

Cristian Preda

## **Curriculum vitae détaillé**

(Janvier 2022)

Activités de recherche, enseignements et collaborations scientifiques



## Table de matières

|   |      |
|---|------|
| <b>Situation actuelle et parcours académique</b>          | p.3  |
| <b>Activité de recherche</b>                              |      |
| Thématiques   | p.3  |
| Publications, conférences                                 | p.4  |
| Encadrement doctoral                                      | p.5  |
| Projets/contrats de recherche                             | p.6  |
| <b>Responsabilités scientifiques</b>                      |      |
| Direction, animation laboratoires et équipes de recherche | p.7  |
| Activités éditoriales                                     | p.7  |
| Organisation de conférences/colloques                     | p.7  |
| Rapporteur/Examineur (doctorat/HDR)                       | p.7  |
| Autres (Expert, Reviewer, sociétés savantes)              | p.7  |
| <b>Responsabilités pédagogiques et administratives</b>    |      |
| Responsabilités pédagogiques (cours, stages)              | p.8  |
| CNU, Conseil de l'Université, Recrutement                 | p.8  |
| <b>Activités d'enseignements</b>                          |      |
| Faculté de Médecine de Lille                              | p.9  |
| Polytech'Lille/Université Lille 1                         | p.9  |
| Biostatistique (Polytech'Lille – CHRU Lille)              | p.10 |
| <b>Collaborations scientifiques et pédagogiques</b>       |      |
| Faculté de médecine de Lille                              | p.11 |
| Unité de Méthodologie-Biostatistique de CHRU de Lille     | p.12 |
| GHICL – Département de Recherche Médicale                 | p.12 |
| INRIA – Equipe Modal                                      | p.13 |
| Mobilité internationale                                   | p.13 |
| <b>Annexe</b>   |      |
| Publications  | p.14 |
| Conférences   | p.17 |



## Cristian PREDA

Professeur des Universités  
Université des Sciences et Technologies de Lille (Lille 1)  
Ecole Polytechnique Universitaire de Lille  
Laboratoire de Mathématiques Paul Painlevé, UMR CNRS 8524

e-mail : cristian.preda@univ-lille.fr

### Situation actuelle et parcours

| Années             | Fonction  | Etablissement   |
|--------------------|---|---|
| 09.2008 – présent  | Professeur (1cl.)<br>(CNU 26 – Stat./Math. Appl.) | Ecole Polytechnique Universitaire,<br>Laboratoire de Mathématiques Paul<br>Painlevé, UMR CNRS 8524<br>Université de Lille 1 |
| 09.2000 – 08. 2008 | Maitre de Conférences                             | Faculté de Médecine<br>Université Lille 2   |
| 02.2000 – 08.2000  | Assistant chercheur                               | Université Libre de Bruxelles   |
| 02.1997 – 08.1999  | Doctorant   | Université Lille 1  |
| 02.1997 – 08.1999  | Assistant Chercheur                               | Institut de Statistique de<br>l'Académie Roumaine   |

## I. Activités de recherche

### 1.1 Thématiques

— Modèles pour données fonctionnelles (analyse de courbes): Données fonctionnelles multivariées et données fonctionnelles qualitatives : classification non-supervisée (modèles de mélange), modèles de régression linéaire avec réponse fonctionnelle (PLS, PCR), modèles non-paramétriques et non-linéaires à l'aide des techniques basées sur les propriétés des espaces de Hilbert à noyau reproduisant (RKHS).


— Biostatistiques : Modèles hiérarchiques à effets aléatoires. Modélisation et estimation pour des données de comptage (modèles Gamma-Poisson). Recherche clinique (calcul du nombre de sujets nécessaires par simulations, modèles bayésiens).

— Scan statistics : Détection des clusters atypiques d'événements à l'aide des statistiques de scan pour les modèles discrets et continus, i.i.d. et dépendance de type block-factor. Estimation par intervalle de confiance de la fonction de répartition de la statistique de scan dans le cas uni, bi- et tri-dimensionnel. Développement des outils d'aide à la décision basés sur les scan statistics.

— Programmation mathématique : Développement des algorithmes pour l'implémentation des techniques de classification non-supervisée pour données fonctionnelles. Algorithmes d'optimisation pour les modèles de régression non-linéaires (statistical learning). Algorithmes de simulation pour les statistiques de scan.

## 1.2 Publications

Profile Google Scholar (01/2022) : <https://scholar.google.fr/citations?user=hCy7I40AAAAJ&hl=fr>



**Cristian Preda**  
 Université Lille 1  
 Adresse e-mail validée de polytech-lille.fr  
[functional data](#) [scan statistics](#) [biostatistics](#)

SUIVRE

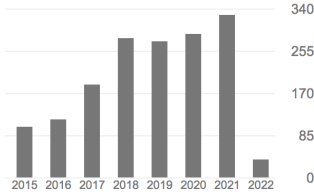
OBTENIR MON PROPRE PROFIL

| TITRE   | CITÉE PAR | ANNÉE |
|---|-----------|-------|
| <b>Functional data clustering: a survey</b><br>J Jacques, C Preda<br>Advances in Data Analysis and Classification 8 (3), 231-255            | 309       | 2014  |
| <b>Model-based clustering for multivariate functional data</b><br>J Jacques, C Preda<br>Computational Statistics & Data Analysis 71, 92-106 | 254       | 2014  |
| <b>Clusterwise PLS regression on a stochastic process</b><br>C Preda, G Saporta<br>Computational Statistics & Data Analysis 49 (1), 99-108  | 221       | 2005  |
| <b>PLS classification of functional data</b><br>C Preda, G Saporta, C Lévêder<br>Computational Statistics 22 (2), 223-235                   | 164       | 2007  |

