



Master mention Mathématiques, Biostatistique

Deux parcours :

Statistiques pour les Sciences de la Vie (SSV)

Statistiques pour les Sciences de la Santé (3S)

Université de Montpellier



Présentation des parcours

« Statistique pour les Sciences de la Vie »

Responsables : Jean-Noël Bacro et Robert Sabatier

Université de Montpellier

« Statistique pour les Sciences de la Santé »

Responsable : Paul Landais

Université de Montpellier et Faculté de Médecine Montpellier-Nîmes



Introduction

Le Public

- Pour le master Statistiques et Sciences de la vie (SSV) ce sont les étudiants de biologie (mécanisme du vivant, écologie, santé, agronomie, génétique,..)
- Pour le master Statistiques et Sciences de la Santé (SSS) ce sont les étudiants en Médecine, Pharmacie, Odontologie, Maïeutique ou Biologie.

Les débouchés

- Pour le master SSV, ils concernent plusieurs métiers : Biostatisticien en recherche en public ou en privé ; Biostatisticien dans une Contract Research Organization (CRO) ; Ingénieur statisticien dans un laboratoire de biologie ; Chargé de recherche dans un organisme de recherche ; Datamanagers dans un organisme de recherche ou un laboratoire de biologie ou une CRO ; Chercheur en biologie avec une forte composante en statistique.
- Pour le master SSS ils concernent les métiers suivants : Enseignant, chercheur, enseignant-chercheur dans l'enseignement supérieur, dans les organismes de recherche, ou en Santé Publique ; Chargé d'étude ou chargé de mission dans les institutions sanitaires et sociales ; Administrateur de santé publique ; Ingénieur de recherche ; Ingénieur d'étude ; Investigateurs ; Chefs de projets ; Assistants de recherche clinique ; Datamanagers ; ou Médecins et praticiens de Santé Publique.

Les enseignements

Ils sont présentés ci-après par unité d'enseignement, selon qu'ils sont partagés par les deux masters ou spécifiques de l'un ou de l'autre.

Le volume horaire

Le nombre d'ECTS à valider en M1 et en M2 est de 60. Chaque crédit est associé à un volume d'enseignement de 8 à 10 heures.

Les étudiants de licence doivent valider 60 ECTS en master 1.

UE de master 1 obligatoires pour les étudiants de la filière santé pour accéder au master 2 SSS

Les étudiants des filières santé valident 40 ECTS par leur cursus 2eme cycle ; ils doivent compléter par 10 ECTS (4 UE de 2.5 ECTS) et un stage d'initiation à la recherche dans un labo d'accueil validant (10 ECTS), pour atteindre les 60 ECTS requis pour le diplôme de Master 1.

Les 4 UE obligatoires sont :

- HMMA232 : Modèle linéaire général (2.5 ECTS)
- HMMA227 : Méthodes en épidémiologie quantitative 2 (2.5 ECTS)
- HMMA229 : Méthodes en recherche clinique : ECR 1 (2.5 ECTS)
- HMMA230 : Méthodes en recherche clinique : diagnostic – pronostic 1 1 (2.5 ECTS)

Aucune autre UE ne sera ouverte aux étudiants ; ils peuvent en revanche suivre les cours des autres UE en candidats libres.

Des équivalences pourront être données après examen du dossier du candidat par le comité pédagogique du master.

Les étudiants de 2eme cycle doivent valider 2 UE libres de 5 ECTS sur 2 ans donc les 4 UE de master valident ces UE libres.

Dans les autres parcours, ces UE libres sont validées avec une note $\geq 8/20$. Pour valider l'UE de master, il faut une note $\geq 10/20$. On adopte le même principe.

Pour les étudiants de second cycle à partir du DFGSM3, sur 2 ans :

- Modèle linéaire général et Méthodes en épidémiologie quantitative 2 la 1ere année
- Méthodes en recherche clinique : ECR 1 et Diagnostic-pronostic 1 la 2ème année
- Stage d'initiation à la recherche d'1 mois temps plein à l'issue dans un terrain de stage validant
 - BESPIM
 - EA2415
 - DIM Montpellier
 - CIC Montpellier
 - Unité INSERM Pathologies du vieillissement

(pour les 3 dernières structures, validation en cours)

Les étudiants qui veulent faire le master SSS seront prioritaires sur le choix de ces unités

Il n'y a pas d'équivalence de crédits pour le M2 et cela 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.

