DMP du projet "PGD Become"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "PGDR type "projet de recherche" IRSN" fourni par IRSN - Institut de Radioprotection et de Sureté Nucléaire.

Renseignements sur le plan

Titre du plan	DMP du projet "PGD Become"
Livrable	livrable ANSES
Version	Version initiale
Objet/périmètre du plan	PGDR de l'étude BECOME exigé par l'ANSES pour bénéficier de la subvention.
Domaines de recherche (selon classification de l'OCDE)	Health sciences
Langue	fra
Date de création	2023-03-31
Date de dernière modification	2023-04-25

Renseignements sur le projet

Titre du projet	PGD Become
Acronyme	Become
Résumé	**Contexte** Le cancer est l'une des principales maladies radio-induites qui peut survenir à de faibles niveaux de dose d'exposition, comme le démontrent de nombreuses études épidémiologiques, notamment chez les travailleurs du nucléaire. Pour les professionnels de la santé, le risque de cancer a été observé pour des expositions prolongées et cumulées sur une longue période. Certaines publications ont rapporté un excès de tumeur cérébrale chez les cardiologues interventionnels, mais le lien avec la profession n'a jamais été formellement prouvé. Cependant, ces études sont peu documentées en termes d'évaluation dosimétrique, avec le plus souvent une reconstitution rétrospective basée en partie sur l'historique de la carrière. **Objectifs** L'étude proposée ici vise à réaliser une étude internationale cas-témoins nichée pour étudier le risque de décès par cancer du système nerveux central (SNC) radio-induit en fonction de la dose de rayonnements ionisants (RI) reçue au niveau professionnel dans trois cohortes de professionnels de la santé exposés aux RI en France, en Corée et aux États-Unis.

Méthodes

Trois cohortes nationales construites pour étudier le risque de maladies radio-induites chez les professionnels médicaux sont impliquées, la cohorte ORICAMs (Occupational Radiation induced Cancer In Medical Staff), mise en place en 2011 par l'Institut français de radioprotection et de sûreté nucléaire, suivant 164 000 individus depuis 2002, la cohorte coréenne des travailleurs médicaux, mise en place en 2017 par l'Université coréenne de Séoul, incluant 94 000 individus suivis depuis 1996 et la cohorte américaine des radiologistes (US. La cohorte américaine des technologues en radiologie (USRT), mise en place en 1983 par le National Cancer Institute, qui comprend environ 110 000 travailleurs médicaux exposés de 1920 à 2010.

Le projet prévoit une analyse cas-témoins conjointe nichée dans les trois cohortes, avec 300 décès par tumeur du SNC (193 dans la cohorte USRT, 67 dans la cohorte ORICAMs et 40 dans la cohorte coréenne) identifiés. Cinq témoins seront appariés de manière aléatoire à chaque cas dans leur propre cohorte afin de prendre en compte l'hétérogénéité de la population étudiée en termes de structure d'âge, de mode de vie et de facteurs environnementaux. Des informations supplémentaires (y compris le diagnostic de la tumeur cérébrale et d'autres comorbidités, la profession, la spécialité, le(s) poste(s) de travail, les indicateurs du statut socio-économique) seront obtenues à partir des cohortes ou des dossiers de santé au travail. Une reconstruction dosimétrique sera effectuée pour estimer les doses cumulées au cerveau, en tenant compte des incertitudes liées aux méthodes de collecte et aux types de badges dans chaque cohorte.

L'odds ratio et les intervalles de confiance basés sur la vraisemblance du décès par tumeur du SNC associés à la dose de radiation seront estimés en comparant l'exposition des cas à celle de leurs témoins appariés, en utilisant des méthodes de régression conditionnelle pour les analyses univariées et multivariées, en tenant compte des facteurs de confusion potentiels. En termes de puissance statistique, avec l'hypothèse d'un niveau de base de sujets "plus exposés" de 17% et 11%, respectivement pour les cas et les contrôles (correspondant principalement aux médecins interventionnels, à la fois cardiologues et radiologues et aux travailleurs de la médecine nucléaire, y compris les manipulateurs radio et les médecins), l'inclusion de 300 cas et de 1500 contrôles appariés permettra de détecter un odds ratio de 1,66 ou plus pour une puissance statistique de 79% et un niveau alpha de signification de 5%.

