

**PROFIL DE POSTE**

**Post-doctorant Biostatisticien(ne)**

**Libellé de la fiche métier de rattachement** : Biostatisticien(ne)

**Sous-famille** : Conception, gestion et analyse des données

**Code métier** : 15D60

(*www.métiers-fonctionpubliquehospitaliere.sante.gouv.fr*)

**DIRECTION/SITE** : DRCI

**PÔLE/SERVICE** : DRCI

**IDENTIFICATION DU POSTE**

**Grade** : ingénieur hospitalier

**Niveau requis :** thèse d’Université en Biostatistiques ou Epidémiologie

**Poste :** post-doctorant

**Type, durée du contrat** : CDD d’un an, période de 3 ans

**Responsables hiérarchiques :**

- Responsables fonctionnels :Pr Marc Berger, responsable médical du projet BIO-TIMER et directeur d’équipe EA 7453 CHELTER, Université Clermont Auvergne (le post-doctorant sera membre de l’équipe de recherche) ; MrBruno PEREIRA, responsable du secteur Biométrie et Médico-Economie de la DRCI.

- Responsable administratif : Mr Xavier BIJAYE: Directeur de la DRCI.

**Collaboration :** Pr Arnaud Guillin, Pr. Pierre Latouche, Pr. Nourdinne Azzaoui, Laboratoire de Mathématiques Blaise Pascal UMR CNRS 6620 (Université Clermont Auvergne);

**Localisation du poste** : Service d’Hématologie Biologique & EA7453 CHELTER UCA,  Hôpital Estaing / DRCI, Secteur Biométrie et Médico-Economie, CHU Clermont-Ferrand / Laboratoire de Mathématiques Blaise Pascal UMR CNRS 6620, UCA, Campus des Cézeaux, Aubière.

**Quotité de travail** : 100%

**Jours travaillés :** du lundi au vendredi

**PRESENTATION DU SERVICE**

L’EA 7453 CHELTER et le service d’Hématologie Biologique s’intéressent aux trajectoires des patients atteints de leucémie, traités par thérapie ciblée et à l’impact du vieillissement individuel sur leur prise en charge.

La Direction de la Recherche Clinique et de l’Innovation (DRCI) créée à la fin des années 90 assure un rôle essentiel en matière d’élaboration et de mise en œuvre de la politique de Recherche du CHU.

**DEFINITION DU POSTE**

L'objectif du poste est de contribuer au projet de recherche clinique BIO-TIMER : Ce projet vise à comprendre l'impact du vieillissement sur la prise en charge thérapeutique des cancers, notamment depuis le développement des thérapies ciblées, enjeu majeur de la médecine personnalisée. Les outils utilisés par les gériatres pour évaluer la fragilité individuelle (FI) sont limités dans leur utilisation, du fait de leur validation dans les populations les plus âgées et de leur manque de sensibilité et de précision dans la phase précoce d’un vieillissement accéléré. Une évaluation biologique complémentaire du vieillissement individuel semble nécessaire.

Le projet BIO-TIMER, financé par la Fondation ARC et impliquant plus de 17 centres, vise à évaluer l'intérêt de combiner l'âge biologique évalué par l'analyse de la méthylation de l'ADN avec l'évaluation de la fragilité clinique individuelle pour personnaliser la prise en charge des patients atteints de cancer et recevant une thérapie ciblée. Nous proposons d'utiliser comme modèle la leucémie myéloïde chronique (LMC) traitée par des inhibiteurs de la tyrosine kinase (ITK), en raison de recul de l'utilisation des ITK, de la standardisation de sa prise en charge et de son aspect physiopathologique particulier. Les résultats préliminaires de l’EA 7453 CHELTER ont montré que l'horloge de méthylation de l'ADN du clone de LMC est modifiée (Lebecque B. et al., Cancers (Basel) 2021, 13(14):3587), ce qui suggère que le vieillissement des cellules du clone malin est altéré. Néanmoins, on ne connaît pas le vieillissement biologique du patient porteur de la LMC et son impact sur la prise en charge. Le projet BIO-TIMER a pour objectif de répondre à cette question. Pour le suivi de leur prise en charge, les patients seront inclus dans l’Observatoire LMC, une base de données en vie réelle multicentrique. Les données de FI et d’âge biologique seront enregistrées dans un e-CRF dédié.