Citée par TOUT AFFICHER

|            | Toutes | Depuis 2017 |
|------------|--------|-------------|
| Citations  | 2108   | 1399        |
| indice h   | 21     | 17          |
| indice i10 | 31     | 20          |



| Publications   | Nombre |
|--|--------|
| Articles dans <b>revues internationales</b><br>(peer reviewed)             | 50     |
| Conférences internationales à <b>comité de lecture et actes publiées</b> ) | 31     |
| Ouvrages collectifs/individuels  | 2      |
| Conférences invitées   | 15     |

Principales revues :

- Statistique et probabilités : Computational Statistics and Data Analysis , Chemometrics, Methodology and Computing in Applied Probability, Neurocomputing, Journal of Statistical Planning and Inference Statistics and Probability Letters
- Biostatistique et Médecine : Journal of Hospital Infection, Journal of Public Health Annals of Surgery, Studies in Health Technology and Informatics Value in Health, Journal of Translational Medicine Molecular Biology, Eur J Neurol

Principales conférences :

- ISI World Statistics Congress
- Applied Stochastic Models and Data Analysis (ASMDA)
- International Workshop on Applied Probability (IWAP)
- Compstat
- Journées de la Société Française de Statistique (SFDS)
- PLS'05 International Symposium (PLS)
- Congress of Romanian Mathematicians

### 1.3 Encadrement doctoral

4 thèses de doctorat des universités soutenues (2010-2016)

3 thèses de médecine (2005,2006, 2007)

1 thèse en cours (dernière thèse soutenue : 22 janvier 2016 ; contrat CIFRE en cours de réalisation) .

Liste des thèses soutenues :

*Florence Loingeville*, Statistique (2012-2016)

Titre : Modèle linéaire généralisé hiérarchique Gamma-Poisson pour le contrôle de qualité en microbiologie

Financement : CIFRE

Date de soutenance : 22 jan 2016

Taux d'encadrement : 80% (co-encadrement avec Julien Jacques - Lyon2)

*Alexandru Amarioarei*, Statistiques (2010-2014)

Titre : Statistiques de scan. Théorie et Applications.

Financement : Contrat Doctoral, Université de Lille 1, Laboratoire Paul Painlevé.

Date de soutenance : 15 Septembre 2014.

Taux d'encadrement : 100 %

*Michael Genin*, Statistiques (2010-2013)

Titre : Scan statistics and Epidemiology

Laboratoire: CERIM, Faculté de Médecine, Université de Lille 2

Financement : Contrat doctoral.

Date de soutenance : le 3 Décembre 2013

Taux d'encadrement : 50 % (co-encadrement avec Alain Duhamel)

*Costin Apostol*, Bioinformatique et Statistique (2007-2010)

Titre : Statistics and bioinformatics for the study of the HGF/MET interaction.

Laboratoire : CERIM, Faculté de Médecine, Université de Lille 2.

en collaboration avec l'Institut de Biologie de Lille (IBL).

Financé par l' Institut National du Cancer.

Date de soutenance : le 21 Décembre 2010

Taux d'encadrement : 50 % (co-encadrement avec Véronique Fafeur)

#### Thèses de Médecine :

*Stéphane Erouart*, Statistiques/Santé Publique.(2005-2007)

Titre : Classification des sites industrielles polluées au plomb dans la région Nord Pas de Calais.

Soutenue le 21 Février 2007, à la Faculté de Médecine de Lille, Université Lille 2.

*Emmanuel Chazard*, Statistiques/Santé Publique.(2004-2006)

Titre : Etude rétrospective de 17 ans d'activité du Centre d'Education pour le traitement du diabète et des maladies de la nutrition de CHG de Roubaix : Traitement de base de données, analyse, indicateurs d'utilité.

Soutenue le 25 Janvier 2006, à la Faculté de Médecine de Lille, Université Lille 2.

*Guillaume Delporte*, Pharmacie (2003-2005)

Titre :Dénombrement de bactéries de la flore intestinale résistantes aux biocides.

Soutenue le 25 Janvier 2005.

## 1.4 Projets/contrats de recherche

- Coordonateur INRIA pour le **projet ANR ClinMine** (2014 – 06/2017) : optimisation du parcours patient. Classification des données fonctionnelles qualitatives (<http://www.agence-nationale-recherche.fr/?Projet=ANR-13-TECS-0009>) : Budget 340000 euros pour INRIA.
- Projets internationales :
  - 2007–2011, **Projet Européen** (FP7 - ICT 2007-1, Health) : Patient Safety through Intelligent Procedures in medication (PSIP). Participant en tant que responsable scientifique pour la partie méthodologie statistique. Budget : 2.000.000 euros (CHRU Lille)
  - 2005–2007, **Action de Coopération Intégrée** avec la Région (Nord-Pas de Calais) et le Ministère de la Recherche et Technologies. Partenariat avec l'Institut de Biologie de Lille et l'Université de Lille 1. Responsable scientifique pour l'Université de Lille 2. Budget : 3000 euros.
- Réalisation des **projets de recherche avec des industriels** en tant que **coordinateur** :
  - PGXIS (2010-2013) : Pharmacogénétique et analyse de grandes bases de données pour la recherche en médicament. Budget : 12000 euros
  - Maia Eolis (2010-2011) : Prévission de la production d'électricité d'un parc éolien. Régression sur données fonctionnelles. Budget : 4000 euros
  - ASEL (2012-2013) : Analyse des clusters de lymphomes dans les départements Nord et Pas-de Calais par les statistiques de scan. Budget : 6000 euros.

## II Responsabilités scientifiques

### 2.1. Direction, animation laboratoires et équipes de recherche

Depuis Décembre 2020 je suis Directeur Scientifique de l'équipe MODAL, INRIA Lille Nord Europe.

De 2011-2020, je suis **responsable permanent (adjoint)** de l'équipe MODAL, INRIA Lille Nord Europe. Dans ce cadre, je coordonne un des axes fondateurs de l'équipe portant sur la modélisation des données fonctionnelles.

De 2005 à 2008 j'ai été le **responsable du Centre de Bio-Informatique** de la Génopole de Lille, pour l'Université de Lille 2.