Résultats attendus

Cette étude permettra d'améliorer les connaissances sur les risques sanitaires associés à de faibles doses prolongées d'IR qui sont des modes d'exposition pertinents pour les professionnels de la santé et le public. Cette analyse conjointe permettra d'étudier l'impact sanitaire de l'exposition professionnelle aux RI chez les travailleurs médicaux qui constituent le plus grand groupe de travailleurs exposés aux RI, avec certains professionnels potentiellement plus exposés que la moyenne. Les résultats contribueront à améliorer les stratégies de radioprotection des travailleurs médicaux et à sensibiliser les professionnels de la santé à l'importance de la radioprotection dans leur

environnement professionnel. D'autres sites de cancer seront étudiés à l'avenir selon la méthodologie et la collaboration développées dans ce projet.

Sources de financement

- Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail : ANSES-22-EST-033

Date de début 2023-05-15

Date de fin 2025-05-19

Partenaires

- Isglobal 0000000417633517
- université de séoul 0000000404705905
- National cancer institute 0000000098268822

Produits de recherche :

1. Bases de données démographiques et dosimétriques (Jeu de données)

Contributeurs

Nom	Affiliation	Rôles
Bernier Marie-Odile - 918873	INSTITUT DE RADIOPROTECTION ET DE SURETE NUCLEAIRE - 0000000114146236	• Coordinateur du projet • Personne contact pour les données • Responsable du plan de gestion de données

Droits d'auteur :

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

DMP du projet "PGD Become"

Informations sur le service de recherche

Nom du service (ou laboratoire, le cas échéant) de recherche

Service de recherche sur les effets biologiques et sanitaires des rayonnements ionisants, laboratoire d'épidémiologie

Acronyme service (ou laboratoire) et appartenance

SESANE/LEPID

Structure IRSN de recherche *(selon répertoire national des structures)*

- URSan 201722594P
- URSan 201722594P

Partenaires impliqués dans le projet de recherche

Mme Isabelle Thierry Chef
Barcelona Institute for Global Health (ISGlobal)
Dr Aiguader 88 08003 Barcelona, Espagne
email: isabelle.thierrychef@isglobal.org

M. Lee Won Jin
Korea University
College of Medicine, Dept Preventi Med.
02841 Seoul, Korea

email :leewj@korea.ac.kr

Mlle Cari Kitahara
National Cancer Institute
Division of cancer epidemiology & genetic
9609 Medical Center Drive 7E456
20892 Bethesda, Maryland, USA

email: meinholdc@mail.nih.gov

Responsable(s) de la gestion et de l'ouverture des données

Marie-Odile Bernier

Informations générales sur les données du projet

Description des données

Description des données	
Type(s) *(numériques, textuelles, images, vidéo...)*	données numériques et textuelles
Format(s)	format texte et format numérique
Volume *(approximatif en espace de stockage par ex., à ré-évaluer en cours de projet)*	faible volume

---

Préciser le mode d'obtention des données produites

- 1- Données d'observation
- 1- Données d'observation

les données ont été collectées à partir de registres de données, de dossiers médicaux, de questionnaires complétés par les personnes incluses dans les études

---

Données utilisées par le projet et provenant de sources extérieures

Les données des cohortes étrangères proviennent des USA , de Corée du sud

---

Logiciel(s) de production ou de traitement nécessaire(s) avec sa (leur) version

Logiciel SAS, version 9.4
Logiciel R, version , version 2

"Criticité" des données : que se passerait-il si les données étaient corrompues ou perdues?

- 2- Elles sont re-générables avec un effort/coût modéré
- 2- Elles sont re-générables avec un effort/coût modéré

les données sources restent dans les organismes qui les détiennent. Une copie est envoyée pour le projet spécifique

---

Documentation et qualité des données au fil du projet de recherche

Qualification des données au fil du projet de recherche

réception de données brutes; elles seront nettoyées (contrôles de cohérence) et traitées (création de doses à l'organe en

fonction des données dosimétriques fournies) à l'IRSN.
Elles seront analysées et interprétées à l'IRSN.
Ces différentes étapes sont tracées dans une fiche de vie de la base de données en interne au LEPID

Règles de nommage des ensembles de données

Pour chaque jeu de données, un fichier précisant le nommage des données est fournie par l'organisme détenteur des données.
ces fichiers se trouvent sur un serveur dédié au LEPID pour le stockage des bases de données.
Z:\ORICAMS\projet BECOME\DATA SOURCE
usrt_brain_cns_cc_20200915_codebook..xlsx : dictionnaire des données pour les USA
Codebook-notes.docx: dictionnaire des données pour la corée du sud
Z:\ORICAMS\1.1__Data management
oricams dico variables 2022.xlsx : pour les données françaises

