

Ingénieur-e statisticien-ne

Concours externes 2022 – Ingénieurs et techniciens - Profil de poste – Concours n°16 – 1 poste

Corps IE - Ingénieur-e d'Etudes

BAP E - Informatique, Statistiques et Calcul Scientifique

Spécialité Statistique - Analyse de données / enquêtes

**RIFSEEP (régime
indemnitare
fonctionnaire)** Fonction : Ingénieur en traitement de données / enquêtes
Groupe : 2
Domaine : Laboratoires

Affectation **Unité 1018 - Centre de recherche en Epidémiologie et Santé des populations (CESP),
Villejuif**

**A propos de la
Structure** Le Centre de recherche en Epidémiologie et Santé des Populations ([CESP](#)) est l'un des principaux acteurs de la recherche épidémiologique en France.
Il bénéficie du soutien de 3 tutelles académiques : Inserm, Université Paris Saclay et UVSQ et de plusieurs partenaires hospitaliers en Ile de France.
Il accueille 11 équipes de recherche dont les thématiques vont des biostatistiques les plus fondamentales aux sciences humaines et sociales en passant par la recherche clinique et les études en population générale.
Son fonctionnement repose sur 2 services communs (affaires générales, plateau informatique) qui assurent les meilleures conditions de travail pour les chercheurs et pour le développement de leurs projets, et un pôle mutualisé de recherche qui conseille les chercheurs en matière de méthodologie statistique.
Le CESP compte près de 600 membres et a un budget annuel de 10 millions d'euros HT.
Le poste à pourvoir est basé sur le site de l'hôpital Paul Brousse à Villejuif (métro Paul Vaillant-Couturier).
<https://cesp.inserm.fr/fr>

Missions A l'interface entre la science des données et les biostatistiques, l'ingénieur-e d'études apportera son expertise à l'analyse des données en grandes dimensions qui pourront être de nature moléculaires, environnementales ou digitales, en ayant pour objectif principal la prédiction ou le partitionnement.

Son poste se situera :

- Au sein de l'équipe « Epidémiologie clinique » du CESP, où pour 50% de son temps il/elle proposera des méthodes statistiques appropriées aux objectifs de recherche, conduira des analyses de données et participera à leur valorisation scientifique en interaction avec l'équipe de recherche et les équipes partenaires au niveau national et international. Il/elle

définira et mettra en œuvre les analyses statistiques portant principalement sur des données cliniques et biologiques recueillies dans le cadre de la cohorte CKD-REIN ou d'autres études de l'équipe. Il développera une expertise concernant le traitement des données -omics (génomiques, transcriptomiques, protéomique) et des données anthropométriques et cliniques issues des entrepôts de données de santé.

- Au sein du pôle Méthodologies et statistiques mutualisé du CESP, où pour 50% de son temps il-elle contribuera aux projets de recherche et aux publications d'autres chercheurs de l'Unité, avec en priorité des projets ayant pour objectif le développement de modèle prédictif ou le partitionnement à partir de données de nature moléculaires, digitales ou environnementales en grande dimension. L'ingénieur-e d'études participera également aux autres activités d'accompagnements de la recherche du pôle, à savoir le conseil en statistique, le transfert de compétence à destination des jeunes chercheurs et ingénieurs ainsi qu'à des missions transversales (animation, outils de synergie inter équipes...).

Activités principales

- Elaborer le plan d'analyses statistiques répondants aux objectifs du projet en collaboration avec les chercheurs.
- Préparer les données et réaliser l'analyse statistique.
- Rédiger les rapports d'analyses et la documentation utile pour les chercheurs sous forme de document computationnel.
- Animer les réunions de présentation des résultats, participer à l'interprétation et contribuer à la valorisation des études .
- Elaborer et réaliser des programmes permettant d'automatiser certains traitements.
- Conseiller les chercheurs et ingénieurs sur la méthodologie et les techniques à mettre en œuvre pour analyser les données.
- Assurer la veille méthodologique sur les méthodes d'apprentissage automatique (machine learning).
- Encadrer sur le plan méthodologique et technique des étudiants et ingénieurs mettant en œuvre des méthodes d'apprentissage automatique.
- Identifier, analyser et comprendre les besoins lors de la conception des projets.
- Effectuer des présentations et des formations pour assurer un transfert de compétences.

Connaissances

- Connaissance approfondie en statistique.
- Connaissance des méthodes de gestion des bases de données.
- Connaissance des études de cohortes et de l'épidémiologie.
- Connaissance de l'environnement institutionnel et de l'organisation de la structure.
- Connaissance de la législation associée au traitement de ces données dont le RGPD, des méthodologies de références CNIL et du référentiel de sécurité des données SNDS
- Intérêt pour la recherche reproductible et les méthodes applicables pour une science transparente.
- Maîtrise de l'anglais : niveau B1 à B2 (CECRL).

Savoir-faire

- Maîtriser la méthodologie et les techniques statistiques et informatique de traitement de données couramment utilisées en recherche biomédicale et en épidémiologie clinique : méthodes d'échantillonnage, modélisation et analyse de données observationnelles, incluant

les méthodes d'inférence causale en situation observationnelle et les méthodes d'analyses adaptées aux données longitudinales, méthodes d'imputations.

- Maîtriser les méthodes de sélection de variables et d'apprentissages statistiques supervisées (régression pénalisée, forêts aléatoires, SVM, gradient boosting).
- Maîtriser les méthodes d'exploration statistique multidimensionnelle (méthodes factorielles) et de partitionnement (k-means, classification hiérarchique).
- Maîtriser les langages R, SAS et Python pour l'analyse statistique et la gestion des données.
- Maîtriser l'environnement Git et du langage Markdown.

Aptitudes

- Esprit d'analyse.
- Autonomie, esprit d'initiative et rigueur.
- Forte capacité d'organisation.

Spécificité(s) et environnement du poste

- Travail sur écran.
- Travail en équipe pluridisciplinaire.
- Déplacement sur les différents sites d'implantation du Centre (CHU de Bicêtre, Hôpital Cochin, Gustave Roussy, Université Paris Saclay et UVSQ ...).

Expérience souhaitée

- Une expérience professionnelle d'au moins 3 ans dans un laboratoire de recherche en épidémiologie ou une agence de santé est souhaitée.
- Une expérience dans l'analyse statistique de larges bases de données longitudinales et des données en grandes dimensions est souhaitée (travaux ayant conduit à une publication scientifique).

Diplôme(s) souhaité(s)

- Master 2 recherche en (bio) statistique ou école d'ingénieur en statistique.

Diplôme requis

- Diplôme de niveau 6 requis (anciennement niveau II).

Informations Générales

Temps de travail

- Temps plein.
- Nombre d'heures hebdomadaires : 38h30.
- 32 jours de Congés Annuels et 13 jours de RTT.

Activités télétravaillables

- OUI * NON
- * 2 jours par semaine hors contexte sanitaire spécifique

Rémunération

Selon la grille indiciaire correspondant au corps administratif de recrutement, une reprise d'ancienneté selon les règles statutaires et un régime indemnitaire (RIFSEEP) correspondant à la fonction occupée.

Pour en savoir +

- Sur l'Inserm : <https://www.inserm.fr/> ; site RH : <https://rh.inserm.fr/Pages/default.aspx>
- Sur la politique handicap de l'Inserm et sur la mise en place d'aménagements de poste de travail, contactez la Mission Handicap : emploi.handicap@inserm.fr