Les Unités d'enseignement

Master 1	7
Semestre 1	7
Unités d'enseignement (UE) mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé.	7
UE Mathématiques générales : F. Marche	7
UE Bases de données - Niveau 1 : M. Galindo / M. Vivien	7
UE Anglais : P. McNish	8
UE Analyse de données - Niveau 1 : R. Sabatier	8
UE Introduction à la statistique inférentielle : JN. Bacro	8
UE Aspects technologiques du recueil de données – omiques : C. Reynès	9
Introduction à la recherche clinique et épidémiologique : P. Fabbro-Peray	9
UE spécifiques parcours Statistiques pour les Sciences de la Vie.	10
UE Programmation : M. Huchard	10
UE spécifiques parcours Statistiques pour les Sciences de la Santé.	10
UE E-santé : P. Landais	10
UE Méthodes en épidémiologie quantitative Niveau 1 : T. Mura	10
Master 1	11
Semestre 2	11
UE mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé	11
UE communication : C. Reynès	11
UE Modèle linéaire général : JN Bacro	11
UE Programmation R et SAS : JN Bacro	11
UE Analyse en biologie des systèmes et e-données : J. Colinge	12
UE Etudes diagnostiques pronostiques niveau 1 : S. Bastide	12
UE Essais cliniques randomisés niveau 1 : JL. Faillie	13
UE Stage ou TER	13
UE spécifiques parcours Statistiques pour les Sciences de la Vie.	13
UE Omiques : AS. Fiston-Lavier / A. Mancheron	13
UE spécifiques parcours Statistiques pour les Sciences de la Santé.	13
UE Méthodes en épidémiologie quantitative Niveau 2 : P. Fabbro-Peray	13
UE Médico-économie niveau 1 : C. Castelli	14
Master 2	15
Semestre 3	15
UE mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé	15

UE Analyse statistique de données –omiques : C. Reynès	15
UE Séminaires de recherche P Fabbro-Peray, : C. Berr	15
UE Analyse des données censurées : P. Landais	15
UE SAS – Programmation niveau 2 : JN Bacro	16
UE Modèle linéaire généralisé et modèle mixte : I. Carrière	16
UE spécifiques du parcours Statistiques pour les Sciences de la Vie	17
UE Base de données – Niveau 2 : M. Galindo / M. Vivien	17
UE Recueils planifiés des données : plans d'expériences et d'échantillonnage : P. Letourmy	17
UE Analyse de données – Niveau 2 : R. Sabatier	17
UE Analyse de données – Niveau 3 : R. Sabatier	18
UE Maîtrise statistique des processus et contrôle qualité : P. Ravel	18
UE Statistique temporelle : JN Bacro	19
UE Statistique spatiale : N. Bez	19
UE spécifiques parcours Statistiques pour les Sciences de la Santé.	19
UE Méthodes en épidémiologie quantitative niveau 3 C. : Berr	20
UE Médico-économie niveau 2 : C. Castelli	20
UE Qualité de Vie : C. Gaujoux-Viala	20
UE Etudes diagnostiques et pronostiques niveau 2 : P. Landais	20
UE Méta-analyses : JP. Daurès	21
UE Essais cliniques randomisés niveau 2 : S. Thézenas	21
UE Méthodes en Imagerie quantitative mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé	21
UE Constitution du signal en imagerie et systèmes bruités (1) : JP. Bérégi	21
UE Méthodes de diffusion et approches statistiques en Imagerie (1) : JP. Bérégi	22
UE Méthodes en Imagerie quantitative neurologique/oncologique : JP. Bérégi	22
UE Méthodes en Imagerie quantitative cardiovasculaire : JP. Bérégi	22
Master 2	23
Semestre 4	23
UE des parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé	23
UE Stage	23

Programme

Statistiques et Sciences de la Vie & Statistiques et Sciences de la santé

Master 1

Semestre 1

Unités d'enseignement (UE) mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé.

UE Mathématiques générales : F. Marche

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 des deux parcours master « Biostatistique en sciences de la vie » et « Biostatistique en santé ».

Le cours revient sous les notions fondamentales de mathématiques que les étudiants ont rencontrées au cours de leur cursus de licence, en les motivant et les exploitant par les contextes d'utilisation en statistique (méthodes d'analyse de données, inférence statistique etc). Le cours se focalise sur les fondamentaux suivants :

- algèbre linéaire (18h) : 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 (6h) : dérivée de fonctions réelles et dérivées partielles de fonctions multivariées ; recherche d'extrema de fonctions différentiables ; développements limités de fonctions différentiables.

UE Bases de données - Niveau 1 : M. Galindo / M. Vivien

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. 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. Ce logiciel est un outil pédagogique d'intérêt pour apprendre le data-management et son choix se justifie pour les raisons suivantes :

- surcouche graphique très efficace et facilement appréhendable permettant de faire du SQL de façon simple ;
- boîte à outils simple et pratique puisqu'en plus des tables et des requêtes (la base du modèle relationnel), Access supporte quasiment tous les modèles d'import-export de données, permet de réaliser des masques de saisie, peut réaliser des modèles de rapports propres et très professionnels (sans perdre trop de temps à les mettre en forme) ;

- Access est programmable et automatisable, ce qui est un atout non négligeable pour approfondir de façon simple les approches data-management.

Le cours du master 1 sera donc axé sur l'utilisation du logiciel Access pour introduire les notions fondamentales du 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.

UE Anglais : P. McNish

L'utilisation de la langue anglaise est devenue incontournable dans le domaine scientifique. L'objectif de cet enseignement est de renforcer les connaissances des étudiants dans le domaine de l'anglais scientifique et technique.

UE Analyse de données - Niveau 1 : R. Sabatier

L'objectif de ce premier enseignement d'Analyse de Données est l'acquisition des méthodes multivariées les plus classiques tant pour l'apprentissage supervisé que 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.

Les méthodes enseignées seront les suivantes :

- Analyses descriptives : Analyse en Composantes Principales (ACP) et Analyse Factorielle des Correspondances (AFC)
- Apprentissage supervisé : Analyse Linéaire Discriminante (LDA)
- Apprentissage non supervisé : k-means ; Classification Hiérarchique Ascendante.