Le post-doctorant jouera un rôle essentiel en tant que statisticien et sera activement impliqué dans le projet scientifique, en particulier dans l'analyse de grands ensembles de données et dans le choix des approches/méthodologies statistiques les plus appropriées. Il/elle sera responsable d'analyses statistiques sophistiquées de données complexes longitudinales et de survie et devra faire face aux problèmes typiques de ce type de données : multicolinéarité, effet centre, points temporels (visites) à différents moments, données manquantes, et plusieurs séquences de traitement. Les défis méthodologiques liés au projet concernent notamment la modélisation des séquences de traitement et de leur association avec l'âge à l'aide de modèles de chaînes de Markov et d'approches bayésiennes. Ces méthodes classiques seront confrontées et enrichies par des approches avancées d'apprentissage automatique, incarnées par l'emploi de réseaux de neurones profonds, notamment les réseaux neuronaux récurrents et de type Long Short-Term Memory (LSTM). Ces derniers sont particulièrement adaptés pour traiter la structure temporelle complexe des séquences de traitements, en capturant les structures de dépendances à long terme. L'approche adoptée se veut intégrative, tirant profit non seulement des données cliniques et biologiques traditionnelles, mais aussi des informations plus générales telles que les données de qualité de vie, les scores cliniques et les données omiques. En outre, l'intégration de techniques d'apprentissage par renforcement pourrait ouvrir la voie à la conception d'algorithmes capables de proposer des stratégies optimisées, en s'adaptant dynamiquement à l'évolution du profil clinique du patient. Très peu de travaux ont été publiés sur ces défis.

Le post-doctorant sera encadré et encouragé à développer de nouveaux modèles et initiatives de recherche en collaboration étroite avec le Laboratoire de Mathématiques Blaise Pascal UMR CNRS 6620 (Université Clermont Auvergne) et le secteur de Biométrie du CHU de Clermont-Ferrand. Par conséquent, le post-doctorant jouera un rôle de facilitateur dans la construction/maintien de l'interface entre les équipes et les membres d'autres workpackages (notamment Gériatrie et Biologie). De plus, il/elle sera activement impliqué(e) dans la rédaction d'articles (publication dans des revues internationales à comité de lecture, présentation des résultats lors de réunions professionnelles) et dans la rédaction de demandes de financement supplémentaires pour le projet.

**MISSIONS PRINCIPALES**

* Développer en collaboration avec les responsables du programme les analyses et les représentations structurées des données cliniques les plus pertinentes pour aider à leur interprétation.
* Interpréter et communiquer les analyses aux différents intervenants du projet en émettant des recommandations.
* Présenter les résultats du projet et de la recherche aux collaborateurs de l’étude.

**PRÉREQUIS**

**Diplômes / qualification :** Doctorat en biostatistiques avec de solides bases en statistiques et mathématiques appliquées à la biologie.

**Expériences professionnelles et connaissances :**

 - Expérience en tant que data-analyst / data-scientist / biostatisticien(ne).

- Maîtrisez la programmation sous R ou STATA.

- Connaissances et/ou expérience en IA

 - S’exprimer en anglais dans le cadre professionnel, à l'écrit comme à l'oral.

**COMPETENCES PROFESSIONNELLES REQUISES**

**Connaissances**

 - Méthodes épidémiologiques et statistiques.

- Capacité de développer et d'appliquer une méthodologie appropriée, de présenter des méthodes et de communiquer les résultats du projet à la fois par écrit et oralement.

 - Gestion éprouvée de grands ensembles de données, de l'acquisition à l'analyse, en particulier pour les données biologiques et cliniques.

- Des connaissances en intelligence artificielle seront appréciées.

**Compétences opérationnelles**

 - Promouvoir les meilleures pratiques et méthodes scientifiques en matière de statistiques avancées et de modélisation pour faciliter la cohérence des informations utiles.

- Identification des besoins et des problématiques : définir avec justesse le périmètre des données, intégrer les objectifs de la recherche, mettre en place un plan d’analyse nécessaire à l’identification des facteurs de prédiction.

