programmer et interpréter ces deux types de modèles d'utilisation très fréquente en épidémiologie et biologie

- **Pré-requis nécessaires :**

Connaître la théorie du modèle linéaire général et son interprétation,
Connaître les bases de la recherche clinique et épidémiologique : types d'enquêtes, mesures d'association, biais et confusion.
Savoir programmer en SAS et R

- **Responsable de l'UE : Isabelle Carrière**

- **Intervenants:**

Isabelle Carrière, INSERM

UE MACHINE LEARNING NIVEAU 1 – APPLICATION AU PRONOSTIC : R SABATIER & T MURA (2.5 ECTS)

- **Description :**

Ce second module d'analyse de données, après Data Mining en M1, a pour objectif de fournir aux étudiants des outils avancés dans l'analyse de jeux de données multivariées. Il permettra d'aborder des méthodes dont l'utilisation s'avère indispensable lorsque les méthodes plus classiques n'ont pas permis de répondre à la question posée, notamment dans un contexte de méthodes pronostiques.

Dans tous les cas, l'accent sera mis sur l'importance d'un choix éclairé et d'une bonne utilisation des méthodes. En particulier, on insistera sur l'impact de l'optimisation des paramètres pour l'ensemble de ces méthodes.

Les principes pédagogiques seront les mêmes que pour l'UE de Data Mining : des cours permettront d'acquérir les éléments théoriques fondamentaux pour maîtriser l'utilisation des méthodes enseignées qui seront mises en pratique lors de TD utilisant les logiciels R

et Python. Ces séances permettront l'approfondissement de la compréhension des différents outils proposés et fourniront également aux étudiants des scripts types qu'ils pourront réutiliser par la suite.

- **Objectifs :**
 - connaître et savoir implémenter des méthodes avancées d'analyse de données
 - maîtriser des méthodes d'optimisation de paramètres
 - savoir appliquer ces méthodes à des problématiques diagnostiques
- **Pré-requis nécessaires :**
acquis de l'UE Data Mining
- **Responsable de l'UE : Robert SABATIER et Thibault MURA**
- **Intervenants**
Thibault MURA, UM, UFR Médecine
Christelle REYNES, UM, UFR Sciences Pharmaceutiques
Robert SABATIER, UM, UFR Sciences Pharmaceutiques

UE SEMINAIRES DE RECHERCHE – P FABBRO-PERAY (2.5 ECTS)

- **Description :**
L'UE « Séminaires » a pour vocation la présentation par des experts d'un domaine méthodologique ou clinique, de leurs travaux de recherche, des enjeux, des orientations et des perspectives qui y sont associés.
L'enseignement sera complété par des séminaires de recherche bibliographique et d'écriture / lecture critique d'articles scientifiques.
- **Objectifs**
A l'issue de l'UE, les étudiants devront être capables de :
 - Mettre en œuvre une recherche bibliographique
 - Evaluer de façon critique une publication scientifique de façon adaptée aux différents types d'études cliniques ou épidémiologiques
 - Utiliser les recommandations internationales d'écriture des publications
 - Appréhender les enjeux et la temporalité d'un programme de recherche
- **Pré-requis nécessaires :**
Aucun
- **Responsable de l'UE : Pascale FABBRO-PERAY**
- **Intervenants:**
 - Pascale Fabbro-Peray, UFR Médecine
 - Nicolas Vermont, UM, BIU
 - Les experts des domaines scientifiques seront amenés à changer chaque année ; ils peuvent appartenir à des universités ou structures de recherche extérieures à l'UM

UE METHODES EN EPIDEMIOLOGIE QUANTITATIVE NIVEAU AVANCE : C BERR (5 ECTS)

- **Description :**
Cette unité d'enseignement est centrée sur l'approfondissement de mesures et de méthodes épidémiologiques orientées sur des sujets spécifiques.

La mesure : elle concerne l'exposition professionnelle, l'exposition aux agents infectieux, l'exposition environnementale ou alimentaire, les mesures subjectives en santé, et la qualité attachée à cette mesure.

Les méthodes abordées concernent :

- les stratégies d'analyse et la causalité, les nouvelles approches de la causalité (approches contrefactuelles, graphes acycliques (Directed Acyclic Graphs DAG), scores de propension,...),
- les designs particuliers (cas-témoins nichés dans les cohortes, cas-contrôles) ;
- les corrélations écologiques, les clusters, l'introduction aux analyses spatio-temporelles.
- L'application des méthodes et des concepts pour le traitement des grandes bases des données épidémiologiques ;
- les particularités méthodologiques des études pharmaco-épidémiologiques.

• **Objectifs :**

A l'issue de cette UE, les étudiants devront être capables de :

- Choisir les méthodes de mesure des expositions et de mettre en œuvre les méthodologies les plus adaptées selon le domaine épidémiologique
- Rédiger un protocole d'étude épidémiologique en fonction de ces spécificités
- Lire de façon critique les publications issues d'études épidémiologiques selon les guidelines adaptés

• **Pré-requis nécessaires :**

Aucun

• **Responsable de l'UE : Claudine BERR**

Intervenants :

- Claudine Berr, INSERM
- Emmanuelle Cadot, IRD
- Isabelle Carrière, INSERM
- Pascale Fabbro, UFR Médecine
- Jean-Luc Faillie, UFR Médecine
- Nicolas Nagot, UFR Médecine
- Fabienne Séguret, CHU Montpellier
- Brigitte Trétarre, Registre des tumeurs de l'Hérault

UE META-ANALYSES P FABBRO-PERAY (2.5 ECTS)

• **Description :**

La méta-analyse (MA) a pris une place importante dans la littérature médicale en matière d'évaluation et de changements de pratiques. Les notions essentielles suivantes seront développées :

