

LOGO partenaire

**Titre du projet**

Titre court/acronyme

Version n°X – DATE

**Porteur de l’étude[[1]](#footnote-1) : Institut National du Cancer**

Président de l’Institut National du Cancer

52, avenue André Morizet

92 513 Boulogne Billancourt Cedex

A compléter si autre porteur

**Responsable du traitement des données[[2]](#footnote-2) :**

**Institut National du Cancer**

52, avenue André Morizet

92 513 Boulogne Billancourt Cedex

A compléter si autre RT

**Type d’étude :** *À compléter Etude épidémiologique observationnelle*

**CONSIGNES**

Le document est en mode modèle (.odtx). Lors du premier enregistrement il faudra définir le nom du document qui sera enregistré au format word (.docx).

Les éléments *en italique surlignées en jaune* sont des consignes qui peuvent être levées au fur et à mesure de la rédaction du document.

En fonction de l’étude, certaines sections peuvent ne pas avoir d’intérêt. Par exemple, la section sur les modalités de recueil de données qui n’a lieu d’être que lorsqu’un retour aux dossiers des personnes ou un questionnaire est prévu. De même la section sur les méthodes et critères d’appariement ne concerne que les études pour lesquelles plusieurs sources de données sont associées. Ces sections peuvent être supprimées.

**/!\Dans le doute** : laisser la section, elle sera supprimée par la suite.

Le document est formaté et **utilise des styles prédéfinis** :

 Pour le texte : Normal

 Pour les titres : 1 TITRE 1 (premier niveau) ; 1 .1 Titre 2 (deuxième niveau) ; 1.1 Titre 3 (troisième niveau)

 Pour les listes à puce : Puce niv 01

 Pour les listes numérotées : 1. Liste à numéro (si vous voulez recommencer la numérotation à 1, sélectionner la ligne concernée, clic de droit avec la souris et cliquer sur « Recommencer la numérotation).

Pour utiliser un style : sélectionner le texte puis cliquez sur le nom du style

Recommandation : les protocoles devant être soumis à l’avis du Cesrees devront rester en dessous des 20 pages (hors annexes)

**Périodes de l’étude :**

Il est important de distinguer différentes périodes de l’étude afin d’éviter tout problème de compréhension.

**1/ Période d’inclusion : (ou période de ciblage cf. définition HDH)**

Période d’inclusion : période choisie pour appliquer les critères de sélection d’une population.

Il s’agit d’identifier la période au cours de laquelle toutes les personnes devant participer à l’étude sont sélectionnées.

Exemple :

* Tous les nouveaux cas de cancer de l’estomac identifiés entre le 1er janvier 2018 et le 31 décembre 2021

**2/ Période nécessaire à la qualification des critères d’inclusion :**

Il s’agit d’identifier la période nécessaire pour s’assurer que chaque personne incluse réponde bien aux critères d’inclusion. Pour éviter les biais de suivi différentiel, il est préférable de considérer un délai. Ce délai peut être à cheval par rapport à la date d’inclusion de la personne lorsqu’un évènement nécessaire à l’inclusion est postérieur à la date d’inclusion. Attention dans ce cas de bien s’assurer que l’évènement postérieur n’influence pas le critère de jugement.

Exemples :

* Afin de s’assurer que le cas est bien un cas incident, recherche dans les 2 ans (soit 730 jours) précédents la date d’inclusion que la personne n’a pas eu de traitement à visé carcinologique
* Une étude portant sur le cancer du sein ayant pour date d’inclusion une mammographie, prend en compte des informations relatives aux hospitalisation dans les 3 mois (90 jours) suivant la date de mammographie afin de caractériser le cancer (le diagnostic de cancer n’étant pas associé à l’information disponible relative à la mammographie)

**3/ Période de suivi des personnes : période sur laquelle porte le recueil d’information.**

Il s’agit de la période de suivi de chaque personne incluse dans l’étude, c’est-à-dire la période nécessaire pour l’établissement des évènements constituant les critères de jugement. Pour éviter des biais de suivi différentiel, il est préférable de considérer une durée à partir de la date d’inclusion. La période de suivi peut être à cheval par rapport à la date d’inclusion de la personne lorsqu’un évènement étudié survient avant l’évènement caractérisant l’inclusion.

Exemple :

* Afin d’étudier les trajectoire thérapeutiques, suivi pendant 3 ans (1 05 jours) à partir de la date d’inclusion de chaque personne
* Afin d’identifier le parcours de soins des personnes ayant conduit au diagnostic de cancer (la date d’inclusion étant la date de diagnostic), suivi quatre mois (120 jours) avant et deux mois (60 jours) après la date d’inclusion des personnes

**4/ Période nécessaire à la qualification de co-facteurs et des critères de suivi  :**

Dans un certain nombre de cas il est nécessaire de reconstituer par des algorithmes des évènements, des indicateurs, des cofacteurs … Il s’agit donc de considérer la période nécessaire à la bonne application des algorithmes. Selon le type d’étude cette période est incluse dans la période de suivi.