Chacune des méthodes enseignées en cours seront appliquées 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.

UE Introduction à la statistique inférentielle : JN. Bacro

L'objectif du cours est de présenter les notions fondamentales et les techniques usuelles de la statistique inférentielle en les situant dans un contexte d'utilisation pratique. Le M1 étant ouvert aux étudiants venant de parcours licence biologie/santé, les compétences en statistique de chacun peuvent être assez hétérogènes et l'un des buts du cours est de les homogénéiser tout en renforçant le socle commun de connaissances. A titre d'exemples, une partie du cours sur l'estimation concerne des rappels (estimation empirique de la moyenne, d'une proportion, d'une variance) mais des techniques qui peuvent être nouvelles pour certains cursus sont aussi présentées et doivent être assimilées (méthode des moments, maximum de vraisemblance). De même pour les tests statistiques ; le fil directeur du cours est donc de resituer dans leur contexte des approches et des résultats éventuellement connus tout en présentant des généralisations ou des extensions permettant de mieux appréhender les méthodes et de s'approprier les techniques.

L'accent est donc mis sur le raisonnement statistique et le cours est axé sur la compréhension et la mise en œuvre correcte de méthodes d'inférence statistique. Les notions présentées sont illustrées au travers de nombreux exemples et exercices. Les exercices sont aussi l'occasion d'une première prise en main du logiciel R, tant pour les tests usuels que pour la simulation de réalisations de variables aléatoires (construction d'échantillons simulés). Après un bref rappel de probabilités mettant en relief les différents outils offerts par les approches probabilistes, le cours présente les méthodes classiques d'estimation et focalise sur la notion de test statistique. Les approches paramétriques et non paramétriques pour un et deux échantillons sont successivement présentées.

UE Aspects technologiques du recueil de données – omiques : C. Reynès

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.

L'enseignement sera composé de visites de plateformes, d'initiation aux principaux logiciels d'acquisition des données – omiques, en privilégiant les logiciels « plate-forme indépendants » tels les outils proposés sur Bioconductor. L'essentiel de l'enseignement sera donc pratique et devra permettre aux étudiants de posséder un certain nombre d'outils et de scripts pour être capables d'une certaine autonomie vis-à-vis d'un tel jeu de données.

Il est important de noter que cette UE a uniquement pour vocation d'initier les étudiants à l'acquisition et au pré-traitement des données – omiques. Leur analyse statistique poussée sera abordée dans une autre UE.

Introduction à la recherche clinique et épidémiologique : P. Fabbro-Peray

Les objectifs pédagogiques concernent l'appropriation des grands concepts qui sont déclinés dans le domaine de la recherche clinique et épidémiologique.

Présentation des principaux champs de l'épidémiologie et de la recherche clinique, les objectifs descriptifs et les objectifs analytiques des études, les grands types d'enquêtes observationnelles et expérimentales ; points clefs des études expérimentales (randomisation, aveugle, intention de traiter) ; intégrer les concepts d'association statistique et de causalité, l'étude d'une association entre une maladie et un facteur de risque (enquêtes transversales, enquêtes de cohorte, enquêtes cas-témoins) ; les différentes mesures de fréquence d'une maladie ou d'un phénomène de santé (prévalence, taux d'incidence, risque) ; les mesures d'association entre une maladie et un facteur de risque (différence de risque, risque relatif, odds ratio), la notion de fluctuation d'échantillonnage, les principaux types de biais, biais de sélection, biais de classification et facteurs de confusion ; connaître les principaux chapitres d'un protocole.

UE spécifiques parcours Statistiques pour les Sciences de la Vie.

UE Programmation : M. Huchard

UE spécifiques parcours Statistiques pour les Sciences de la Santé.

UE E-santé : P. Landais

L'objectif de cette UE est d'introduire les grands concepts et méthodes utilisés en e-Santé, les principes de collecte, de gestion et d'échange de données numérisées. Introduction à la nomenclature et aux classifications médicales, notion d'ontologies ; Banques d'information et intégration de données ; Bases de données en santé ; Systèmes d'information hospitaliers ; Dossier Patient Informatisé et recherche clinique, approches CDISC, HL7 et BRIDG ; « Case Report Form » électronique, introduction aux principaux logiciels d'acquisition de données en recherche épidémiologique et clinique ; Dossier médical et réseaux de soins ; Interopérabilité en santé ; Objets connectés en santé ; Aspects juridiques et éthiques pour le traitement des données ; Tarification à l'activité ; Sécurisation de la prescription médicamenteuse ; Introduction à la bioinformatique.

UE Méthodes en épidémiologie quantitative Niveau 1 : T. Mura

L'objectif de cette UE est de fournir aux étudiants des connaissances théoriques approfondies sur la nature, les propriétés et les méthodes d'estimation des principaux indicateurs et mesures utilisés dans les enquêtes épidémiologiques. L'UE abordera notamment les aspects statistiques des méthodes de sondage, les propriétés des indicateurs de santé (incidence, prévalence...) et les méthodes de standardisation, les biais, les interactions, l'analyse univariée des mesures de risque et d'association, l'ajustement et les principes généraux des analyses multivariées, la notion de puissance et le principe du calcul du nombre de sujets nécessaires. La seconde partie de cette UE abordera de façon théorique les différents éléments constitutifs du contenu d'un protocole de recherche en épidémiologie et l'introduction à la lecture critique d'article fondée sur la grille STROBE (STrengthening the Reporting of OBservational studies in Epidemiology). L'enseignement s'attachera à fournir pour chacun de ces éléments 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.