- Justification d'une revue de littérature avec ou sans méta-analyse, définition, objectifs, variables principales, recherche et sélection de la littérature, évaluation de la qualité & analyse des biais au moyen de grilles validées.
- Types de méta-analyse : MA sur données agrégées, MA sur données individuelles, MA en réseau.
- Synthèse statistique des résultats.
- Méthodes de mesure de l'effet moyen :
- en fonction du type de MA et du type de variable : binaire, continue, censurée, du paramètre adéquat : odds ratio, risque relatif, hazard ratio ; prévalence, incidence, courbe ROC , coefficient de corrélation

- en fonction du type d'études & de la problématique : essais cliniques randomisés, études cas-témoin, études de cohorte, étude de prévalence, études d'évaluation diagnostique.;
 - Méthodes de Recherche de l'hétérogénéité. Méthodes de modélisation à effets fixes ou à effets aléatoires ;
 - Méthodes de recherche du biais de publication
 - Applications pratiques sur logiciel R . Discussion des résultats.
 - Ecriture d'un protocole de méta-analyse et registre PROSPERO.
 - Lecture critique d'une méta-analyse fondée sur les recommandations PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), pour les essais randomisés et fondée sur le MOOSE (Meta-analysis Of Observational Studies in Epidemiology)
- **Objectifs :**
A l'issue de cette UE, les étudiants devront être capables de :
 - Choisir la méthodologie adaptée au calcul méta-analytique en fonction de l'objectif
 - Rédiger un protocole de méta-analyse
 - Mettre en œuvre l'analyse statistique d'une MA sur données agrégées ou sur données individuelles
 - Lire de façon critique une publication de méta-analyse
 - **Pré-requis nécessaires :**
Les acquis des UE :
 - « Essais cliniques randomisés », « Méthodes en épidémiologie quantitative niveau de base », « Méthodes en épidémiologie quantitative niveau avancé »,
 - « Communication » et « Séminaires » pour la partie Recherche bibliographique et LCA
 - **Responsable de l'UE : Pascale FABBRO-PERAY**
 - **Intervenants**
 - Pascale FABBRO-PERAY, UFR Médecine
 - Florian NAUDET, Université de Rennes
 - Bob-Valery OCCEAN, CHU de Nîmes

BLOC OPTIONNEL 1 : UE METHODOLOGIES SPECIFIQUES

Le choix du bloc optionnel 1 concerne l'ensemble de 4 UEs pour les étudiants du sous parcours « Données de Santé ».

UE MEDICO-ECONOMIE : S CHKAIR – S BOUVET (2.5 ECTS)

- **Description :**
Les objectifs pédagogiques de cette UE sont de comprendre les enjeux médico-économiques dans les projets de recherche clinique en France et d'être en mesure de mener à bien une réflexion méthodologique et statistique autour d'un projet médico-économique.
L'UE est composée de 7 cours qui permettront de :
 - Poser le contexte et comprendre les leviers économiques en santé autour d'un cours sur l'historique des systèmes de santé.
 - Définir les grands concepts et les recommandations méthodologiques pour de l'évaluation médico-économique

- Comprendre l'enjeu crucial relatif aux données ; savoir où et comment obtenir les données nécessaires au montage des projets médico-économiques
- Appréhender les analyses médico-économiques par modélisation grâce au logiciel TreeAgePro.
- Connaître les différentes méthodes de costing pour estimer le coût d'un acte médical (ABC, ICR, micro-costing)
- Savoir mettre en œuvre une Analyse d'Impact Budgétaire (AIB) et découvrir l'évaluation des technologies de santé
- Comprendre la méthodologie des choix discrets (DCE) et son recours en médico-économie.
- **Objectifs :**
Les étudiants devront être en mesure, à l'issue de cette UE, de rédiger un protocole de recherche en médico-économie sur un sujet de leur choix. Ils devront être en mesure d'identifier le type d'étude à mettre en œuvre ainsi que tous les aspects techniques définissant la recherche médico-économique, les données à mobiliser et les outils statistiques nécessaires à l'analyse.
Les étudiants devront également être capables de comprendre des résultats d'une étude et d'analyser une problématique médico-économique.
- **Pré-requis nécessaires :**
Aucun
- **Responsables de l'UE : Sihame Chkair et Sophie Bouvet**
- **Intervenants :**
Thierry Blayac UM, UFR Economie
Sophie Bouvet CHU Nîmes
Sihame Chkair CHU Nîmes

UE QUALITE DE VIE : C. GAUJOUX-VIALA (2.5 ECTS)

- **Description**
La notion de qualité de vie et son historique seront abordés. Qu'est-ce que la qualité de vie liée à la santé ? Pourquoi mesurer la qualité de vie ? Quelles études pour mesurer la qualité de vie ? Comment mesurer la qualité de vie ? Qu'est-ce qu'un critère rapporté par le patient (Patient Reported Outcome) ? Approche de différents types d'instruments et de leurs propriétés. Les échelles de mesure de qualité de vie seront présentées : principe, construction et taxonomie. Les scores et leurs mesures seront introduits : validité, fiabilité, sensibilité. Analyse et mesure des différentes propriétés psychométriques. La méthodologie de construction des échelles de mesures de qualité de vie liées à la santé seront présentées. Puis la méthodologie de validation de ces échelles. Illustration pratique du développement et de la validation d'un score d'impact en rhumatologie. Autre illustration au travers des essais thérapeutiques en oncologie. Enfin, l'utilisation de la qualité de vie en pratique médicale courante et la place dans le processus de décision seront exposés.
- **Objectifs :**
 - comprendre le concept de qualité de vie et son intérêt
 - définition et utilisation d'un Patient Reported outcome
 - Connaître et savoir analyser les différentes propriétés psychométriques d'un instrument de qualité de vie
 - Connaître la méthodologie de construction des échelles de mesures de qualité de vie
 - Connaître la méthodologie de validation des échelles de mesures de qualité de vie

- **Pré-requis nécessaires :**
Aucun
- **Responsable de l'UE : Pr Cécile Gaujoux-Viala**
- **Intervenants**
Frédéric Fiteni, UM, UFR Médecine
Cédric Lukas, UM, UFR Médecine
Cécile Gaujoux-Viala, UM, UFR Médecine
Laure Gossec, Université Paris VI