De 2002 à 2008 j'ai **organisé et coordonné le séminaire** des « Applications statistiques et Informatiques en Santé de l'EA 2694 de l'Université de Lille 2.

### 2.1. Activités éditoriales

*Editeur associé* pour les journaux :

01/2013 - présent : Methodology and Computing in Applied Probability  
(<http://www.springer.com/statistics/journal/11009>)

01/2013 - présent : Romanian Journal of Mathematics and Computer Science  
(<http://www.rjm-cs.ro>)

### 2.2 Organisation de conférences/colloques :

Co-organisation des 13 conférences internationales (ASMDA, IWAP, COMPSTAT, SfdS, etc).

### 2.3 Rapporteur/Examineur (doctorat/HDR)

- *Rapporteur* de 7 thèses de doctorat : Yousri Henchiri (Univ. Montpellier 2, octobre 2013), Florentin Emanuel Olariu (UAIC-Iasi, Roumanie, septembre 2012), Bianca Roman (Faculté de Mathématiques, Université de Bucarest, Septembre 2012), Muhammad N. Sheraz (Faculté de Mathématiques, Université de Bucarest, Avril 2014) , Alin Rusu (Faculté de Mathématiques, Université de Bucarest, Avril 2015). Irina Bancescu (Faculté de Mathématiques, Université de Bucarest, Mai 2018), Amandine Schmutz, (Université Lyon 2, novembre 2019).
- *Rapporteur de 3 HDR*, Christophe Crambes (Université de Montpellier 2, Juillet 2020), Raluca Vernic (Université de Constanta, Roumanie, Juillet 2019), Mihaela Paun (Université de Bucharest, Roumanie, Juin 2015).
- *Membre des jurys* d'HDR : Sophie Dabo Niang (Univ. Lille 3, 2008) , Julien Jacques (Univ. Lille 1, 2012), C. Iliescu (Univ. Lille 1, Décembre 2013).
- *Membre des jurys* de thèses de doctorat : Yacouba Boubacar (Univ. Lille 3, novembre 2009), Tawfik Hamadeh (Univ. Lille 3, Juin, 2010), Camille Ternynck (Univ. Lille 3, 2014), Mohamed Salem Ahmed (Univ. Lille 3, Décembre 2017).

## 2.4 Autres responsabilités scientifiques (expert, rapporteur)

- Expert pour Executive Agency for Higher Education, Research, Development and Innovation Funding (Romanian Government). Evaluation des projets : 2016 funding call "Experimental Demonstrative Project" (<http://uefiscdi.gov.ro>)

- Rapporteur pour les journaux : Journal of American Statistical Association (JASA), Bernoulli, Journal of Statistical Planning and Inference (JSPI), Computational Statistics and Data Analysis (CSDA), Journal of Multivariate Analysis (JMVA), Canadian Journal of Statistics, Journal of Multivariate Behavioral Research, etc.

## III. Responsabilités pédagogiques et administratives

### Responsabilités pédagogiques :

- 09/2014 – 11/2017 : Responsable du Pole Mathématiques de Polytech'Lille.
- De 2013 à 2014 j'ai partagé avec Olivier Caron (enseignant Polytech'Lille, informatique) la responsabilité des projets de fin d'études pour les étudiants en dernière année d'ingénieur dans le Département Génie Informatique et Statistique (GIS).
- De 2010 à 2012 j'ai été responsable de la 3ème année d'ingénieurs GIS.

Création d'enseignements (Responsable de modules):

- j'ai créé en 2008 l'option de **Biostatistique** en 5ème année (GIS5). Elle ouvre chaque année depuis 2008. Une forte collaboration est ouverte depuis entre Polytech'Lille et la plateforme méthodologique de Biostatistique du CHRU de Lille (Professeur Duhamel) à travers des stages et emplois d'ingénieurs en biostatistique.
- j'ai introduit en 2008 à Polytech'Lille l'enseignement du langage R pour les statistiques. Aujourd'hui c'est le logiciel de base de la formation GIS.

### Administration :

**CNU** : 2006-2007 : membre titulaire élu du Conseil National des Universités, Section 26 -- Mathématiques appliquées et applications des mathématiques.

**Conseil de l'Université** de Sciences et Technologies de Lille (Lille 1). Membre élu du Conseil Académique et de la Commission de Recherche (2015-2017)

### Recrutement :

-2014-2015 : Président du Comité de sélection pour le poste de Professeur en statistiques ouvert à l'Ecole Polytechnique Universitaire de Lille (EPU).

- 2014-2015 : Président du Comité de sélection pour le poste de Maître de conférences en statistique à l'Université de Lille 2, ILIS.

- 2009- 2014 : Membre des comités de sélection pour la 26ème section : Université Lille 3 (2009-2014),  
Université de Lille 1 (2012)

## IV. Activités d'enseignements

### 4.1 Faculté de Médecine (2000 - présent)

De 2000 à 2008 j'ai été Maître de Conférences à la Faculté de Médecine de L'Université de Lille 2. J'ai assuré pendant cette période (et pour certains cours, encore aujourd'hui) des enseignements dans les filières suivantes :

- PCEM1 (PACES),
- MED2 – Enseignements thématiques (Biostatistique),
- DU de Biostatistique,
- Master 2 IIRS (Ingénierie Informatique dans les Réseaux de Soins) ,
- Master 1 Biologie et Santé.

**PCEM 1** (2001-2010): j'ai assuré une partie du cours de Probabilité (tests statistique, régression linéaires) ainsi que deux groupes de ED. Volume de l'enseignement (24h cours, 24h TD/année). Enseignement réalisé en collaboration avec les Professeurs Régis Beuscart et Alain Duhamel.

**MED-2** (2012 – présent) Enseignements thématiques (Bio-statistique) : j'assure chaque année, depuis 2012, quatre heures d'enseignements (2 séances) sur les thèmes : Analyse en Composantes Principales, Régression Linéaire et Analyse de survie. Enseignement réalisé en collaboration avec les Professeurs Régis Beuscart, Alain Duhamel et Emmanuel Chazard.

