

Ingénieur d'études - Data-Manager

L'équipe « Générations et Santé » base l'essentiel de ses travaux de recherche sur les grandes études de cohorte dont elle a la charge : E3N (www.e3n.fr), E3N-EPIC et E4N (www.e4n.fr). Basée à Gustave Roussy (Villejuif) et dirigée par Gianluca Severi, l'équipe comprend une cinquantaine de personnes et fait partie du Centre de recherche en Epidémiologie et Santé des Populations (CESP, Inserm U1018).

E3N (Etude Epidémiologique auprès des femmes de l'Education Nationale) est une étude de cohorte à laquelle participent environ 100 000 femmes volontaires adhérentes à la MGEN (Mutuelle Générale de l'Education Nationale). Les informations concernant leur mode de vie et l'évolution de leur état de santé (notamment cancer et maladies chroniques) sont recueillies par auto-questionnaires tous les 2-3 ans depuis 1990. E3N est également la partie française de EPIC, vaste étude européenne coordonnée par le Centre International de Recherches sur le Cancer portant sur près de 500 000 européens dans 10 pays. E4N vise à prolonger E3N en suivant les conjoints et les descendants des femmes de l'étude E3N (enfants et petits-enfants). L'objectif de cette vaste cohorte familiale sur 3 générations est de mieux comprendre, dans l'apparition des maladies, ce qui relève de la part génétique, de l'environnement familial et de l'environnement extra-familial.

Le système d'information E3N-E4N comprend une infrastructure de virtualisation hébergeant un système de gestion électronique de documents (GED), un système de lecture automatique de documents (LAD) pour l'acquisition et la gestion des données « papier », ainsi qu'une plateforme informatique web pouvant agréger des données de sources multiples: questionnaires en ligne, données issues de technologies numériques (smartphones, capteurs, objets et dispositifs médicaux connectés, etc.) et appariement avec de grandes bases de données médico-administratives .

Au sein du Service technique de l'équipe (20 personnes), le groupe Data Management (6 à 7 data managers) gère cette grande base de données E3N-E4N et met des jeux de données à disposition des chercheurs.

Missions

Dans ce cadre, la mission du Data Manager sera de participer à la gestion d'ensemble des bases de données E3N et E4N et de prendre en charge le data management de projet sur la **maladie de Parkinson**. Il travaillera sous la direction de la responsable du service technique et, pour les activités concernant des projets spécifiques, en étroite relation avec les chercheurs et responsables concernés.

Activités principales

Le Data Manager aura les activités spécifiques suivantes :

- Effectuer les mailings aux participants et aux médecins :
 - Sélection de la population concernée et des médecins traitants
 - Conception des questionnaires
 - Paramétrage de système de lecture automatique des documents (LAD)
 - Supervision de l'acquisition des questionnaires
 - Création des tables de données SAS et mise à disposition
- Organiser le nettoyage des bases de données brutes, une fois la collecte des données effectuées (cohérence, correction des valeurs aberrantes)
- Rédiger des rapports faisant état de l'ensemble des nettoyages, modifications et créations de base de données effectués
- Mise à jour des données dans la base de données de suivi d'envoi (RedCap)
- Participer aux réunions avec les chercheurs et le responsable du projet
- Répondre aux besoins des chercheurs en proposant les solutions informatiques

Activités associées

II/Elle confiera aux assistants de data management des tâches de validation, visualisation, correction, saisie et en assurera la supervision. II/Elle rédigera des rapports faisant état de l'ensemble des nettoyages, modifications et créations de base de données effectués.

II/Elle participera à l'élaboration des procédures d'assurance qualité du data management et veillera à leur mise en œuvre.

Connaissances

Connaissance approfondie

- des méthodes de recueil et de traitement des données dans le domaine des sciences et de la santé,
- des outils adaptés au contrôle de cohérence des données.

Savoir-faire

- Maîtrise indispensable
 - des bases de données relationnelles et du langage SQL (requêtes de sélection et de manipulation des tables)
 - des langages informatiques pour l'exploitation des bases de données relationnelles (SQL, SAS base et macros)
- Une première expérience dans la gestion de données de santé ou la gestion de très grandes bases de données seraient un plus

Aptitudes

- Qualité relationnelle et facilité à travailler en équipe
- Rigueur, organisation, autonomie
- Capacité à rendre compte de son activité

Pour postuler

Les dossiers de candidature comprendront : un CV détaillé, une lettre de motivation, et si possible les coordonnées de 1 ou 2 personnes de référence. Ils devront être transmis par mail sous la référence "Data manager_2019" à alexis.elbaz@inserm.fr et anita.kowal@gustaveroussy.fr.