UE METHODES EN RECHERCHE QUALITATIVE : A OUDE-ENGBERINK (2.5 ECTS)

- **Description :**
La recherche utilisant des méthodes qualitatives en science de la santé regroupent des méthodologie de recueil et d'analyse de données issues des Sciences Humaines et Sociales qui permettent de comprendre les phénomènes complexes relevant de l'expérience vécue des acteurs du soin (patients, aidants, soignants ...). Ces méthodologies rigoureuses répondent à des critères scientifiques spécifiques de recueil et d'analyse de données de terrain. Elles apportent des connaissances complémentaires à celles des approches quantitatives et biostatistiques en leur donnant du sens. A une logique souvent vérificatoire, elles substituent une logique de découverte et d'innovation. Elles permettent de « repenser le concept d'EBM » (Sackett) en apportant des connaissances sur « l'expérience des soignants » et les « perspective des patients ». L'enjeu est important aujourd'hui où, plus que jamais, nous devons disposer d'une information multidimensionnelle pour comprendre et décider. Parmi les phénomènes étudiés par les approches qualitatives, le plus important en santé est « l'expérience vécue de la maladie par les patients et leur entourage », mais aussi, les comportements en santé (déli, (in)observance, addictions..), la relation à l'autre (alliance thérapeutique), les processus (ETP), les organisations (les interactions sociales, nouvelles organisations de soins), la santé environnementale (l'exposome), les conditions d'implémentation d'une intervention.
La finalité de ces approches est d'apporter de la connaissance nouvelle et diversifiée pour l'amélioration des pratiques de soins.
Exemple d'objectif de recherche qualitative (RQI) : (ou phénomènes étudiés)
 - Comprendre les effets de l'activité physique adaptée sur la fatigue des femmes atteintes d'un cancer du sein sous chimiothérapie incluse dans un essai thérapeutique.
 - Comprendre la décision vaccinale de parents au sujet de leurs enfants.
 - Comprendre et prévenir les risques liées à l'épilepsie à partir de l'expérience de patients, de proches, de proches endeuillés.
 - Explorer les freins et leviers à l'utilisation d'un outil de télémédecine par les professionnels et des patients.
 - Comprendre l'impact de la crise sanitaire sur le parcours de soin de patients DT2.
- L'UE s'articulera autour de 5 modules :
- 1 - Principes et Bases épistémologiques de la Recherche Qualitative
 - 2- Problématisation en RQI : de la question de départ à la construction d'un objet de recherche, en intégrant le cadre réglementaire et éthique.
 - 3 -Recueil de données en RQI : principales méthodes de recueil : observation participante, entretien semi directif, individuel ou collectif, avec un accent sur l'entretien d'explicitation de l'expérience vécue d'un phénomène. Un atelier sera

centré sur la construction d'un dispositif de recueil : guide d'entretien et données descriptives.

4- Bases et méthodes de l'Analyse de données en RQI : comprendre le continuum méthodologique à travers les opérations d'analyse de verbatim (matériau textuel issu de la transcription des entretiens). Connaître les principales approches méthodologiques (analyse thématique, phénoménologique-pragmatique, Théorisation ancrée, méthodes mixtes), leur pertinence en fonction de la question de recherche, le rendu des résultats.

5- la rédaction d'article en RQI.

- **Objectifs**

A l'issue de l'UE les étudiants devront être capables d'analyser la démarche en recherche qualitative : problématisation, recueil et analyse des données qualitatives à partir de publication de travaux de recherche qualitative.

- **Pré-requis nécessaires :**

Aucun

- **Responsable de l'UE : Oude Engberink Agnès**

- **Intervenants :**

Bourrel Gérard, UM, UFR Médecine
Lognos Béatrice, UM, UFR Médecine
Clary Bernard , UM, UFR Médecine
Million Elodie, UM, UFR Médecine
Pavageau , UM, UFR Médecine
Serayet Phillipe, UM, UFR Médecine
Kinné Mélanie chargée d'enseignement, CHU Nîmes
Minet Mathilde chargée d'enseignement
Hermine Saguy , UM, UFR Médecine
Cloé Brami Oncologue (intervenant extérieur)

UE MACHINE LEARNING NIVEAU 2 –: R SABATIER (2.5 ECTS)

- **Description :**

Les outils proposés lors de cette UE sont transversaux par rapport à ceux des deux premiers niveaux. Il s'agit de montrer aux étudiants des techniques qui sont développées en amont ou en aval des méthodes précédemment vues. La première partie concernera les jeux de données pour lesquels il est nécessaire de réaliser une sélection de variables, que ce soit pour des raisons de taille du jeu de données ou pour fournir un modèle plus facilement interprétable pour les utilisateurs. On abordera les méthodes de type filtres (pré-tri indépendant de la question posée), de type enveloppantes (déterministes, gourmandes ou heuristiques) et de type embarqué. Dans une seconde partie, on s'intéressera à la validation de modèle dont l'objectif sera d'évaluer la performance d'un modèle et de la quantifier de sorte à pouvoir réaliser une sélection de modèle.

Les principes pédagogiques seront les mêmes que pour les autres UEs d'analyse de données : des cours permettront d'acquérir les éléments théoriques fondamentaux pour maîtriser l'utilisation des méthodes enseignées qui seront mises en pratique lors de TD utilisant les logiciels R et Python. Ces séances permettront l'approfondissement de la compréhension des différents outils proposés et fourniront également aux étudiants des scripts types qu'ils pourront réutiliser par la suite.

- **Objectifs :**

connaître et savoir implémenter des méthodes avancées d'analyse de données

savoir programmer des adaptations de méthodes classiques

- **Pré-requis nécessaires :**
acquis de l'UE Data Mining et Machine Learning 1
- **Responsable de l'UE : Robert SABATIER**
- **Intervenants:**
Christelle REYNES, UM, UFR Sciences Pharmaceutiques
Robert SABATIER, UM, UFR Sciences Pharmaceutiques

BLOC OPTIONNEL 2 : UE RADIOMICS : L'IMAGERIE QUANTITATIVE ET L'IA EN RADIOLOGIE

Ces unités d'enseignement sont réservées aux étudiants prenant une orientation vers l'imagerie quantitative & l'Intelligence artificielle en radiologie. Le choix concerne alors l'ensemble de ces 4 UEs pour les étudiants du sous parcours « Données de Santé ».