- Faire progresser le programme de recherche en guidant les conceptions expérimentales et en garantissant les méthodes statistiques appropriées au contexte des objectifs scientifiques des études.

 - Restituer les travaux de manière synthétique et pertinente auprès des responsables.

**Compétences comportementales**

- Capacité de développer de nouvelles approches statistiques.

- Capacité d'apprendre et d'appliquer de nouvelles méthodes.

- Capacité à travailler de manière indépendante et collaborative au sein de groupes de recherche.

- Capacité d'interagir au sein des équipes multifonctionnelles.

- Etre force de propositions et innovant, et présentant un goût de l'innovation et des challenges.

- Reporting régulier auprès de la hiérarchie.

- Rigueur.

**Personnes à contacter**

Nom-Prénom : BERGER Marc

Fonction : chef de service HEMATOLOGIE BIOLOGIQUE, directeur équipe EA7453 CHELTER UCA

Coordonnées : mberger@chu-clermontferrand.fr

Nom-Prénom : PEREIRA Bruno

Fonction : responsable du secteur Biométrie et Médico-Economie - DRCI

Coordonnées : bpereira@chu-clermontferrand.fr

***Le profil de poste peut être réajusté en fonction****: des évolutions réglementaires et organisationnelles, des réajustements du projet pédagogique des Instituts et des projets institutionnels, des contraintes environnementales*



**JOB PROFILE**

**Biostatistician post-doctoral fellow**

**Wording of the job description:** biostatistician

**Subfamily:** data design, management and analysis

**Trade code:** 15D60

(*www.métiers-fonctionpubliquehospitaliere.sante.gouv.fr*)

**JOB IDENTIFICATION**

**Required level**: science thesis (PhD) in biostatistics or epidemiology

**Contract type**: CDD

**Grade**: Hospital engineer

**Duration:** 3 years

**Direction**: department of Clinical Research and Innovation (DRCI)

**Reporting line(s):**

- Functional managers:Prof. Marc Berger, medical manager of BIO-TIMER trial and et director of EA 7453 CHELTER group, Université Clermont Auvergne (the doctoral position will be a member of the team) ; M.Bruno PEREIRA, manager of Biometrics and Medical-Economics sector of DRCI department.

- Administrative manager: M. Xavier BIJAYE: Directeur de la DRCI.

**Collaboration:** Pr Arnaud Guillin, mathematics laboratory Blaise Pascal UMR CNRS 6620 (Université Clermont Auvergne) ;

**Job location**: Hematology (Biology) department & EA7453 CHELTER UCA, Hôpital Estaing / DRCI, Biometrics and Medical-Economics sector, hôpital G. Montpied, CHU Clermont-Ferrand / Blaise Pascal mathematics laboratory UMR CNRS 6620, UCA, Campus des Cézeaux, Aubière.

**Working time:** 100%.

**Days worked:** Monday to Friday, 7 hours/day. This organisation may vary according to the unit needs

**PRESENTATION OF DEPARTMENTS**

- The Clinical Research and Innovation Department (DRCI) was set up at the end of the 1990s and plays a key role in developing and implementing the CHU's research policy.

- EA 7453 CHELTER and the Biological Haematology Department are interested in the trajectories of leukaemia patients treated with targeted therapy and the impact of individual ageing on their care.

**PROJECT PRESENTATION AND JOB INTEGRATION**

The aim of the post is to contribute to the BIO-TIMER clinical research project: This project aims to understand the impact of ageing on the therapeutic management of cancer, particularly since the development of targeted therapies, a major challenge for personalised medicine. The tools used by geriatricians to assess individual frailty (IF) are limited in their use, due to their validation in the oldest populations and their lack of sensitivity and precision in the early phase of accelerated ageing. A complementary biological assessment of individual ageing appears necessary.

The BIO-TIMER project, funded by the Fondation ARC and involving >17 centres, aims to assess the value of combining biological age assessed by DNA methylation analysis with the assessment of individual frailty to personalise the management of cancer patients receiving targeted therapy. We propose to use chronic myeloid leukaemia (CML) treated with tyrosine kinase inhibitors (TKIs) as a model, because of the long history of TKI use, the standardisation of its management and its particular pathophysiological aspect. Preliminary results from the EA 7453 CHELTER group showed that the DNA methylation clock of the CML clone is modified (Lebecque B. *et al*., Cancers (Basel) 2021, 13(14):3587), suggesting that ageing of malignant clone cells is altered. However, the biological aging of the patient with CML and its impact on treatment remain unkown. The BIO-TIMER project aims to answer this question. Patients will be included in the CML Observatory, a multicentre real-life database, to monitor their care. IF and biological age data will be recorded in a dedicated e-CRF.