D'une façon globale, l'EU « Méthodes en épidémiologie quantitative Niveau 1 » fournira aux étudiants les connaissances théoriques nécessaires à la compréhension approfondie et à la conceptualisation d'une recherche épidémiologique de qualité.

Master 1

Semestre 2

UE mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé

UE communication : C. Reynès

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 Modèle linéaire général : JN Bacro

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).

Le plan du cours est le suivant : Qu'est-ce qu'un modèle linéaire ; Ecriture matricielle $Y=Xq+E$, distributions de Y et de E (résiduelle) ; Ecriture complète d'exemples concrets (régressions simple et multiple, ANOVA à 1 facteur, 2 ou plus facteurs croisés) ; Estimation par moindres carrés ; maximum de vraisemblance (exemple d'un échantillon gaussien) ; Estimation de q et distribution de l'estimateur (explicitée sur deux exemples de l'introduction) ; Estimation de la variance de E (introduction de la loi de chi 2) ; Intervalles de confiance (introduction de la loi de Student) ; Tests d'hypothèses multiples : Présentation géométrique ; Cas des différents types de somme dans SAS ; Tests. Introduction de la loi de Fisher.

Tous les points précédents sont illustrés de traitement de données (régressions quadratiques, régressions quadratiques combinées avec une variable qualitative). Les sorties « standard » offertes par SAS sont « disséquées » (tableau d'analyse de variance, sommes de carrés de type I et III, tableau des estimations et leurs écarts types et tests t correspondants, etc).

UE Programmation R et SAS : JN Bacro

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 (15h), le logiciel R est présenté : les concepts de base sont introduits : lecture et sauvegarde de données ; manipulations, constructions et opérations sur les vecteurs ; utilisation des fonctions statistiques de base internes ; sorties graphiques élémentaires ; simulation de variables aléatoires ; introduction aux fonctions. Les étudiants travailleront les notions sous forme d'exercices à traiter avec R et le

lien avec les UE de statistique enseignées en parallèle et utilisant R (UE Statistique de base, Analyse des données) sera fait.

La seconde partie de cours (9h) est consacrée à une introduction au logiciel statistique SAS. Trois séances de 3h permettront une première prise en main du logiciel : structure des programmes SAS : étapes DATA et PROC, fichiers saisie, log et output ; lecture de données selon les principaux formats rencontrés en pratique ; données continues et discrètes (MEANS, FREQ) ; calculs statistiques de base sur un échantillon (FREQ, UNIVARIATE) et procédures SAS pour les tests usuels sur un échantillon (NPAR1WAY ; TTEST) ; approches régression et anova (REG, ANOVA).

UE Analyse en biologie des systèmes et e-données : J. Colinge

Introduction des concepts clefs de la biologie des systèmes et illustration de leur application par des exemples en médecine et santé. Définition du champ de la biologie des systèmes puis d'un élément central de sa méthodologie : la notion de réseau d'interaction et sa modélisation en tant que graphe. Des exemples communs de réseaux biologiques avec les procédures expérimentales permettant de les connaître seront évoqués.

Forts de ces notions, nous étudierons la topologie des réseaux biologiques classiques, p.ex. interaction physique de protéines (PPI). Nous mettrons en avant leur redondance et discuterons certaines implications considérant l'hétérogénéité inter-individuelle des mutations tumorales, la stratification des tumeurs et des patients, la létalité synthétique. Nous verrons aussi des opportunités en termes de puissance statistique et cohortes de petite taille.

Un second grand type d'analyse consistera en l'intégration de jeux de données sur un réseau biologique, qui a généralement la propriété d'associer gènes ou protéines participant à une même fonction biologique. Ceci peut faciliter l'identification de processus biologiques régulés parmi des données complexes ou de grande taille. Notre exemple modèle s'appliquera à prédire les effets secondaires d'inhibiteurs de kinases utilisés en cancérologie.

Le cours se terminera par la modélisation quantitative et dynamique de réactions enzymatiques ou de l'activité de facteurs de transcription. Nous verrons que certaines structures types apparaissent de manière récurrente dans les réseaux biologiques et qu'elles ont des propriétés bien particulières. Nous montrerons comment des modèles de ce type permettent de simuler des traitements en cancérologie afin de trouver des solutions optimales.

Les techniques utilisées dans les exemples du cours seront applicables à maints autres problèmes. Des outils et ressources publics seront présentés (Cytoscape, librairies R). Cette UE est articulée autour de 3 chapitres : les données, structures et annotations ; la modélisation ; la notion d'hétérogénéité des systèmes.

UE Etudes diagnostiques pronostiques niveau 1 : S. Bastide

L'objectif est de présenter les protocoles d'étude permettant d'évaluer les méthodes diagnostiques. Estimation de la sensibilité, de la spécificité et des ratios de vraisemblance d'un test diagnostique. Valeurs prédictives positives et négatives dans la population à laquelle est appliqué le test. Méthodes d'estimation et interprétation d'une courbe ROC. Probabilité pré-test de la maladie, odds de la maladie. Odds post-test d'une affection en fonction du résultat du test et passage à la probabilité post-test. Place des tests diagnostiques dans le contexte de l'analyse de la décision médicale : introduction des notions d'utilité. Détermination d'un seuil décisionnel. Méthodes de choix et d'estimation du seuil de positivité d'un test diagnostique maximisant l'utilité du test. Construction d'un arbre de décision portant sur le choix entre deux méthodes diagnostiques. Lecture critique d'articles et utilisation de la grille STARD.