UE CONSTITUTION DU SIGNAL MORPHO-PHYSIOLOGIQUE EN IMAGERIE MEDICALE : F PEREIRA (2.5 ECTS)

- **Description**
Les objectifs pédagogiques concernent l'appropriation des phénomènes physiques utilisés à la description d'un phénomène biologique à partir de la constitution de l'image. Les notions techniques seront approfondies dans l'UE « Traitement numérique d'image médicale » et les applications cliniques seront étudiées dans les UEs « Méthodes en imagerie quantitative vasculaire, cardiaque et en radiologie interventionnelle » et « Méthodes en imagerie quantitative neurologique et oncologique » :
 - L'imagerie par radiation de haute énergie (CT/Rx) ;
 - L'imagerie par radio-pharmaceutiques (PET/SPECT) ;
 - L'imagerie par ondes mécaniques (Échographie) ;
 - L'imagerie par le principe électromagnétique (EEG et MEG) ;
 - L'imagerie par résonance magnétique (IRM) ;
 - L'imagerie par radiation de basse énergie (optique) ;
 - Reconstitutions multidimensionnelles (axiales, tridimensionnelles, multiplanaires, espace-temporelle et tensorielle).
 - L'imagerie dans l'espace de fréquence ;
 - Le sinogramme et le produit de convolution ;
 - L'algorithme & langage de programmation (Matlab/Python) pour la reconstruction d'un signal biologique ;
 - Le format compression DICOM – Nifti ;
 - L'archivage, télétransmission et télédiagnostic
- **Objectifs**
L'objectif est l'introduction à la génération du signal biomédical, les bases physiques des signaux dans un système vivant, les formulations mathématiques et les limitations pour permettre les traitements des signaux et une potentielle inférence clinique. Les étudiants devront connaître l'analyse spectrale et les méthodes classiques de manipulation des images, savoir définir les formats des images et les particularités pour la recherche clinique (ie anonymisation, qualité ...), la base d'archivage, de télétransmission et de télédiagnostic avec une approche informatique, juridique, médicale et financière. On vise aussi l'application de langage de programmation et développement en routine pour le traitement des images.

- **Pré-requis nécessaires :**
Connaissance des bases en physique et mathématique
Connaissance des bases en « line de command » (Bash, Shell, etc)
- **Pré-requis recommandés :**
Connaissance de langage de programmation (Python, R ou Matlab)
- **Responsable de l'UE : Fabrício PEREIRA**
- **Intervenants:**
BEREGI Jean Paul, UM, UFR Médecine
GREFFIER Joel, CHU de Nîmes
PEREIRA Fabricio, CHU de Nîmes
SANS Remi, CHU de Nîmes

UE TRAITEMENT NUMERIQUE DE L'IMAGE MEDICALE : F PEREIRA (2.5 ECTS)

- **Description:**
Les objectifs pédagogiques concernent une introduction aux différentes méthodes utilisées pour le traitement numérique des images médicales du point de vue algorithmique et mathématique. Les applications cliniques seront étudiées dans les UEs « Méthodes en imagerie quantitative vasculaire, cardiaque et en radiologie interventionnelle » et « Méthodes en imagerie quantitative neurologique et oncologique » :
 - Qualité en imagerie médicale;
 - Différentes méthodes de filtrages (régularisation, fréquentielle, convolutionnelle) ;
 - Segmentation semi-automatique;
 - Détection de contours et de lignes ;
 - Techniques de seuillage ;
 - Restauration, reconstruction et recalage d'images médicales ;
 - Fusion d'image et co-enregistrement ;
 - Modelage et simulation par méthodes stochastique (théorie des graphes et simulation de Monte-Carlo);
 - Méthodes d'inférence causale (causalité de granger, modèle de causalité dynamique et modèle d'équation structurales)
 - Imagerie quantitative : tenseur de diffusion ; résonance magnétique fonctionnelle ; morphométrie ; analyses de séries multidimensionnelles
 - Introduction à intelligence artificielle appliquées à l'imagerie médicale ;
 - Reconnaissance de forme par les algorithmes d'apprentissage et d'exploration
- **Objectifs:**
L'objectif est l'introduction aux algorithmes pour le rehaussement et la restauration d'images médicales en utilisant diverses méthodes numériques.
Après avoir suivi le cours, l'étudiant devrait être en mesure d'appliquer des méthodes d'investigation moderne utilisées pour le traitement d'images médicale, telles que les simulations, l'intelligence artificielle, l'intégration avec métadonnées, les modelages, et le radiomics
- **Pré-requis nécessaires :**
Connaissance de langage de programmation (Python, R ou Matlab)
Connaissance des line de command (Bash, Shell, etc)
- **Pré-requis recommandés :**
Connaissance des bases de l'inférence bayésienne
Connaissance des bases de modèle non-linéaire

- **Responsable de l'UE: Fabricio PEREIRA**
- **Intervenants:**
BEREGI Jean Paul, UM, UFR Médecine
GREFFIER Joel, CHU de Nîmes
PEREIRA Fabricio, CHU de Nîmes

UE METHODES EN IMAGERIE QUANTITATIVE NEUROLOGIQUE/ONCOLOGIQUE : JP. BEREGI – N MENJOT DE CHAMPFLEUR (2.5 ECTS)