The post-doctoral fellow will play a central role as a statistician and will be actively involved in the scientific project, particularly in the analysis of our large real-life dataset and choice of the most suitable statistical approaches/methodology. He/she will be responsible of sophisticated statistical analyses of complex longitudinal and survival data and will have to deal with the typical issues of this type of data: multicollinearity, centre effect, time-points (visits) at different times, missing data, and several treatment sequences. The methodological challenges linked to the project involve notably modelling treatment sequences and their association with age using Markov chain models and Bayesian approaches. These classical methods will be compared and enriched by advanced machine learning approaches, embodied by the use of deep neural networks, in particular Long Short-Term Memory (LSTM) recurrent neural networks. These are particularly well suited to dealing with the complex temporal structure of processing sequences, by capturing long-term dependency structures. The approach adopted is intended to be integrative, taking advantage not only of traditional clinical and biological data, but also of more general information such as quality of life data, clinical scores and omics data. In addition, the integration of reinforcement learning techniques could pave the way for the design of algorithms capable of proposing optimised strategies, adapting dynamically to changes in the patient's clinical profile. Very few works have been published on these challenges.

The postdoctoral fellow will be mentored and encouraged to develop new models and research initiatives in collaboration with the Mathematics UMR CNRS 6620 Unit (University Clermont Auvergne). Therefore, the post-doctoral fellow will play a facilitating role in building/maintaining the interface between the teams and members of other workpackages. Moreover, he/she will be actively involved in writing papers (publication in international peer-reviewed journals, presentation of findings at professional meetings) and in drafting of additional funding requests for the project.

**GENERAL JOB DESCRIPTION**

- To develop, in collaboration with the programme managers, the most relevant analyses and structured representations of clinical data to assist in their interpretation.

- Interpret and communicate the analyses to the various project partners, making recommendations.

- Present the results of the project and research to the study collaborators.

**PREREQUISITES**

**Diplomas/qualifications**:PhD in biostatistics or epidemiology with a solid grounding in statistics and mathematics applied to biology.

**Professional experience and knowledge:** Experience as a data analyst / data scientist / biostatistician; Proficiency in programming in R or STATA; Ability to speak and write in English for professional purposes.

**PROFESSIONAL EXPERIENCE**

**Skills required**

- Knowledge of epidemiological and statistical methods.

- Ability to develop and apply appropriate methodology, present methods and communicate project results both in writing and orally.

 - Proven management of large data sets, from acquisition to analysis, particularly for biological and clinical data

- Knowledge of artificial intelligence would be appreciated

**Strengths required**

Promoting best practice and scientific methods in advanced statistics and modelling to facilitate the consistency of useful information.

- Identifying needs and issues: accurately defining the scope of the data, integrating the research objectives, setting up the analysis plan needed to identify predictive factors.

- Advancing the research programme by guiding experimental designs and ensuring the statistical methods are appropriate to the context of the scientific objectives of the studies.

 - To report on the work in a concise and relevant manner to those in charge.

**Behavioural skills**

- Ability to develop new statistical approaches.

- Ability to learn and apply new methods.

- Ability to work independently and collaboratively within research groups.

- Ability to interact within cross-functional teams.

- Be innovative and have a taste for innovation and challenges.

- Regular reporting to line management.

- Thoroughness.

**Contact person**

Full name: BERGER Marc

Position: chef de service HEMATOLOGIE BIOLOGIQUE, directeur équipe EA7453 CHELTER UCA

Mail: mberger@chu-clermontferrand.fr

Full name : PEREIRA Bruno

Position : responsable du secteur Biométrie et Médico-Economie - DRCI

Mail: bpereira@chu-clermontferrand.fr

***The job profile can be readjusted according to: regulatory and organisational developments, readjustments to the Institute educational project and institutional projects, environmental constraints***