**DU Biostatistique** (2000-2009) : j'ai enseigné dans ce cursus les statistiques multidimensionnelles (ACP, Classification, Analyse Discriminante) pour un volume d'heures de cours de 12h/année. Enseignement en collaboration avec le Professeur Alain Duhamel et le biostatisticien Patrick Devos.

**Master 2 IIRS** (2000-2008) : j'ai été responsable et assuré les cours de data-mining (statistique multidimensionnelle) pour un volume de 24h de cours et 36 h de TP et projets par année. J'ai encadré chaque année au moins un mémoire de stage.

**Master 1 Biologie et Santé** (2000-2008) : j'ai assuré les cours de statistique inférentielle (12h) et analyses multi-dimensionnelles (12h). Enseignement en collaboration avec les Professeurs Mohamed Lemdani et Alain Duhamel et Monsieur Patrick Devos.

### 4.2 Ecole Polytechnique Universitaire, Polytech'Lille (2008 – présent)

Polytech'Lille est une école d'ingénieurs (ancienne EUDIL) faisant partie du réseaux Polytech (14 écoles au niveau national). Je suis rattaché au département Génie Informatique et Statistique. Dans cette formation je suis responsables de plusieurs cours dont les Probabilités, Modélisation Avancée, Régression linéaire et Biostatistiques.

Voici une liste exhaustive des mes enseignements à Polytech'Lille ainsi que dans le Master Recherche de Mathématiques (M2) de l'Université de Lille 1.

| Formation  | Niveau               | Intitulée                                    | Volume (C/TD/TP) |
|--|----------------------|--|------------------|
| Ingénieurs (Polytech'Lille)<br>formation initiale      | L3 – 1ère année ing. | Probabilités (2008-2014)                     | 32/24/8          |
| Ingénieurs (Polytech'Lille)<br>formation initiale      | L3 – 1ère année ing. | Stat. Inf. (2014-2015)                       | 20/0/32          |
| Ingénieurs (Polytech'Lille)<br>formation initiale      | L3 – 1ère année ing. | Reg. Lin. (2014-présent)                     | 8/0/12           |
| Ingénieurs (Polytech'Lille)<br>formation initiale      | M1 – 2ème année ing. | Mod. Avancée (2014-présent)                  | 20/0/24          |
| Ingénieurs (Polytech'Lille)<br>formation initiale      | M1 – 2ème année ing. | Stat. Explo. (2008-2014)                     | 20/0/18          |
| Ingénieurs (Polytech'Lille)<br>formation initiale      | M2 – 3ème année ing. | Biostatistique (2008-présent)                | 14/0/10          |
| Ingénieurs (Polytech'Lille)<br>formation apprentissage | M2 – 3ème année ing. | Plans d'exp. (2014-présent)                  | 20/0/20          |
| Ingénieurs (Polytech'Lille)<br>formation apprentissage | M1 – 3ème année ing. | Analyse de don. (2011-2015)                  | 20/0/20          |
| UFR Mathématiques                                      | M2 recherche         | Analyse de donnés fonctionnelles (2010-2014) | 24/12/12         |

### 4.3 Biostatistique

En 2008 j'ai créé à Polytech'Lille l'option de Biostatistique pour la dernière année d'ingénieur en Génie Informatique et Statistique. Cela est, au niveau régional et à l'exception des enseignements de biostatistique dirigés par le Professeur Alain Duhamel à la Faculté de Médecine de Lille, le seul module de Biostatistique au niveau Master.

Les thématiques que j'ai proposées dans cet enseignement sont clairement orientées vers l'utilisation des statistiques en Recherche Clinique :

- calcul du nombre de sujets nécessaires dans un test,
- tests de diagnostique : construction et évaluation,
- plans d'expériences : ANOVA,
- Mesures répétées et effet aléatoires,
- Analyse de survie,
- notions de base de recherche clinique (étapes dans la mise en place d'un protocole).

L'option est suivie chaque année par une quinzaine d'étudiants. Dès l'ouverture de l'option, une collaboration avec la plate-forme Méthodologie Biostatistique dirigée par le professeur Alain Duhamel s'est mise en place par la participation des membres de la plate-forme aux enseignements (Julia Salleron, Carole Langlois) et par la réalisation des stages de fin de formation au CHRU de Lille.

Plusieurs étudiants de cette option ont été recrutés, à la suite de leur stage dans l'unité de biostatistique, comme ingénieurs d'études biostatisticiens dans ce même service. Ce sont des étudiants que j'ai co-encadré avec le Professeur Alain Duhamel. Je citerais : Julia Salleron, Hélène Behal, Elodie Drumez, Mucheron Francois, Emilie Scrobola.

Michael Genin (Maître de conférences à la Faculté de Médecine de Lille) a été également un étudiant de Polytech'Lille qui a suivi l'option de Biostatistique en 2009. De 2010 à 2013, avec le Professeur Alain Duhamel, nous co-encadrons sa thèse de doctorat des universités sur un sujet très actuel des statistiques de scan. La thèse a été soutenue en Décembre 2013. Il était ensuite recruté comme MCU à la Faculté de Médecine où il développe actuellement, avec beaucoup de succès, la statistique spatiale et ses applications en épidémiologie.

Chaque année et aujourd'hui encore, l'Unité de Biostatistique du CHRU de Lille accueille au moins un étudiant de Polytech'Lille sur le thème de biostatistiques dans le cadre des stages de fin d'études.