- **Description :**
Les objectifs pédagogiques sont orientés vers la connaissance des méthodes en radiologie et imagerie neurologique pour l'évaluation des pathologies neurologiques et oncologiques :
 - IRM cérébrale morphologique et quantification des éléments cérébraux morphologiques, cartographie du cerveau
 - IRM cérébrale de perfusion : technique et intérêts
 - IRM fonctionnelle : techniques et applications
 - IRM DTI : technique et identification des fibres du cerveau ou dans le corps
 - Cours sur les techniques d'acquisition du fiber tracking
 - Diffusion du corps entier
 - Analyse tumorale par imagerie volumique
 - Applications dans les pathologies tumorales, dégénératives pour des applications diagnostiques, prédictives et pronostiques
- **Objectifs:**
L'objectif est de connaître les méthodes de quantifications des lésions neurologiques et oncologiques tant morphologiques que fonctionnelles (activité ?) pour adapter le traitement. La description anatomique précise ne suffit plus ou n'est pas suffisante dans de nombreuses situations. La description d'une lésion volumique tumorale doit être complétée par une approche plus précise par la radiomique pour en connaître les caractéristiques génétiques et permettre une prédiction dans la réponse à une thérapeutique ou éviter une biopsie. L'analyse des lésions doit répondre également à son caractère actif ou non. De même en pathologie nerveuse, l'atteinte des nerfs n'est pas visible anatomiquement mais fonctionnellement. Il est donc nécessaire de connaître les moyens d'en approcher l'analyse fonctionnelle pour déceler un trouble de la conduction par exemple ou une démyélinisation.
- **Pré-requis nécessaires :**
Connaissance de l'anatomie cérébrale et nerveuse
Connaissance des bases en IRM et TDM
- **Pré-requis recommandés :**
Connaissance des tumeurs bénignes et malignes
- **Responsables de l'UE: Jean Paul BEREGI / Nicolas MENJOT DE CHAMPFLEUR**
- **Intervenants:**
Jean-Paul BEREGI, UM, UFR de Médecine
C Cassinotto,
Boris GUIU, UM, UFR de Médecine
Nicolas MENJOT DE CHAMPFLEUR, UM, UFR de Médecine
Ingrid MILLET-CENAC, UM, UFR de Médecine

UE METHODES EN IMAGERIE QUANTITATIVE CARDIOVASCULAIRE & EN RADIOLOGIE INTERVENTIONNELLE: JP. BEREGLI – H KOVACSIK (2.5 ECTS)

- **Description :**

Les objectifs pédagogiques sont orientés vers la connaissance des méthodes en radiologie et imagerie vasculaire, cardiaque et radiologie interventionnelle pour l'évaluation et le traitement des pathologies neurologiques, oncologiques, vasculaires, cardiaques :

- Définition et quantification d'une sténose
- Définition et quantification d'un trouble de la perfusion
- Définition et quantification de la viabilité d'un territoire
- Discussion des méthodes de quantification par imagerie volumique et interventionnelle
- Cartographie des lésions pré-interventionnelle, simulation, impression 3D, programmation d'une intervention basée sur l'image
- Applications dans les pathologies tumorales, oncologiques, vasculaire, cardiaques pour des applications diagnostiques, prédictives et pronostiques

- **Objectifs :**

L'objectif est d'approcher la pathologie ischémique avec des techniques permettant de confirmer la réalité de la maladie et de traiter non pas l'image d'une lésion sténosante mais bien une souffrance d'un territoire qui entraîne des symptômes et une pathologie voire un décès. La quantification des lésions permet une programmation des interventions réduisant le temps interventionnel et donc les complications. La médecine personnalisée est possible avec la nécessité d'adapter les traitements en fonction des besoins et de la réalité anatomique. L'évaluation du retentissement d'une atteinte artérielle sur l'organe d'aval est primordiale. De nombreuses approches sont possibles avec des avantages et des limites.

- **Pré-requis nécessaires:**

Connaissance de l'anatomie vasculaire et cardiaque
Connaissance des lésions de perfusion, d'ischémie, de viabilité, de nécrose
Connaissance des bases en IRM, TDM

- **Pré-requis recommandés :**

Pathologies cardiaques, vasculaires, tumorales, ...

- **Responsables de l'UE : Jean-Paul BEREGLI / Hélène KOVACSIK**

- **Intervenants:**

Jan-Paul BEREGLI, UFR de Médecine
Vincent COSTALAT, UFR de Médecine
Julien FRANDON, CHU de Nîmes
Hélène KOVACSIK, UFR de Médecine
Patrice TAOUREL, UFR de Médecine

SEMESTRE 4

UE STAGE : T MURA (30 ECTS)

- **Description:**

L'UE stage 2 se déroulera au quatrième semestre du master « santé » parcours EDSB. Il s'agit d'un stage de 4 à 6 mois, sur un lieu d'exercice de l'épidémiologie (hôpitaux, unité de recherche, institut de surveillance épidémiologique, ONG, laboratoire ou structure privée...). Ce stage devra permettre aux étudiants d'acquérir les compétences permettant la mise en application des connaissances acquises au cours du master. L'ensemble du stage devra porter sur un projet d'analyse épidémiologique (épidémiologie de terrain ou de recherche, clinique ou populationnelle à visée descriptive, analytique ou évaluative).

- **Objectifs**

Au cours de leur stage, les étudiants devront notamment :

- Participer à la réflexion visant à justifier le projet, notamment en réalisant (ou complétant) une analyse bibliographique sur le sujet du stage.
- Formuler les objectifs du projet.
- S'assurer de la validité des données fournies (contrôles de cohérence),
- Mettre en œuvre les analyses statistiques de bases (à l'aide des outils informatiques SAS, R ou Python)
- Mettre en œuvre la modélisation statistique des données après s'être assuré de l'adéquation entre la question posée et le modèle utilisé (à l'aide des outils informatiques SAS, R ou Python).
- Interpréter les résultats issus du projet et des modélisations.
- Mettre en perspective les résultats obtenus en regard des connaissances existantes.
- Rédiger un mémoire au format « article de recherche » (structure : Introduction/Justification, Résultats, Discussion/Interprétation, Conclusion)

- **Pré-requis nécessaires:**

Maîtriser le contenu théorique du master

- **Responsable de l'UE: Thibault MURA-TODESCO**

MASTER 2 SOUS PARCOURS « DATA ANALYST POUR LES SCIENCES DU VIVANT »