UE Essais cliniques randomisés niveau 1 : JL. Faillie

Cette UE est destinée à évoquer les buts et principes généraux des essais thérapeutiques ; différentes phases d'évaluation du médicament ; formulation des hypothèses ; définition des traitements, des malades et des critères de jugement ; randomisation (méthodes et réalisation), contrôle, aveugle, placebo.

Le nombre de sujets nécessaire et calcul de puissance pour un essai d'efficacité sur deux groupes parallèles. Techniques d'ajustement, interaction. Planifications autres que deux groupes parallèles : principes, intérêts, limites ; en particulier, plan factoriel 2x2 : design, intérêt, analyse. Ecriture d'un protocole, conduite de l'essai, ICH guidelines (The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use). Analyse et discussion des résultats ; analyses intermédiaires, analyses de sous-groupes, Aspects éthiques et réglementaires ; Introduction au CONSORT (Consolidated Standards of Reporting Trials) et au SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), lecture critique d'article. L'exemple d'un projet de recherche sera décliné avec les étapes du data management : depuis l'attribution d'un projet jusqu'à l'archivage des données. Ces étapes concernent : la création d'un Clinical Report Form (CRF), la création d'un e-CRF, la création d'un cahier des charges de data management, le contrôle de qualité à la saisie, l'importation de données externes, le contrôle de cohérence des requêtes, la gestion des corrections de l'e-CRF, la réconciliation des événements indésirables graves avec la base de pharmacovigilance, le contrôle de qualité de la base de données, la préparation de la revue des données, le gel de base, l'export des données destinées à l'analyse, et enfin l'archivage de tous les documents afférents. Ces différentes étapes seront déclinées en utilisant un logiciel de gestion de données, le logiciel RedCap, mis en open access par l'université Vanderbilt à Nashville, USA.

UE Stage ou TER

UE spécifiques parcours Statistiques pour les Sciences de la Vie.

UE Omiques : AS. Fiston-Lavier / A. Mancheron

UE Interventions de professionnels

UE spécifiques parcours Statistiques pour les Sciences de la Santé.

UE Méthodes en épidémiologie quantitative Niveau 2 : P. Fabbro-Peray

L'UE « Méthodes en épidémiologie quantitative Niveau 2 » correspond à la mise en pratique des connaissances théoriques acquises dans l'UE « Niveau 1 ». Cette UE doit permettre aux étudiants d'acquérir des connaissances approfondies sur les principes et les enjeux des étapes de la planification des différents types d'études épidémiologiques (transversale, cas-témoin, cohorte). Les étudiants devront, à terme, savoir appliquer ces principes dans la construction d'un protocole, le calcul du nombre de sujets nécessaire et évaluer leur pertinence dans le cadre d'une lecture critique d'article, notamment à l'aide de l'utilisation d'une grille de lecture d'article épidémiologique (Grilles STROBE). L'UE abordera également les différentes sources

de données existantes en épidémiologie (registres, PMSI, SNIIRAM, CEPIDC...), leurs modes de fonctionnement et leurs limites. L'enseignement sera réalisé, en grande partie, à l'aide d'exercices pratiques portant sur la construction d'un protocole de recherche épidémiologique et sur la lecture critique d'article.

UE Médico-économie niveau 1 : C. Castelli

L'objectif de cette UE est d'introduire les principaux types d'études médico-économiques utilisés en recherche clinique. Principaux concepts utiles en médico-économie : le coût, les principaux types de cout, méthodologie de recueil des couts, notion d'utilité, et théorie des préférences (QALY, DALY), l'étude du lien entre la qualité de vie et l'utilité : utilité individuelle/sociétale. Introduction au système de soins et aux principaux opérateurs. Plusieurs schémas d'études sont présentés. Il s'agit des études de minimisation des coûts, les études cout-efficacité, cout utilité et cout conséquences. Les études de micro-costing sont également abordées. Elles consistent à évaluer le coût « réel » d'un acte par opposition au tarif. Ce type d'étude est particulièrement utile dans le cas d'une innovation dont le coût est, par définition, inconnu et qui ne dispose donc pas encore de tarification. Les analyses d'impact budgétaire sont aussi introduites. Elles aident à estimer l'incidence financière sur le budget d'un acteur du système de santé de l'introduction du retrait ou de la modification de stratégies médicales liées à une innovation technique ou à une modification règlementaire. Elles peuvent être réalisées *ex-ante* et dans ce cas elles peuvent être vues comme un outil d'aide à la décision d'achat d'un dispositif médical au sein d'un établissement de santé. A ce titre, ces études font partie intégrante des études dites de « Health Technology Assessment » (HTA). Ces analyses peuvent être également réalisées *ex-post* après validation de l'efficacité. Ce type d'étude est alors un outil d'aide à la décision de remboursement et diffusion. Ces différents concepts et les techniques/outils qui leur sont associés sont présentés. Lecture critique d'article et introduction au CHEERS (Consolidated Health Economic Evaluation Reporting Standards) statement.

