

□

# Offer #2024-07165

## PhD Position F/M [Campagne Allocation Région 2024] High-dimensional data analysis in the framework of survival modelling in presence of competing risks.

**Contract type :** Fixed-term contract

**Level of qualifications required :** Graduate degree or equivalent

**Fonction :** PhD Position

### About the research centre or Inria department

The Inria University of Lille centre, created in 2008, employs 360 people including 305 scientists in 15 research teams. Recognised for its strong involvement in the socio-economic development of the Hauts-De-France region, the Inria University of Lille centre pursues a close relationship with large companies and SMEs. By promoting synergies between researchers and industrialists, Inria participates in the transfer of skills and expertise in digital technologies and provides access to the best European and international research for the benefit of innovation and companies, particularly in the region.

For more than 10 years, the Inria University of Lille centre has been located at the heart of Lille's university and scientific ecosystem, as well as at the heart of Frenchtech, with a technology showroom based on Avenue de Bretagne in Lille, on the EuraTechnologies site of economic excellence dedicated to information and communication technologies (ICT).

### Context

The PhD candidate will join both the MODAL team based in the Inria center at the University of Lille (France) and the METRICS unit research based in the Faculty of Medicine at the University of Lille.

Inria MODAL team is specialised in statistical learning and most of her members also belong to Paul Painlevé laboratory, whose research covers a broad spectrum of mathematical research.

METRICS is a multidisciplinary public health research team that brings together clinicians and methodologists in the fields of ergonomics, biostatistics, medical informatics, data science and health economics.

### Assignment

The aim of the PhD thesis is to develop a new statistical method for variable selection in the context of survival analysis, with competing risks, where measurements are repeated over time on thousands of variables simultaneously.

For example, the aim will be to identify new omics markers (transcriptomics, metabolomics, proteomics) for predictive purposes in order to adapt patients' treatment (Precision Health). The originality of the statistical approach consists in simultaneously addressing three scientific challenges: i) high-dimensional analysis (more variables than individuals), ii) analysis of competing risks in the context of survival models and iii) analysis of longitudinal data (repeated measurements on the same individuals).

Results are expected both in terms of biostatistical methodology and software to provide a clinical decision support tool.

### Main activities

The main objective is to develop a statistical method for selecting variables in a context of survival prediction in the presence of competing risks.

Secondary objectives are:

- to improve the search for biomarkers in analyses with repeated omics data
- to develop a decision-support software tool that can be generic and can be used in any context of high-dimensional data analysis with survival models in the presence of competing risks.

PhD candidates who want to understand the concepts necessary for this PhD can read the following references:

- related to statistics

[1] Fine and Gray (1999) A proportional hazards model for the subdistribution of a competing risk. Journal

of the American statistical association, 94(446):496–509.

[2] Chi and Ibrahim (2006) Joint models for multivariate longitudinal and multivariate survival data. *Biometrics*, 62(2):432–445.

[3] Binder, Allignol, Schumacher, and Beyersmann (2009) Boosting for high-dimensional time-to-event data with competing risks. *Bioinformatics*, 25(7):890–896.

- related to one application which motivated the PhD subject (proteomics data analysis)

[4] Heyse, Vandewalle, Marot, Amouyel, Bauters, and Pinet (2023) Identification of patient subtypes based on protein expression for prediction of heart failure after myocardial infarction. *Iscience*, 26(3).

## Skills

The candidate must hold a MSc in Biostatistics.

Technical skills and level required :

- Proficiency in R and at least one scripting language (e.g. python)
- In-depth knowledge of statistical learning to handle high-dimensional data
- Basic knowledge of biology (molecular and cellular biology, genomics and genetics)

Other valued appreciated : creativity, independence, team working and communication skills

## Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

## General Information

- **Theme/Domain** : Computational Biology Statistics (Big data) (BAP E)
- **Town/city** : Villeneuve d'Ascq
- **Inria Center** : [Centre Inria de l'Université de Lille](#)
- **Starting date** : 2024-10-01
- **Duration of contract** : 3 years
- **Deadline to apply** : 2024-04-30

## Contacts

- **Inria Team** : [MODAL](#)
- **PhD Supervisor** :  
Marot Guillemette / [Guillemette.Marot@inria.fr](mailto:Guillemette.Marot@inria.fr)

## About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

**Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

## Instruction to apply

Please send us your CV and cover letter.

**Defence Security** :

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

**Recruitment Policy :**

As part of its diversity policy, all Inria positions are accessible to people with disabilities.