SEMESTRE 3

UE ANALYSE DES DONNEES CENSUREES : T MURA (2.5 ECTS)

- **Description**

L'analyse des données censurées a pris une place majeure en santé publique et en médecine. L'objectif de ce cours est de présenter les notions de base sur l'analyse de données censurées (analyse de survie) et de détailler les modèles les plus communément utilisés pour leur analyse. Les notions de censure à droite, à gauche et par intervalle seront présentées, puis les notions de censure non informative et de censure informative. L'analyse de la survie proprement dite sera abordée, à l'aide de méthode non paramétrique (estimateur de Nelson Aalen, de Kaplan Meier et actuarielle) et survie paramétrique (modèle exponentiel, Weibull, Weibull Généralisé), survie semi paramétrique (modèle de Cox) et comparaison de survies. L'introduction de covariables dans les cas paramétriques et semi paramétrique, en particulier le modèle de Cox, sera ensuite abordée. Les cas particuliers des modèles à entrée retardée, l'utilisation de variables dépendantes du temps et les modèles de survie à risques compétitifs seront également traités. Le lien entre l'approche de survie avec le processus Markovien ou non Markovien sera introduit. Le calcul du nombre de sujets à inclure dans le cas de la comparaison de courbes de survie sera également présenté.

- **Objectifs**

Les étudiants devront être capable, à l'issue de cette UE de manipuler les principales méthodes d'analyses de données censurées, en les appliquant à des données réelles :

- en choisissant le modèle le plus adapté aux données et à la question posée,
- en calculant à la main les estimateurs de bases (Nelson Aalen, de Kaplan Meier et actuarielle)
- en calculant à la main un test du log-Rank
- en programmant les analyses basées sur modèles non paramétrique, paramétriques et semi paramétriques sur les logiciels R ou SAS,
- en étant capable de vérifier les conditions de validité des modèles
- et étant capable d'interpréter les résultats qui en sont issus

- **Pré-requis nécessaires :**

Notion de base (import de données, gestion et manipulation de données, statistiques de base) sur les logiciels SAS et R.

- **Responsable de l'UE : Thibault MURA-TODESCO**

- **Intervenants:**

Thibault MURA-TODESCO, UM, UFR Médecine
Laure-Anne Gutierrez, INSERM

UE MODELE LINEAIRE GENERALISE ET MODELE MIXTE : I. CARRIERE (5 ECTS)

- **Description :**

L'objectif du cours est d'étendre les acquis du cours modèle linéaire général à deux classes de modèles : les modèles linéaires généralisés et les modèles linéaires mixtes.

Les modèles linéaires généralisés permettent de modéliser une variable réponse dont la distribution autour de son espérance n'est plus contrainte à être gaussienne mais peut appartenir à la famille exponentielle. Celle-ci comprend en sus de la distribution normale, un certain nombre de lois classiques notamment pour l'étude des données catégorielles (loi binomiale, multinomiale), de comptage (loi de poisson). La théorie des modèles tels que la régression logistique, la régression multinomiale, le modèle log-binomial et le modèle de poisson est présentée ainsi que des exemples d'applications dans le domaine de l'épidémiologie. On détaille les interprétations des différentes fonctions de lien et des paramètres des modèles. La stratégie de construction d'un modèle multivarié est présentée ainsi que l'évaluation de l'adéquation du modèle sélectionné, dans le cas de la régression logistique. La méthode des scores de propension qui s'appuie sur le modèle logistique est aussi présentée et appliquée à des exemples.

Les modèles linéaires mixtes sont une extension du modèle linéaire général à une variable réponse continue de distribution approximativement gaussienne avec des observations non indépendantes. Le cours développe plus particulièrement l'application aux données longitudinales répétées dans le temps. En introduction, les biais de confusion spécifiques aux données longitudinales sont abordés. La catégorisation des données manquantes est présentée et les différentes méthodes d'imputations sont développées. La théorie des modèles linéaires à structure de covariance paramétrée est ensuite détaillée puis celle des modèles linéaires mixtes. Une large part du cours concerne l'interprétation des paramètres des modèles et les critères de choix de modèles illustrés par des exemples dans le domaine clinique et biologique.

La mise en œuvre avec SAS et R est abordée de façon détaillée sous forme de programmes et sorties commentées.

- **Objectifs :**

Le cours permettra aux étudiants de construire, programmer et interpréter ces deux types de modèles d'utilisation très fréquente en épidémiologie et biologie

- **Pré-requis nécessaires :**

Connaître la théorie du modèle linéaire général et son interprétation,

Connaître les bases de la recherche clinique et épidémiologique : types d'enquêtes, mesures d'association, biais et confusion.

Savoir programmer en SAS et R

- **Responsable de l'UE : Isabelle Carrière**

- **Intervenants:**

Isabelle Carrière, INSERM

UE MACHINE LEARNING NIVEAU 1 – APPLICATION AU PRONOSTIC : R SABATIER & T MURA (2.5 ECTS)

- **Description :**

Ce second module d'analyse de données, après Data Mining en M1, a pour objectif de fournir aux étudiants des outils avancés dans l'analyse de jeux de données multivariées. Il permettra d'aborder des méthodes dont l'utilisation s'avère indispensable lorsque les méthodes plus classiques n'ont pas permis de répondre à la question posée, notamment dans un contexte de méthodes pronostiques.

Dans tous les cas, l'accent sera mis sur l'importance d'un choix éclairé et d'une bonne utilisation des méthodes. En particulier, on insistera sur l'impact de l'optimisation des paramètres pour l'ensemble de ces méthodes.

Les principes pédagogiques seront les mêmes que pour l'UE de Data Mining : des cours permettront d'acquérir les éléments théoriques fondamentaux pour maîtriser l'utilisation des méthodes enseignées qui seront mises en pratique lors de TD utilisant les logiciels R et Python. Ces séances permettront l'approfondissement de la compréhension des différents outils proposés et fourniront également aux étudiants des scripts types qu'ils pourront réutiliser par la suite.