## V. Collaborations Scientifiques et Pédagogiques

### 5.1 Faculté de Médecine de l'Université de Lille 2

Collaborations scientifiques :

Plusieurs collaborations en recherche avec des membres de l'EA 2694 sont actuellement en cours. Je mentionnerais les suivants :

1. Avec le Professeur Alain Duhamel nous travaillons sur l'approche bayésienne dans les essais cliniques séquentiels. Nous avons monté ensemble un groupe de travail autour de cette thématique auquel ont adhéré aussi bien les chercheurs en statistique de l'Université de Lille 2 (Faculté de Médecine, ILIS) que les bio-statisticiens de l'Unité de Biostatistique de CHRU de Lille. Ce groupe de travail se réunit avec une fréquence d'une réunion par mois et est souvent en collaboration avec le séminaire de statistique et informatique médicale organisé au CERIM (Centre d'Etude et de Recherches en Informatique Médicale, Faculté de Médecine, Lille 2).
2. Avec le Professeur Régis Beuscart dans le cadre de la SFR TSM (Technologie de la Santé et Médicaments) nous nous intéressons à l'utilisation des nouvelles techniques d'analyse statistique des grandes bases de données hospitalières et en particulier les données connectés (courbes, images).
3. Avec Dr. Michael Genin (Maître de Conférences à la Faculté de Médecine, Lille 2) sur l'utilisation des statistiques de scan en santé. Des problématiques théoriques (en continuité des ses travaux de thèse) sont au cœur de nos travaux communs aujourd'hui. Les collaborations avec d'autres chercheurs (le Professeur Hervé Hubert, Dr. Corinne Gower ) sont également avancées sur le plan de publications (Journal of Public Health – voir Annexe Publications).
4. Avec Dr. Emmanuel Chazard sur les problématiques liées à l'exploitation de grandes bases de données hospitalières. La mise en place des algorithmes statistiques (deep-learning, machine learning) pour les grandes bases de données hospitalières est une thématique très actuelle en santé publique. Emmanuel Chazard a effectué deux stages dans notre équipe INRIA Lille /Modal (2016 et 2017) à la suite desquels nous avons co-écrit un papier (Value in Health).

5. Avec Dr. Guillemette Marot sur l'analyse des données fonctionnelles et leurs applications en génétique. Nous partageons déjà des sujets de recherche en collaboration avec la plateforme bi-lille dont elle est responsable.

Collaborations pédagogiques :

Avec le Professeur Alain Duhamel nous avons depuis longtemps collaboré sur les enseignements des biostatistiques à la Faculté de Médecine aussi bien en PACES que des le Master de Biologie et Santé et le DU de Biostatistiques. Le contenu de ces cursus sont le fruit d'une étroite collaboration tant sur un plan méthodologique que sur un plan applicatif (logiciel).

Avec le Dr. Emmanuel Chazard et Dr. Gregoire Ficheur nous intervenons dans le module Biostatistiques du MED-2. Nous privilégions les aspects théoriques fondamentales nécessaires et les applications en médecine.

## **5.2 Unité de Méthodologie-Biostatistique de CHRU de Lille**

Cette collaboration s'est initialement mise en place tout naturellement par mon appartenance à l'équipe de Biostatisticiens du CERIM lorsque j'étais MCU à la Faculté de Médecine de l'Université de Lille (2000 - 2008). Elle a continué après mon départ comme Professeur à Polytech'Lille et s'est matérialisée par l'encadrement de stages des étudiants ingénieurs de Polytech'Lille et l'Unité de Biostatistique. Je rappelle ici les étudiants formés à Polytech'Lille et qui aujourd'hui sont acteurs majeurs dans le fonctionnement de l'unité de Biostatistiques : Julia Salleron (partie en 2013 à Nancy), Helene Behal, Elodie Drumez, Francois Muchron, Emilie Scrobola, Michael Genin.

Avec le responsable de la Plateforme, le Professeur Alain Duhamel, nous avons développé des collaborations scientifiques ayant comme point de départ des projets portés par des médecins du CHRU Lille. C'est ainsi que j'ai participé à de nombreuses consultations dans le cadre de l'unité de Biostatistique pour apporter une aide méthodologique aux questions des porteurs des projets. Actuellement ma participation dans le cadre de l'unité de biostatistique est formalisé par la réalisation d'une séance de consultations (un après-midi par semaine). Cela m'a permis de connaître le fonctionnement de cette unité dans son ensemble. L'application des statistiques dans le contexte de la recherche clinique est une activité que j'ai privilégié ces dernières années. Cela s'est réalisé par des discussions hebdomadaires avec le Professeur Alain Duhamel, les biostatisticiens Julien Labreuche et Patrick Devos et les autres ingénieurs de la plate-forme.

J'ai ainsi participé à la publication de plusieurs articles (Annals of Surgery, Journal of Public Health, Eur. J. Neurol). Voir liste de publications en annexe.

## **5.3 GHICL – Département de Recherche Médicale**

C'est en tant que Professeur à l'Université de Lille 1 que j'ai proposé et co-encadré en 2011 des étudiants ingénieurs en biostatistique (Polytech'Lille) dans le cadre de leur stage de fin d'études (4 à 6 mois) en recherche clinique dans le Département de Recherche Médicale de GHICL (Florence Loingeville, Pierre Lejeune, Zen Cao).

Depuis Septembre 2012, j'ai participé à la création et au développement de l'Unité de Biostatistique du Département de Recherche Médicale de GHICL par le recrutement des ingénieurs bio-statisticiens Laurene Norberciak et Stephane Verdun. Nous sommes aussi co-auteurs de plusieurs articles et conférences nationales (EPICLIN, Journées des Statisticiens des CLCC, etc) ou internationales (FAAM, ICB Seminar, etc).

## 5.4 INRIA

En tant que chercheur à l'Université de Lille 1, j'ai participé à la création en 2011 de l'équipe projet MODAL dans le cadre du Centre INRIA Lille. Je suis le responsable permanent de cette équipe depuis sa création et le responsable scientifique de plusieurs axes dans l'équipe (analyse de données fonctionnelles).

C'est dans le cadre de cette équipe que nous avons accueilli en 2012 Dr. Guillemette Marot dans le cadre de la chaire d'excellence INRIA-Lille 2.