Master 2

Semestre 3

UE mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé

UE Analyse statistique de données –omiques : C. Reynès

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. Par exemple, la recherche univariée de gènes différentiellement exprimés fait appel à des modèles linéaires classiques couplés à des tests de comparaisons multiples. D'autre part, la recherche d'une signature génétique d'une réponse à un traitement s'apparente à un problème de sélection de variables. Cette UE fera donc appel à des méthodes abordées dans différentes autres UEs et leur donnera une application concrète dans le contexte des données –omics.

D'un point de vue pédagogique, les cours permettront de donner les éléments théoriques nécessaires à l'adaptation de méthodes classiques au contexte particulier des données –omics (contexte qui aura été explicité dans l'UE « Aspects technologiques du recueil de données –omics »). Ensuite, les méthodes seront appliquées lors de TD utilisant le logiciel R et des packages issus de Bioconductor. 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.

UE Séminaires de recherche P Fabbro-Peray, : C. Berr

Des professionnels de différents horizons viennent présenter sous forme de conférences leur travail actuel, les enjeux qui s'y rattachent, les orientations et les développements à venir. Ces interventions concernent par exemple la biostatistique en milieu pharmaceutique (SANOFI), l'analyse sensorielle et les enquêtes sur panel (INRA), le data-management (INCA).

La rédaction de CV et la préparation d'un entretien d'embauche sont aussi abordées (N. Camus, consultante en RH).

UE Analyse des données censurées : P. Landais

L'objectif de ce cours est de présenter les notions de base sur les données de survie et de détailler les modèles les plus communément utilisés pour leur analyse. L'analyse des données censurées a pris une place majeure en santé publique et en médecine. La problématique de la survie sera présentée ainsi que sa place en médico-économie. Les notions de censure à droite, à gauche et par intervalle seront présentées, puis les notions de censure non informative et de censure informative. La notion de risque instantané de décès brut et net sera introduite. L'analyse de la survie proprement dite sera abordée, survie non paramétrique et survie paramétrique (lois de Weibull, Weibull Généralisé, ..), survie semi paramétrique et comparaison de survies. L'introduction de covariables dans les cas paramétriques et semi paramétrique, en particulier le

modèle de Cox, sera ensuite abordée. Notion de survie nette et de survie relative. Le lien de 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. La mise en œuvre des programmes avec R sera présentée sous forme de programmes et de sorties commentées.

UE SAS – Programmation niveau 2 : JN Bacro

L'UE SAS – Niveau fait suite à l'UE « **Programmation R et SAS** » du M1 et a pour objectif d'approfondir la maîtrise du logiciel SAS, logiciel référence dans le domaine de la biostatistique. La construction des tables de données SAS sera complétée : la compréhension fine de la façon dont sont les tables de données sont créées, formatées, modifiées, gérées ou exploitées, tables sur lesquelles les procédures SAS travaillent ensuite, est fondamental pour une utilisation optimale et pertinente des possibilités offertes par le logiciel. Les procédures de base vues en M1 seront complétées par des procédures spécifiques aux modèles linéaires à effets fixes et effets aléatoires (GLM, MIXED) et les diagnostics de modélisation (validation croisée, graphiques). La programmation en langage macro SAS, permettant par exemple de construire des fonctions des utilisations récurrentes, sera aussi étudiée et mise en pratique sur des exemples concrets. Les liens possibles sous SAS avec le data-management (masques de saisie en programmation macro SAS) seront aussi étudiés.

UE Modèle linéaire généralisé et modèle mixte : I. Carrière

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

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 dans ce contexte d'application ainsi que la gestion des notions épidémiologiques de facteurs de confusion et/ou modificateurs d'effets. 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.

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 ainsi que la catégorisation et les méthodes d'imputations des données manquantes. 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 est abordée de façon détaillée sous forme de programmes et sorties commentées.

UE spécifiques du parcours Statistiques pour les Sciences de la Vie

UE Base de données – Niveau 2 : M. Galindo / M. Vivien

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.

UE Recueils planifiés des données : plans d'expériences et d'échantillonnage : P. Letourmy

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 « Plans d'expériences » 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 confusion 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.

UE Analyse de données – Niveau 2 : R. Sabatier

Ce second module d'analyse de données 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.

Nous présenterons des méthodes non linéaires à noyau tel SVM (Support Vector Machine) ou méthodes de type « arbre » tels CART ou les Forêts Aléatoires. Des extensions de la régression linéaire multiple seront également abordés : on peut citer PLS (Partial Least Squares) ou encore les régressions pénalisées. Des notions sur l'analyse de données multi-tableaux (type STATIS) seront également présentées.

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 Analyse de Données – Niveau 1 : 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 le logiciel R. 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.

UE Analyse de données – Niveau 3 : R. Sabatier

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 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 le logiciel R. 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.

UE Maîtrise statistique des processus et contrôle qualité : P. Ravel

La Maîtrise Statistique des Procédés fait partie des outils du Management de la Qualité. Cette approche statistique puissante est utilisée depuis les années 1930 dans le cadre du suivi de procédés industriels. Fort de son succès, cette méthode s'est généralisée dans l'évaluation du contrôle de services depuis plus d'une vingtaine d'années.