Exemple :

* Afin de caractériser l’existence d’un diabète insulinorequérant au moment du diagnostic, un recul de 6 mois (180 jours) par rapport à la date d’inclusion est nécessaire. Il permet d’identifier les soins (traitements, hospitalisations, …) caractéristiques

**5/ Période d’extraction (ou historique des données cf. définition HDH) :**

Période qui couvre toutes les données nécessaires à l’étude, le cas échéant du début de la période nécessaire à la qualification des critères d’inclusion, jusqu’à la fin de la période nécessaire à la qualification de co-facteurs et des critères de suivi.

Point d’attention : Délai pour l’obtention et la mise à disposition des données

Les bases de données sont actualisées périodiquement (généralement annuellement). Lors de la conception de l’étude, il faut donc s’assurer que les délais et périodes identifiées dans le cadre du protocole soient compatibles.



**Matrice de responsabilité**

Une matrice de responsabilité ci-dessous permet de définir avec l’Institut le rôle de chaque partie dans le projet :

|  |
| --- |
| AIDE A LA DETERMINATION DE LA CO-RESPONSABILITE DE TRAITEMENT DES DONNEES A CARACTERE PERSONNEL |
|   | Critères | INCa | XXX | XXX | Commentaires |
| 1 | Initiative de mener le traitement des données |   |   |   |   |
| 2 | Détermination des finalités du traitement (qui a décidé de la mise en œuvre du traitement, du « pourquoi » le traitement est mis en place, de ses objectifs, du résultat attendu, des éléments de cadrage) |   |   |   |   |
| 3 | Détermination des sources de données nécessaires |   |   |   |   |
| 4 | Détermination des données nécessaires |   |   |   |   |
| 5 | Détermination de la population concernée |   |   |   |   |
| 6 | Expertise permettant de décider des moyens à mettre en place pour atteinte l’objectif du traitement (autonomie, marge de manœuvre importante) |   |   |   |   |
| 7 | Détermination de la méthode d'appariement (le cas échéant) |   |   |   |   |
| 8 | Détermination de la méthode d'analyse |   |   |   |   |
| 9 | Détermination des moyens financiers |   |   |   |   |
| 10 | Détermination des moyens techniques (disponibilité, pseudonymisation, traçabilité, absence d’accès non autorisés…) |   |   |   |   |
| 11 | Détermination des moyens humains |   |   |   |   |
| 12 | Maîtrise et contrôle des opérations de traitement |   |   |   |   |

**TABLE DES MATIÈRES**

[1 équipe projet 12](#_Toc161759368)

[1.1 Responsables de traitement et service chargé de la mise en œuvre des traitements 12](#_Toc161759369)

[1.2 Équipe coordinatrice 12](#_Toc161759370)

[1.3 Partenaires 13](#_Toc161759371)

[1.4 Comité d’experts 13](#_Toc161759372)

[1.5 Délégué à la protection des données (DPO) 13](#_Toc161759373)

[1.6 Organisme financeur et contractualisation 13](#_Toc161759374)

[2 contexte de l’étude, recherche ou évaluation 13](#_Toc161759375)

[3 Objectifs de l’étude 13](#_Toc161759376)

[3.1 Objectif principal de l’étude 13](#_Toc161759377)

[3.2 Objectifs secondaires 13](#_Toc161759378)

[4 Respect de l’éthique 13](#_Toc161759379)

[5 justification d’intérêt public de l’étude, la recherche ou l’évaluation 14](#_Toc161759380)

[6 Publication des résultats et valorisation 15](#_Toc161759381)

[7 Type d’étude 15](#_Toc161759382)

[8 Population concernée ou Cohorte 16](#_Toc161759383)

[8.1 Description de la population concernés par l’étude 16](#_Toc161759384)

[8.1.1 Critères d’inclusion 16](#_Toc161759385)

[8.1.2 Critères de non inclusion 16](#_Toc161759386)

[8.1.3 Critères d’exclusion 16](#_Toc161759387)

[8.2 Taille de la population – Nombre de sujets nécessaires – Effectifs prévisionnels 16](#_Toc161759388)

[9 Gestion des données 16](#_Toc161759389)

[9.1 Origine des données de santé à caractère personnel (Source des données) 17](#_Toc161759390)

[9.1.1 « Plateforme des données en cancérologie » 17](#_Toc161759391)

[9.2 Modalités de recueil des données 17](#_Toc161759392)