En collaboration avec CHRU de Lille, le laboratoire CRISTAL (Professeur Clarisse Dhaenens) le CHRU de Montpellier et la société Alicante, nous avons monté le projet ANR ClinMine (2014-2017) financé à hauteur d'un million d'euros pour modéliser la parcours du patient à l'hôpital. J'ai été le coordonateur pour INRIA dans ce projet.

J'assure également à INRIA des encadrements de thèses de doctorat CIFRE et mémoires de Master M2 ainsi que des projets de recherche avec diverses entreprises (par la plateforme Eura-Technologies).

## 5.5 Mobilité internationale

### Invitations dans des universités étrangères :

1. Department of Mathematics of the University of Granada (Spain), June 30 - 5 July, 2014. Research with Professor Ana Maria Aguilera.
2. School of Physical & Mathematical Sciences, University Nanyang of Singapore, December 1--15, 2012. Invitation for research with Professor Lian Heng.
3. Department of Mathematics of the University of Granada (Spain), May 28 - June 3, 2011. Teaching (Master degree) and research.
4. Department of Mathematics of the University of Granada (Spain), May 23 - May 29, 2010. Teaching (Master degree) and research.
5. Department of Mathematics of the University of Luxembourg, March 15 -- May 15, 2010. Research project : Time series prediction via functional data methods. Financed by FNR Luxembourg.
6. Mission à la Faculté de Mathématiques et Informatique de l'Université de Bucarest, Roumanie, du 11 avril au 19 avril 2007, Programme ERASMUS. Cours de analyse de données fonctionnelles, Master "Statistique et Informatique", 2ème année.

## Annexe

### Publications :

1. **C. Preda**, Q. Grimonprez, V. Vandewalle (2021), Categorical Functional Data Analysis, *Mathematics* 2021, 9(23) ; <https://doi.org/10.3390/math9233074>.
2. A. Amarioarei and **C. Preda** (2020), One Dimensional Discrete Scan Statistics for Dependent Models and Some Related Problems, *Mathematics* 2020, 8(4), 576; <https://doi.org/10.3390/math8040576>.
3. Sophie Dabo-Niang, **Cristian Preda**, Vincent Vandewalle (2020), Clustering spatial functional data. *Geostatistical Functional Data Analysis : Theory and Methods*. Editors: Jorge Mateu, Ramon Giraldo, In press.
4. A Amarioarei, M Genin, **C Preda** (2019), Two dimensional discrete scan statistics with arbitrary window shape, *Handbook of Scan Statistics*, Springer.
5. C. Hirel, L Berton, **C. Preda**, Y Lambert, O Richard, F Pico (2019), Air Pollution and Humidity as Triggering Factors for Stroke. Results of a 12-year Analysis in the West Paris Area, *Rev. Neurol*, Volume 175, Issue 10, December 2019, Pages 614-618.
6. G Haiman, **C. Preda** (2018) Scan Statistics Viewed as Maximum of 1-Dependent Random Variables, *Handbook of Scan Statistics*, Springer.
7. ClinMine ANR consortium, **C. Preda** (2018), ClinMine: Optimizing the management of patients in hospital, *IRBM*, 39 (2), 83-92/
8. S Dabo-Niang, **C. Preda**, V Vandewalle (2018) Clustering spatial functional data, In, *Geostatistical Functional Data Analysis: Theory and Methods*. Editors: Jorge Mateu, Ramon Giraldo.
9. E. Hebbinckuys, J.-P. Marissal, **C. Preda** , V. Leclercq (2018), Assessing the burden of Clostridium difficile infections for hospitals, *Journal of Hospital Infection*, 98 (1), 29-35.
10. E. Chazard, G. Ficheur, J-B. Beuscart, **C. Preda** (2017), How to Compare the Length of Stay of Two Samples of Inpatients? *Value in Health*, Vol. 20, Issue 7, 992–998.
11. A Dermoune, **C. Preda** (2017), Parametrizations, fixed and random effects, *Journal of Multivariate Analysis* 154, 162-176.
12. A. Dermoune, **C. Preda**, Estimation of noisy cubic spline using a natural basis, *Annals of the University of Craiova - Mathematics and Computer Science Series*, Vol 43, No 1, 33-52 (2016).
13. N. Ajili, J.P. Decroix, **C. Preda**, J. Labreuche, F. Bourdain, B. Lapergue (2016), Impact of thrombolysis in acute ischaemic stroke without occlusion: an observational comparative study. *Eur J Neurol.*, 23(8):1380-6. doi: 10.1111/ene.13042.
14. V. Raverdy, V G. Baud, M. Pigeyre, **C. Preda**, F. Pattou (2016), Incidence and Predictive Factors of Postprandial Hyperinsulinemic Hypoglycemia After Roux-en-Y Gastric Bypass: A Five year Longitudinal Study, *Annals of Surgery* . 264(5):878-885