L'objectif du cours est l'utilisation de la maîtrise statistique des procédés MSP/SPC au sens large (production ou service) dans le cadre du suivi au cours du temps de ces procédés.

La maîtrise Statistique des procédés est l'ensemble des méthodes statistiques permettant d'évaluer de façon mesurables les caractéristiques d'un procédé de production ou d'un service, afin de les maintenir sous contrôle par rapport aux spécifications retenues et aussi d'améliorer les performances du procédé ou du service étudié.

Les thèmes abordés au cours de cette formation seront les suivants : Identification des paramètres du processus à maîtriser ; Diagramme des 5M ; Les fluctuations et l'échantillonnage ; Cartes de Contrôle par attributs NF X06-31-2 ; Cartes de contrôle pour variable quantitative NF X06-31-1 ; Capacité d'un Procédé, indices de performances CP, Cap, Cpk ; Carte de contrôle de type Shewhart, intégration des limites de spécifications clients, NF X06-31-1 ; Gestion des alarmes sur les risques cartes de contrôle : Risques clients/ Fournisseurs, période opérationnelle moyenne ; Cartes de contrôle pour détecter les dérives lentes et progressives : Carte EWMEA NF X06-31-3, Carte CUSUM, NF X06-31-4 ; Cartes de contrôle pour les problèmes à plusieurs dimensions : Carte multi Variée T2-Hotelling ; Carte Arima .

UE Statistique temporelle : JN Bacro

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,..) 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.

UE Statistique spatiale : N. Bez

Les données spatialisées sont de plus en plus utilisées en pratique et concernent aujourd'hui de nombreux et très variés domaines d'application. L'objectif est ici d'initier les étudiants au caractère particulier des données spatialisées et à la nécessité d'utiliser des approches appropriées pour le traitement de celles-ci afin d'éviter toute utilisation erronée de méthodes adaptées au cadre indépendant. Dans un premier temps, une présentation générale des méthodes est proposée (données sur lattice, données sur espace continu, notions de corrélation et d'hétérogénéité spatiales, ...) et des indices généraux de caractérisation de la dépendance spatiale sont présentés (Moran, Geary). L'exemple typique de la recherche, et de la caractérisation et de la prise en compte de dépendances spatiales dans des résidus de régression est étudié et traité de façon approfondie. Dans une seconde partie, les problèmes de prédiction spatiale et de construction de cartes sont abordés. Des méthodes déterministes (tessellation, distance inverse) et géostatistique (krigeage) sont étudiées et comparées. Les notions de fonction de covariance spatiale et de variogramme sont présentées et explicitées pour une bonne utilisation en pratique. Des exemples traités avec le logiciel R, dont des packages dédiés à la modélisation spatiale et au traitement des données spatiales, illustrent les techniques vues en cours.

UE spécifiques parcours Statistiques pour les Sciences de la Santé.

UE Méthodes en épidémiologie quantitative niveau 3 C. : Berr

Cette unité d'enseignement est centrée sur l'approfondissement de mesures et de méthodes orientées sur des sujets spécifiques.

La mesure : elle concerne l'exposition professionnelle, 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, nouvelles approches de la causalité (approches contrefactuelles, graphes acycliques (Directed Acyclic Graphs DAG), scores de propension,...), designs particuliers (cas-témoins nichés dans les cohortes, cas-contrôles) ; corrélations écologiques, clusters, introduction aux analyses spatio-temporelles. Application des méthodes et des concepts pour le traitement des grandes bases des données épidémiologiques ; Particularités méthodologiques des études pharmaco-épidémiologiques.

UE Médico-économie niveau 2 : C. Castelli

L'objectif de cette UE est de présenter certaines méthodes avancées en médico-économie appliquées à la recherche hospitalière. Plusieurs types d'approches seront abordés dans ce cadre : l'utilisation de variables auxiliaires dans l'estimation des coûts moyens, la prise en compte des données de coûts absentes ou tronquées par les techniques d'imputation, la prise en compte de la censure informative dans l'estimation du coût moyen, l'utilisation des scores de propension dans les analyses médico-économiques, l'approche des méta analyses de données médico-économiques, l'application de la théorie de la loterie et de l'utilité dans les événements courts de type diagnostic, et enfin l'utilisation des Processus Markoviens en médico-économie.

UE Qualité de Vie : C. Gaujoux-Viala

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 ? Différents types d'instruments et 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 et sensibilité. 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. Le lien avec la théorie de la performance sera introduit. L'analyse factorielle et la modélisation par des équations structurelles sera également présentée ainsi que la structure des réponses aux items. Enfin, l'utilisation de la qualité de vie en pratique médicale courante (essais thérapeutiques) et la place dans le processus de décision seront exposés.

UE Etudes diagnostiques et pronostiques niveau 2 : P. Landais

Prédiction : Introduction, aspects méthodologiques généraux ; Méthodologie des études pour la prédiction. Les protocoles d'études pour les modèles de prédiction en recherche clinique ; prédicteur et variable de jugement, concordance des prédicteurs, variable de jugement, puissance statistique.