- **Objectifs :**
connaître et savoir implémenter des méthodes avancées d'analyse de données
maîtriser des méthodes d'optimisation de paramètres
savoir appliquer ces méthodes à des problématiques diagnostiques
- **Pré-requis nécessaires :**
acquis de l'UE Data Mining
- **Responsable de l'UE : Robert SABATIER et Thibault MURA**
- **Intervenants**
Thibault MURA, UM, UFR Médecine
Christelle REYNES, UM, UFR Sciences Pharmaceutiques
Robert SABATIER, UM, UFR Sciences Pharmaceutiques

UE MACHINE LEARNING NIVEAU 2 –: R SABATIER (2.5 ECTS)

- **Description :**
Les outils proposés lors de cette UE sont transversaux par rapport à ceux des deux premiers niveaux. Il s'agit de montrer aux étudiants des techniques qui sont développées en amont ou en aval des méthodes précédemment vues. La première partie concernera les jeux de données pour lesquels il est nécessaire de réaliser une sélection de variables, que ce soit pour des raisons de taille du jeu de données ou pour fournir un modèle plus facilement interprétable pour les utilisateurs. On abordera les méthodes de type filtres (pré-tri indépendant de la question posée), de type enveloppantes (déterministes, gourmandes ou heuristiques) et de type embarquées. Dans une seconde partie, on s'intéressera à la validation de modèle dont l'objectif sera d'évaluer la performance d'un modèle et de la quantifier de sorte à pouvoir réaliser une sélection de modèle.
Les principes pédagogiques seront les mêmes que pour les autres UEs d'analyse de données : des cours permettront d'acquérir les éléments théoriques fondamentaux pour maîtriser l'utilisation des méthodes enseignées qui seront mises en pratique lors de TD utilisant les logiciels R et Python. Ces séances permettront l'approfondissement de la compréhension des différents outils proposés et fourniront également aux étudiants des scripts types qu'ils pourront réutiliser par la suite.
- **Objectifs :**
 - connaître et savoir implémenter des méthodes avancées d'analyse de données
 - savoir programmer des adaptations de méthodes classiques
- **Pré-requis nécessaires :**
acquis de l'UE Data Mining et Machine Learning 1
- **Responsable de l'UE : Robert SABATIER**
- **Intervenants:**
Christelle REYNES, UFR Sciences Pharmaceutiques
Robert SABATIER, UFR Sciences Pharmaceutiques

UE SEMINAIRES DE RECHERCHE – P FABBRO-PERAY (2.5 ECTS)

- **Description :**

L'UE « Séminaires » a pour vocation la présentation par des experts d'un domaine méthodologique ou clinique, de leurs travaux de recherche, des enjeux, des orientations et des perspectives qui y sont associés.

L'enseignement sera complété par des séminaires de recherche bibliographique et d'écriture / lecture critique d'articles scientifiques.

- **Objectifs**

A l'issue de l'UE, les étudiants devront être capables de :

- Mettre en œuvre une recherche bibliographique
- Evaluer de façon critique une publication scientifique de façon adaptée aux différents types d'études cliniques ou épidémiologiques
- Utiliser les recommandations internationales d'écriture des publications
- Appréhender les enjeux et la temporalité d'un programme de recherche

- **Pré-requis nécessaires :**

Aucun

- **Responsable de l'UE : Pascale FABBRO-PERAY**

- **Intervenants:**

- Pascale Fabbro-Peray, UM, UFR Médecine
- Nicolas Vermont, BU, UM
- Les experts des domaines scientifiques seront amenés à changer chaque année ; ils peuvent appartenir à des universités ou structures de recherche extérieures à l'UM

UE BASE DE DONNEES – NIVEAU 2 : M. VIVIEN (2.5 ECTS)

- **Description :**

Le cours de data-management de seconde année du master est une mise en application des notions vues en première année et se focalise sur le recueil de données et leur administration. La notion de base de données permet de faire le lien entre certaines des matières enseignées dans le master : pour faire une base de données, il faut

- un recueil pertinent de données (module de recueil) ;
- organiser la saisie des données (base de données+ formulaires de saisie et contrôle de cohérence) par exemple sur un site internet sécurisé (chaque étudiant fait un formulaire et va saisir ses données) ;
- croiser ces données (pourquoi ne pas les géocoder par exemple), ou les lier à des bibliothèques (ex sensométrie) ;
- extraire les données pour les faire analyser dans les autres modules ;
- récupérer les résultats et les envoyer automatiquement aux décideurs.
- Le cours sera organisé autour d'exemples de mise en situations réelles, dans un profil big-data. Une partie de ces applications sera réalisée sous oracle et mysql.

- **Objectifs:**

- maîtriser des concepts avancés de bases de données
-

- **Pré-requis nécessaires :**

acquis de l'UE Bases de Données niveau 1

- **Responsable de l'UE : Myrtille VIVIEN**
- **Intervenants :**
Myrtille VIVIEN, UM, UFR Sciences Pharmaceutiques

UE RECUEIL PLANIFIE DES DONNEES : E GOZE (2.5 ECTS)

- **Description :**
Deux étapes sont cruciales dans toute étude statistique : le recueil des données et l'analyse de ces données. Souvent l'accent est mis sur l'ensemble des méthodes d'analyse statistique des données, mais la question des méthodes de recueil est souvent négligée, bien qu'essentielle.
On considère deux modes de recueil de l'information : l'expérimentation et l'échantillonnage. L'objet de l'UE « Recueil planifié de données » est d'apprendre à maîtriser les techniques de contrôle de l'hétérogénéité ou du choix des objets comparés en expérimentation, et les façons d'obtenir la représentativité en échantillonnage. Différents types de plans d'expérience et différentes stratégies d'échantillonnage sont exposées en relation avec les objectifs visés.
Sont abordées les notions de randomisation, de bloc (complet ou incomplet), d'interaction pour les expériences factorielles, de confounding et de répétitions fractionnaires, enfin de plans à plusieurs échelles et de plans optimaux. En échantillonnage, sont abordées les notions de stratégie d'échantillonnage, de plans de sondage à un ou plusieurs degrés, de plans séquentiels et de plans basés sur des modèles.
- **Objectifs :**
à maîtriser les techniques de contrôle de l'hétérogénéité ou du choix des objets comparés en expérimentation, et les façons d'obtenir la représentativité en échantillonnage
- **Pré-requis nécessaires :**
acquis de l'UE Modèle Linéaire Général
- **Responsable de l'UE : Eric GOZE**
- **Intervenants:**
Eric GOZE, CIRAD