15. AM Aguilera, MC Aguilera-Morillo, **C Preda** (2016) Penalized versions of functional PLS regression, *Chemometrics and Intelligent Laboratory Systems* 154, 80-92.
16. A. Amarioarei, **C. Preda** (2015), Approximation for the Distribution of Three-dimensional Discrete Scan Statistic, *Methodology and Computing in Applied Probability*, Volume 17, Issue 3, pp 565-578.
17. F. Herbert, J. Jacques, **C. Preda**, S. Pied (2015), Evidence of IL-17, IP-10, and IL-10 involvement in multiple-organ dysfunction and IL-17 pathway in acute renal failure associated to Plasmodium falciparum malaria, *Journal of Translational Medicine*, vol. 13 :369, doi : 10.1186/s12967-015-0731-6 (2015).
18. A. Amarioarei, **C. Preda** (2014) Approximations for two-dimensional discrete scan statistics in some block-factor type dependent models, *Journal of Statistical Planning and Inference*, 151, 107–120.
19. J. Jacques and **C. Preda** (2014), Functional data clustering: a survey, in *Advances in Data Analysis and Classification*, 8, (3), 231-255.
20. J. Jacques, **C. Preda** (2014), Model-based clustering of multivariate functional data, *Computational Statistics and Data Analysis*, 72, 92-106.
21. A. Amarioarei, **C. Preda** (2013), Approximation for the Distribution of Threedimensional Discrete Scan Statistic, *Methodology and Computing in Applied Probability*, in press, DOI : 10.1007/s11009-013-9382-3.
22. G. Haiman, **C. Preda** (2013), One dimensional scan statistics generated by some dependent stationary sequences, *Statistics and Probability Letters*, Volume 83, Issue 5, 1457- 1463.
23. M. Genin, A. Duhamel, **C. Preda** (2013), Space-time clusters of Crohn's disease in northern France, *Journal of Public Health*, DOI: 10.1007/s10389-013-0580-9.
24. J. Jacques, **C. Preda** (2013), Funclust : a curves clustering method using functional random variable density approximation, *Neurocomputing*, 112, 164-171.
25. C. Iliescu, C. Sauvage, A. Decoster, **C. Preda**, M.-C. Castelain (2013), L'enfant allergique à l'arachide : une approche thérapeutique personnalisée, *Revue Française d'Allergologie*, 53 (1), 5-10.
26. **C. Preda**, C.Apostol (2010), Méthodes PLS pour l'analyse des données fonctionnelles, *Romanian Journal of Pure and Applied Mathematics*, Vol. 55, No. 6, p. 431-446.
27. A. Aguilera, M. Escabias, **C. Preda**, G. Saporta (2010), Using basis expansion for estimating functional PLS regression : application with chemometric data, *Chemometrics and Intelligent Laboratory Systems*, Vol. 104, Issue 2, p. 289-305.
28. G. Haiman, **C. Preda** (2010), A new Method of Approximating the Probability of Matching Common Words in Multiple Random Sequences, *Methodology and Computing in Applied Probability*, Vol. 12, No. 4, p. 775-795.
29. **C. Preda**, G. Saporta, Mohamed Mbarek (2010), The NIPALS algorithm for missing functional data, *Romanian Journal of Pure and Applied Mathematics*, Vol. 55, No. 4, p. 315-326.
30. Chazard E, **Preda C**, Merlin B, Ficheur G; PSIP consortium, Beuscart R.(2009), Datamining-based detection of adverse drug events, *Stud. Health Technol. Inform.*, Vol. 150, p.552-556.

31. Chazard E, Ficheur G, Merlin B, Genin M, **Preda C**, Beuscart R. (2009), Detection of adverse drug events detection: data agregation and data mining, *Stud. Health Technol. Inform.* Vol. 148, p. 75-84.
32. Pamelard F, Even G, Apostol C, **Preda C**, Dhaenens C, Fafeur V, Desmet R, Melnyk O. (2009), PASE: a web-based platform for peptide/protein microarray experiments, *Methods in Molecular Biology*, Vol. 570, p. 413-30 (2009).
33. **C. Preda**, G. Saporta, C. Lévêder (2007) PLS classification for functional data, *Computational Statistics*, 22, p. 223-235.
34. **C. Preda** (2007), Regression models for functional data by reproducing kernel Hilbert spaces methods, *Journal of Statistical Planning and Inference*, Vol. 137, 3, p. 829-840.
35. G. Haiman, **C. Preda** (2006), Estimation for the distribution of two-dimensional discrete scan statistics, *Methodology and Computing in Applied Probability Journal*, 8, 373-382.
36. **C. Preda** (2006), L'approche PLS pour l'analyse de données fonctionnelles, *Bulletin de la Société de Sciences Médicales de Luxembourg*, Vol. 2, p. 171-187.
37. [27] A. Duhamel, P.Devos, J.L. Bourriez, **C. Preda**, L. Lefebvre (2006), Functional data analysis for gait curves study in Parkinson's disease, *Studies in Health Technology and Informatics*, 124, p. 569-574.
38. **C. Preda**, G. Saporta, (2005), Clusterwise PLS regression on a stochastic process, *Computational Statistics and Data Analysis*, 49, p. 99-108.
39. **C. Preda**, G. Saporta (2005), PLS regression on a stochastic process, *Computational Statistics and Data Analysis*, 48, p. 149-158.
40. **C. Preda**, A. Duhamel (2005), Tools for statistical analysis with missing data. An application to a large medical dataset, *Studies in Health Technology and Informatics*, 116, p. 181-186.
41. G. Haiman, **C. Preda** (2002), A new method for estimating the distribution of scan statistics for a two-dimensional Poisson process, *Methodology and Computing in Applied Probability Journal*, 4 (4), p. 393-407.
42. **C. Preda**, G. Saporta (2002), Régression PLS sur un processus stochastique, *Revue de Statistique Appliquée*, XLX (2), p. 3-23.
43. **C. Preda** (2000), Approximation par moyennage de l'analyse en composantes principales d'un processus stochastique, *Comptes rendus de l'Académie des Sciences de Paris*, T. 330, Série I, p. 1-6.
44. **C. Preda**, (1998) Analyse harmonique qualitative des processus markoviens de sauts stationnaires, *Scientific Annals of « Alexandru Ioan Cuza » University of Iasi (Romania)*, Computer Science Section, Tome VII, p. 5-18.
45. **C. Preda** (1995), Structural characterization of Dedekind-continuous partial ordered sets, *Mathematical Reports*, Tome 47, p. 81-96, Bucarest.