Pronostic : Méthodologie des études pronostiques ; applications des méthodes d'analyse de Survie, comparaison de survie ; analyse hiérarchique ;

Le cours des événements de santé est abordé dans le contexte de la nature et la qualité des soins de santé prodigués (recherche pronostique fondamentale) ; Etude de certains facteurs (biomarqueurs) qui sont associés au pronostic (recherche de facteurs pronostiques) ; Le développement, la validation et l'impact des modèles statistiques qui permettent de prédire un risque individuel d'un futur résultat (recherche de modèle pronostique) ; L'utilisation de

l'information pronostique pour aider à adapter les décisions de traitement à une personne ou un groupe d'individus ayant des caractéristiques semblables (recherche stratifiée).
Lecture critique d'article fondée sur l'approche PROGRESS (Prognosis Research Strategy) et sur les recommandations du TRIPOD.

UE Méta-analyses : JP. Daurès

La méta-analyse a pris une place importante dans la littérature médicale en matière d'évaluation. Les notions essentielles suivantes seront introduites : définition, objectifs, variables principales, analyse de la littérature, estimation de la qualité ; Analyse des biais ; Synthèse des résultats ; Mesure des résultats en fonction du type de variable : binaire, continue, Odds Ratio, Risque Relatif, temps dépendante ; Fonction du type d'étude : Cas-témoin, cohorte, prévalence, survie, appariée, courbe ROC, ... ; Analyse de l'hétérogénéité ; Méthodes fixes ou aléatoires ; Analyse du biais de publication ; Applications et discussion des résultats. Lecture critique d'une méta-analyse fondée sur le Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA Statement.

UE Essais cliniques randomisés niveau 2 : S. Thézenas

Cette unité d'enseignement vise à introduire des méthodes récentes utilisées dans certains types d'essais thérapeutiques. Seront abordés, le principe de la randomisation adaptative, les essais en cluster, le Step-wedge design ; les analyses de futilité ; les études de phase I : principe, réalisation et analyse ; essais d'équivalence et de non-infériorité ; les développements méthodologiques dans le domaine des essais cliniques fondés sur une évaluation économique. Seront également abordées les études non pharmacologiques, non médicamenteuses comme les essais en chirurgie, les dispositifs médicaux, les interventions éducatives, les interventions non médicamenteuses. Ces approches seront illustrées par des exemples tirés de la littérature internationale. La lecture d'articles suivra les principes du CONSORT.

UE Méthodes en Imagerie quantitative mutualisées entre les parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé

Ces unités d'enseignement sont réservées aux étudiants prenant une orientation vers l'imagerie quantitative. Le choix concerne alors l'ensemble de ces 4 UEs pour les étudiants de *Statistiques et Sciences de la Santé*. Certaines UEs peuvent être choisies par les étudiants de *Statistiques et Sciences de la Vie* après avis favorable du responsable du parcours.

Les principaux objectifs de ces UEs concernent la méthodologie dans les domaines suivants : le signal en imagerie, la quantification morphologique et fonctionnelle, la perfusion, la viabilité, l'IRM de diffusion, l'approche par Diffusion Tensor Imaging (DTI) ; enfin, l'intérêt des logiciels de traitement et l'usage du Cloud.

UE Constitution du signal en imagerie et systèmes bruités (1) : JP. Bérégi

- A- Bases physiques du signal en imagerie ; Méthodes d'obtention des images (acquisition/inj/reco) en CT, MR, Echo, PET, SPECT, Rx, Angio, ... ; Les reconstructions axiales, tridimensionnelles, multiplanaires, ...
- B- K Space lecture Plan de Fourier ; Sinogramme ; Format compression DICOM – Analyse – Nifti ; Logiciels de programmation : Matlab, ... ; Archivage / Télétransmission / Télé Diagnostic

UE Méthodes de diffusion et approches statistiques en Imagerie (1) : JP. Bérégi

A- IRM de diffusion ; Diffusion Weighted Images (DWI) ; Diffusion Tenseur Images ; Diffusion Spectrum Images

B- Logiciel Iktimap ; Décrire Méthodes : segmentation/détection de contours quantification dans 1 image ; Filtrage ; Vessel tracking ; Fusion d'image ; Reconstruction itérative

UE Méthodes en Imagerie quantitative neurologique/oncologique : JP. Bérégi

A- /Fiber tracking ; DfMRI : IRM fonctionnelle ; Perfusion cérébrale

B- Oncologie : Quantitatif Morphologie/ Recist/ ; Quantitatif fonctionnel ; Diffusion du corps entier Neurologie : Cartographie du cerveau ; Diffusion Tensor Tractography (DTT) neurologique; Cas cliniques

UE Méthodes en Imagerie quantitative cardiovasculaire : JP. Bérégi

A- Quantitatif sténose ; Quantitatif anévrisme / Dissection aortique ; Quantitatif d'un infarctus du myocarde (perf/viab) ; Quantitatif de la fonction VG - Couplage ventriculo-artériel ; Quantitatif de la structure ; Rythmologie – Electrophysiologie

B- Carto pré-interventionnelle ; Imagerie de guidage (cancéro) ; Simulation/Imagerie ; Evaluation post opératoire.

Master 2

Semestre 4

UE des parcours Statistiques pour les Sciences de la Vie & Statistiques pour les Sciences de la Santé

UE Stage

Le stage est obligatoire et d'une durée de 4 à 6 mois selon les parcours.

Un mémoire est rédigé et une soutenance devant jury est programmée à l'issue du stage.