UE STATISTIQUE TEMPORELLE : JN BACRO (2.5 ECTS)

- **Description :**
En biologie, les données à traiter présentent très souvent une dépendance temporelle qui doit être prise en compte dans les traitements statistiques des données ; typiquement : mesures horaires (ou toute autre période de temps) d'une quantité caractérisant l'évolution d'une population, relevés temporels de variations d'une variable réponse selon des valeurs d'entrée dans un système etc. Le temps peut jouer un rôle particulier dans les phénomènes étudiés (cycles, valeurs liées au passé etc) et des méthodes spécifiques doivent être utilisées pour traiter ce type de données. Après avoir motivé la nécessité d'approches intégrant la prise en compte d'une éventuelle dépendance au temps dans des relevés chronologique de données, le cours présentera les principaux outils et modèles pour le traitement statistique des séries chronologiques. Les aspects descriptifs et modélisation d'une série seront successivement abordés. Les notions de processus stochastique, d'auto-corrélations, de fonction de corrélation (partielle ou non) seront présentées et explicitées. Le plan du cours est le suivant :

- introduction : motivation, données, questions ;
- décomposition d'une série chronologique ;
- les différentes composantes pour décrire une série chronologique ;
- lissage et prédiction : moyenne mobile ; lissage exponentiel ;
- modélisation stochastique : modèles auto-régressifs $AR(p)$; modèles moyennes mobiles $MA(q)$;
- généralisation : modèles $ARIMA(p,q)$; $SARIMA(p,d,q)$.

Les notions introduites en cours seront appliquées sur des séries simulées et des jeux de données réelles avec le logiciel R.

- **Objectifs :**
 - savoir prendre en compte dans les différents modèles la variable « temps » avec différents types de prise en compte
 -
- **Pré-requis nécessaires :**
 - acquis de l'UE Data Mining et Machine Learning 1
- **Responsable de l'UE : Jean-Noël BACRO**
- **Intervenants:**
 - Jean-Noël BACRO, UM, UFR Médecine

UE STATISTIQUE POUR L'INDUSTRIE : ? (2.5 ECTS)

- **Description :**

Cette UE sera constituée d'intervention de professionnels de la biostatistique qui travaillent dans différents types d'industrie. Ils feront part de leur expérience et proposeront des applications pratiques de méthodologies particulièrement adaptées à leur domaine de compétences.
- **Objectifs :**
 - découvrir de nouveaux domaines d'application des biostatistiques
 - savoir implémenter des méthodes spécifiques à l'industrie et respectant les contraintes réglementaires et autres de ce secteur.
 -
- **Pré-requis nécessaires :**
 - Aucun
- **Responsable de l'UE : Robert SABATIER**
- **Intervenants:**
 - en cours

UE ETUDE DE CAS NIVEAU 2 : R SABATIER (2.5 ECTS)

- **Description :**

L'étude de cas se place dans le cadre de l'apprentissage par projet et a pour objectif de mettre l'étudiant dans une situation professionnalisante lui permettant de suivre un projet d'analyse de données de la planification de l'expérience au rendu des résultats interprétés.

La première partie de l'étude, en M1, a permis l'émergence d'une question en lien avec la biologie et la mise en place d'un plan de recueil des données qui a été suivi du recueil en lui-même.

La deuxième partie de cette étude consistera à la mise en œuvre du plan d'analyse et à l'organisation d'une communication de ces résultats auprès d'un public non spécialiste de biostatistiques.

- **Objectifs :**

- savoir choisir et mettre en œuvre les méthodes d'analyse de données adaptées à un type de données et à une question spécifique
- savoir prendre du recul sur les résultats afin de proposer d'éventuelles adaptations des données
- savoir rebondir sur les résultats pour proposer des approfondissements et de nouvelles perspectives de questions nécessitant le recueil de nouvelles données.
- savoir communiquer sa méthodologie et ses résultats à un public pluridisciplinaire non spécialiste

- **Pré-requis nécessaires :**

acquis de l'ensemble des UEs du Master

- **Responsable de l'UE : Robert SABATIER**

- **Intervenants:**

Christelle REYNES, UM, UFR des Sciences Pharmaceutiques
Robert SABATIER, UM, UFR des Sciences Pharmaceutiques

SEMESTRE 4

FORMATION INITIALE

UE STAGE : R SABATIER

- **Description :**

Le stage de Master 2 a pour objectif de s'immerger dans une structure qui utilise les statistiques pour la biologie (public/privé) pendant une durée de 6 mois. Comme dans le stage de M1, l'étudiant doit prendre en main une problématique biologique, réaliser une étude bibliographique autour du sujet, proposer et mettre en œuvre un plan d'analyse statistique pour répondre à la question de départ, dans le respect des pratiques sur le lieu de stage. Toutefois, en master 2, plusieurs questions biologiques ou des questions complexes peuvent être abordées. La comparaison entre différentes méthodologies sera abordée quand cela est possible. Une prise de recul sur la méthodologie et les résultats sera attendue.

- **Objectifs :**

- s'approprier une problématique biologique
- savoir réaliser et communiquer les résultats d'une étude bibliographique
- savoir argumenter le choix d'une méthode d'analyse
- savoir mettre en œuvre un plan d'analyse préalablement validé
- savoir réaliser la comparaison de méthodes
- savoir présenter et interpréter les résultats d'une étude statistique

- **Pré-requis nécessaires :**

acquis de l'ensemble des UEs du master

- **Responsable de l'UE : Robert SABATIER**

VERSION DE TRAVAIL