

# Ingénieur-e d'études en production, traitement, analyse de données et enquêtes

Concours externes 2022 – Ingénieurs et Techniciens - Profil de poste – Concours n°15 – 1 poste

**Corps** IE - Ingénieur-e d'Etudes

**BAP** D - Sciences Humaines et Sociales

**Spécialité** Cohortes épidémiologiques

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

**Affectation** **Unité de Service 11 - Cohortes épidémiologiques en population, Villejuif**

**A propos de la  
Structure** L'unité mixte de service a la responsabilité scientifique et technique de 2 cohortes en population ouvertes à la communauté des chercheurs : Gazel et Constances. L'équipe est constituée d'une trentaine de personnes de différentes disciplines (épidémiologistes, statisticiens, informaticiens, etc.) et est structurée autour de 9 « *Work Packages* » thématiques, chacun porté par un responsable qui pilote les activités et les différents chefs de projet sur son périmètre, incluant la gestion et la coordination des centres participant aux études, le recueil et la gestion des données de questionnaires, le stockage et la gestion des échantillons biologiques, arbitrage des projets scientifiques, etc. Les responsables se réunissent de manière bimensuelle. Avec 220 000 volontaires les données massives de Constances permettent de répondre à diverses questions de recherche et de surveillance épidémiologique. À l'inclusion, les volontaires répondent à un large questionnaire portant sur plusieurs thématiques et événements de vie et bénéficient d'un examen médical dans un Centre d'Examen de Santé de la sécurité sociale. Les suivis incluent :

- Un auto-questionnaire annuel
- Un appariement régulier avec les bases du Système National des Données de Santé (SNDS) et de la Caisse Nationale d'Assurance Vieillesse (CNAV)
- Un examen de santé régulier proposé dans les CES (tous les 3 à 5 ans)

Des recueils complémentaires pour des études spécifiques sont régulièrement effectués. Une biobanque est également en cours de constitution.

Lancée en 1989, la cohorte Gazel est la "petite GRANDE sœur de Constances". Constituée à l'origine de 20 000 agents d'EDF-GDF en activité, elle est toujours active. Depuis 2018, les questionnaires entre les 2 cohortes sont identiques.

**Missions** L'ingénieur-e d'études sera rattaché-e et encadré-e d'un point de vue technique par le WP traitement de données et d'un point de vue métier par l'un des WP thématique en fonction du

type de données. L'ingénieur-e d'études sera impliqué-e dans l'ensemble du processus de gestion de la qualité des données : de la saisie à la mise à disposition aux utilisateurs :

- Assurer la supervision et la mise en œuvre des plans de gestion de la qualité des données saisies (vidéo codage ...).
- Assurer la supervision et la mise en œuvre des procédures de nettoyage des données et des tests post-intégration.
- Mettre à jour le catalogue de données (Excel ou via application spécifique) permettant les extractions puis mises à disposition de données aux utilisateurs.
- Effectuer la requête dans les bases de données MySQL et Access.
- Traiter des demandes d'analyse de données et de calcul d'effectifs venant de la direction.
- Créer des tableaux croisés et autres indicateurs à la demande de la direction.

### Activités principales

- Superviser l'ensemble des activités de vidéocodage sur Constances et GAZEL sur le périmètre de remontées par le prestataire en charge de la saisie (l'Imprimerie Nationale).
- Réaliser la requête et la création de procédures automatisées dans les bases de données (Access, MySQL) : vérification globale de la qualité des données (nettoyage et tests-post intégration), calcul d'effectifs, création de tableaux de bord...
- Mettre à jour et effectuer le maintien du catalogue des données (en interaction avec l'ensemble des membres de l'équipe impliqués dans la production de données, l'ouverture, le SI et l'expert de données, les extracteurs) : assurer l'enrichissement du catalogue, l'alimentation des libellés en base, le contrôle qualité, l'enregistrement centralisé des anomalies remontées la traçabilité des décisions prises pour les corriger.
- Assurer la documentation des variables nettoyées.
- Analyser et traiter les données de la cohorte dans le cadre du plan de nettoyage des données (données médicales, données biologiques ...) en s'adaptant aux besoins et spécificités des données massives (ex. analyse de la qualité en transversale inter centre et en longitudinal intra centre pour les examens de santé et neuropsychologie ...).
- Créer des indicateurs de suivi d'activité autour de la qualité des données (à destination de la direction de la cohorte) et effectuer leur maintien.

### Activités associées

- Saisir et intégrer les abandons de volontaires dans les bases de données et remontée des points d'arbitrage.
- Gérer la correction des erreurs dans les consentements biobanques et inclusion.
- Gérer les relations avec les partenaires (Imprimerie Nationale...).

### Connaissances

- Connaissances des méthodes de gestion d'enquêtes épidémiologiques.
- Connaissance des données de santé.
- Connaissance des méthodes d'analyse statistique.

### Savoir-faire

- Maîtriser au moins un logiciel de traitement statistique SAS, Stata ou R.
- Avoir une bonne connaissance du langage SQL serait un plus.
- Maîtriser la programmation pour des tests de cohérence, les extractions de données chiffrées et les statistiques descriptives.
- Maîtriser les outils graphiques et de mise en forme d'édition.
- Rédiger des rapports d'analyse et des documents méthodologiques.
- Savoir tenir un planning.

### Aptitudes

- Capacité d'analyse et rigueur.
- Sociabilité et aptitude au travail collectif.
- Autonomie et adaptabilité.
- Sens de l'organisation.
- Sens de la communication.

### Spécificité(s) et environnement du poste

Être tenu au secret professionnel et garantir la sécurité des données.

**Expérience souhaitée**

- Une expérience de 3 ans minimum est souhaitée.

**Diplôme(s) souhaité(s)**

- Diplôme d'Ingénieur ou de formation scientifique (Bac+5) en statistiques.
- Master en épidémiologie, en santé publique...

**Diplôme requis**

- Diplôme de niveau 6 requis au minimum (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

\* Possibilité de 2 jours de télétravail

**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](mailto:emploi.handicap@inserm.fr)