**Conférences, congrès et colloques à communication (Conférences internationales à comité de lecture et actes publiées) :**

1. **C. Preda** (2021) Categorical Functional Data Analysis, The cfda R Package, ERCIM 2021, London, UK, December 17 – 21, 2021.
2. A. Amarioarei, **C. Preda** (2019) Dependent models for scan statistics, IMS China 2019, Dalian, China, June 6 – 10, 2019.
3. **C. Preda**, A. Amarioarei (2018) : One dimensional scan statistics associated with some dependent models, SMTDA 2018 – Chania, Greece, June 14-18, 2018.
4. **C. Preda**, P. Bastien (2017), Functional models applied to data from image analysis, 61st World Statistics Congress - ISI2017, Marrakech, Maroc, 16-21 July, 2017
5. A. Amarioarei, **C. Preda** (2017) Scan statistics for some dependent models and applications The 17th Conference of the ASMDA, London, 6-9 june, 2017.
6. **C. Preda** (2017) "Gilbert Saporta's contributions to functional data analysis", The 17th Conference of the ASMDA, London, UK, 6-9 june, 2017.
7. A. Amarioarei, **C. Preda** (2017) "Approximation for the scan statistics distribution of a three dimensional Poisson process", IMS China 2017, Nanning, China, June 28 - July 3, 2017
8. A. Duhamel, **C. Preda** (2017), Meta-analyses and indirect comparisons 154-th ICB Seminar on Statistics and Clinical Practice, Warsaw, Poland, 11-13 May, 2017
9. **C. Preda**, G. Saporta PLS regression with multivariate functional data SPSR 2017, Romania, Brasov, 28-29 april 2017
10. **C. Preda** : *Approximation of the scan statistics distribution using copulas. A numerical study.* International Workshop on Applied Probability (IWAP2016), 20-23 June, 2016, Toronto, Canada.
11. **C. Preda**, Vincent Vandewalle : *Clustering categorical functional data*, 48e Journées de Statistique de la Société Française de Statistique, Montpellier, June 2016.
12. **C. Preda**, Vincent Vandewalle : *Clustering categorical functional data*, 147th ICB Seminar Tenth International Seminar on statistics and clinical practice, May 15 - 18, 2016, Warsaw, Poland.
13. **C. Preda**, *Regression with categorical functional data*, A 19-a Conferinta a Societatii de Probabilitati si Statistica din Romania, Universitatea Tehnică de Construcții București, 27 mai 2016.
14. A. Amarioarei, **C. Preda**: *Scan statistics for some dependent models. Applications.*, The 16<sup>th</sup> Conference of the Applied Stochastic Models and Data Analysis (ASMDA 2015), Athens, Greece, 2015
15. **C. Preda**, PLS regression for multivariate functional data. The 7th International Conference of the ERCIM WG on Computational and Methodological Statistics (ERCIM 2014), University of Pisa, Italy, 6-8 December 2014.

16. **C. Preda**, Influence of the scanning window shape on the distribution of two-dimensional discrete scan statistics. The 7th International Workshop on Applied Probability (IWAP 2014), Antalya, Turkey, 16-19 June 2014.
17. **C. Preda**, PLS for multivariate functional data. The 8th International Conference on Partial Least Squares (PLS'14) and Related Methods, 26-28 May 2014, Paris, France.
18. **C. Preda**, Space-time clusters of lymphoma in northern France a scan statistics approach. Symposium on Modern Biotechnological Advances for Human Health BAHH May, 20-22, 2014, Bucharest, Romania
19. A. Amarioarei, **C. Preda** (2013), Approximations for two-dimensional discrete scan statistics in some dependent models, Proceedings of the 15th ASMDA International Society, Mataro (Barcelona), Spain, 25-28 June, 2013.
20. J. Jacques, **C. Preda** (2012) Clustering multivariate functional data, in COMPSTAT' 2012, Proceedings of 20th International Conference on Computational Statistics, Limassol, Chypre, August 2012, p. 353-366 (Ref:[http://compstat2012.org/Proceedings\\\_COMPSTAT2012.pdf](http://compstat2012.org/Proceedings\_COMPSTAT2012.pdf)).
21. J. Jacques, **C. Preda** (2012) Curves clustering with approximation of the density of functional random variables, in ESANN 2012 Proceedings, 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, April 2012, p. 459--464. (Ref: <http://www.i6doc.com/livre/?GCOI=28001100967420>).
22. J. Jacques, **C. Preda** (2012) Functional data clustering using density approximation, in 44èmes Journées de Statistique, Société Française de Statistique, Bruxelles. (Ref: <http://jds2012.ulb.ac.be/prog.php>).
23. **C. Preda**, J. Schiltz (2011), Functional PLS regression with functional response, ASMDA 2011, Rome, Italy.
24. **C. Preda**, Gilbert Saporta, M.H. Ben Hadj Mbarek(2009), PLS Regression with Functional Predictor and Missing Data, Proceedings of the 6th International Conference on Partial Least Squares and Related Methods, September 4-7, 2009, Beijing, China, pp. 17-22.
25. G. Haiman, **C. Preda** (2009), Approximations pour les statistiques de scan uni et bidimensionnelles, Proceedings of the 6th Congress of Romanian Mathematicians, June 28 - July 4, 2007, Bucharest, Romania. Ed. Romanian Academy, pp. 461- 466.
26. G. Saporta, **C. Preda** (2008), Adaptive forecasting on functional data, Proceedings of Italian Statistical Society 44th Meeting, University of Calabria, June 25--27, 2008, pp. 173-180, (2008).
27. **C. Preda**, G. Saporta, D. Costanzo(2006), Anticipated prediction in discriminant analysis on functional data for binary response, In Compstat2006, 17th Symposium on Computational Statistics and Data Analysis, Rome, August 28 - September 1st, pp. 821-828, Physica Verlag.
28. **C. Preda**, G. Saporta, C. Lévédér (2005), PLS classification for functional data, In PLS and related methods, Proceedings of the PLS'05 International Symposium, pp. 164-174, Barcelone, September 2005.

29. **C. Preda**, G. Saporta (2005), PLS discriminant analysis for functional data, In ASMDA 2005, XIth International Symposium on Applied Stochastic Models and Data Analysis, Brest, pp. 653-661.
30. **C. Preda**, G. Saporta (2004), PLS approach for clusterwise linear regression on functional data, In IFCS 2004 Proceedings of the IXth Conference of the International Federation of Classification Societies, Eds. Springer-Verlag, Chicago, pp. 167-176, (2004).