---

Arborescence de classement

Les données techniques sont sur un serveur sécurisé dédié au LEPID pour les bases de données : Z:\ORICAMS\projet BECOME\
Les données administratives sont sur le serveur O:\SANTE\SESANE\LEPID\AAA_gestion de doc LEPID\Contrats LEPID\personnel medical\BECOME
Les publications et la traçabilité des résultats sont dans le fichier O:\SANTE\SESANE\LEPID\AAA_gestion de doc LEPID\Production scientifique LEPID\Publis LEPID

Gestion des données pendant le projet de recherche

Stockage des données tout au long du processus de recherche

Données	Brutes	Traitées/nettoyées	Analysées	Interprétées/figées
A	réseau Z sécurisé IRSN	réseau sécurisé Z IRSN	réseau sécurisé Z IRSN	réseau sécurisé Z IRSN
B				
.../...				

Indiquer l'espace de stockage

Plan de sauvegarde des données

sauvegarde supervisée par D3NSI, la sauvegarde (et la restauration le cas échéant) est garantie par les procédures d'exploitation fixées dans un référentiel (pour info: le référentiel Ouvre dans une nouvelle fenêtreIRSN/S3A/REF-057)

Conservation des données au terme du projet de recherche

Définition des critères de tri et de sélection des données à conserver

les données se divisent
- en données individuelles qui vont être analysées dans le projet de recherche pour laquelle la durée de conservation est de 5 ans à la fin de la recherche
- en données informatives sur la recherche sans données personnelles (protocole, résultats de la recherche: publications,données anonymisées, etc...) pour lesquelles un stockage indéfini est prévu

---

Durée de conservation des jeux de données

L'ensemble des données personnelles est conservé jusqu'à 5 ans suivant la dernière publication des résultats.
les données non personnelles sont conservées indéfiniment après la recherche

---

Espace de conservation des données

les données personnelles sont conservées sur un réseau privé sécurisé de l'IRSN (réseau Z).
les données non personnelles sont stockées sur le réseau général de l'IRSN (réseau O)

Y-a-t-il besoin de ressources supplémentaires particulières pour la conservation des données?

Non

Référencement des données dans le catalogue des données de l'institut

sans objet

Exigences légales et éthiques, code de conduite

Nature des données traitées dans le projet: s'agit-il de ?

- Données à caractère personnel
- Données de santé
- Données à caractère personnel
- Données de santé

Si des données personnelles sont contenues dans les données du projet, peuvent-elles être:

- pseudonymisées au moment de la collecte
- pseudonymisées au moment de la collecte

Y -a-t-il besoin de traiter des considérations éthiques ou juridiques par rapport aux données?

non

Y-a-t-il des questions relatives à la propriété intellectuelle (propriété, droits d'usage,...)?

non

Partage des données en Open data

Les données du projet de recherche peuvent-elles être publiées en Open Data?

Non, car elles sont seulement pseudonymisées et couvertes par des accords CNIL restreints d'utilisation pour les données françaises.
l'accord signé avec le NCI pour l'utilisation des données américaines ne couvre pas l'ouverture en open data.
les données coréennes, étant elles aussi pseudonnymisées ne peuvent être diffusées en open data

---

Constitution des jeux de données mis en Open data

sans objet

Données F.A.I.R.

Choix des métadonnées pour des données Faciles à trouver

sans objet

Attribution d'un identifiant unique et permanent de chacun des jeux de données

sans objet

Qui a créé le(s) jeux de données ?

sans objet

Interopérabilité technique des données

données brutes : format DAT, CSV
données texte : DOCX

Quelles licences pour les données réutilisables?

sans objet

Coût des dépôts de données et coût du maintien des données dans les entrepôts de données?
Ces coûts ont-ils été évalués?
Sont-ils intégrés dans le budget du projet ou non?
Par qui seront-ils portés, le cas échéant?

Les données sont déposées dans un entrepôt de données, le CASD, qui permettra de recevoir les données de mortalité de la cohorte française ORICAMS qui est intégrée dans le projet BECOME. Les analyses seront réalisées sur le serveur du CASD car les données de mortalité ne peuvent pas être extraites de ce serveur. Le service est évalué à environ 7000 euros par an.

Ce coût n'est pas intégré dans le budget du projet. Ce coût additionnel sera supporté par l'IRSN dans le cadre du budget pour la maintenance des bases de données au LEPID