[9.3 Catégories de données recueillies (Ciblage des données pertinentes) 17](#_Toc161759393)

[9.3.1 Justifie l’utilisation de données dites sensibles : 18](#_Toc161759394)

[9.3.2 Spécificité des données de la Plateforme de données en cancérologie par rapport au SNDS 18](#_Toc161759395)

[9.3.3 Méthodes et critères d’appariement 19](#_Toc161759396)

[9.4 Préparation des données 19](#_Toc161759397)

[10 Analyses des données 20](#_Toc161759398)

[10.1 Critère de jugement principal 20](#_Toc161759399)

[10.2 Critère de jugement secondaire 20](#_Toc161759400)

[11 Limites de l’étude 20](#_Toc161759401)

[12 Calendrier prévisionnel, faisabilité du projet 20](#_Toc161759402)

[13 Information des personnes 21](#_Toc161759403)

[13.1 Information individuelle 21](#_Toc161759404)

[13.2 Information collective 21](#_Toc161759405)

[13.3 Droit des personnes 21](#_Toc161759406)

[14 Sécurité et circuit des données 22](#_Toc161759407)

[14.1 Gestion du risque d’identification des personnes ou de ré-identification des personnes 22](#_Toc161759408)

[14.1.1 Pseudonymisation (code personne) 22](#_Toc161759409)

[14.2 Hébergement et sécurité de la base de données (Support de données) 22](#_Toc161759410)

[14.3 Circuit des données à caractères personnels et modalité de protection de leur confidentialité 23](#_Toc161759411)

[14.3.1 Circuit des données 23](#_Toc161759412)

[14.3.2 Transmission des données 23](#_Toc161759413)

[14.3.3 Intégration des données 24](#_Toc161759414)

[14.4 Transfert des données de l’étude 24](#_Toc161759415)

[15 Bibliographie 24](#_Toc161759416)

[Annexes 25](#_Toc161759417)

[Annexe 1 : Transfert de données vers la Plateforme de données en cancérologie 25](#_Toc161759418)

[Annexe 2 : Présentation de la plateforme de données en cancérologie 29](#_Toc161759419)

[Annexe 2 : Aspects réglementaires 38](#_Toc161759420)

[1. Avis du Comité scientifique et éthique de la Plateforme de données en cancérologie, du CESREES et autorisation de la CNIL 38](#_Toc161759421)

[2. Accès aux données et confidentialité 38](#_Toc161759422)

[3. Audit 38](#_Toc161759423)

[4. Archivage des documents et Rapport final de l’étude 38](#_Toc161759424)

[5. Délibération de la CNIL - Délibération 2023-040 du 20/04/2023 39](#_Toc161759425)

[6. Délibération de la CNIL - Délibération 2019-083 du 20/06/2019 47](#_Toc161759426)

[7. Avis du Comité scientifique et éthique de la plateforme de données en cancérologie 52](#_Toc161759427)

**GLOSSAIRE**

*À compléter*

AAP Autorisation d’accès précoce

AAC Autorisation d’accès compassionnel

ALD Affection longue durée

ATU Autorisation temporaire d’utilisation

CCAM Classification commune des actes médicaux

CESREES Comité Éthique et Scientifique pour les Recherches, les Études et les Évaluations dans le domaine de la santé

CMUc Couverture mutuelle universelle complémentaire

CNAM Caisse nationale de l’assurance maladie

CNIL Commission nationale de l’informatique et des libertés

CRCDC Centres régionaux de coordination des dépistages des cancers

CSP Code de la santé publique

CSS Complémentaire santé solidaire

DCIR Datamart de consommation inter-régime

HAD Hospitalisation à domicile

HDH Health data hub (plateforme de données de santé)

Institut Institut national du cancer

MCO Médecine, chirurgie, obstétrique

PMSI Programme de médicalisation des systèmes d’informations

RGPD Règlement général sur la protection des données

RIM-P Recueil d'information en psychiatrie

SNDS Système national des données de santé

SSR Soins de suite et de réadaptation

**Résumé du protocole**

*Rédaction conjointe Institut/Partenaire une fois le protocole finalisé*

**Titre de l’étude :**

**Équipe projet :**

**Contexte – justification de l’étude :**

**Objectif principal :**

**Objectif secondaire :**

**Ethique :** L’étude est une étude observationnelle, à partir de bases de données constituées par ailleurs, sans intervention sur la personne. L’étude sera conduite conformément aux dispositions règlementaires, en respectant les conditions d’accès aux données issues du SNDS. Les résultats seront publiés et mis à dispositions selon les règles définies par le HDH.

**Intérêt général** : Fréquence de la pathologie ; Sévérité de la pathologie ; Lien avec la Stratégie décennale de lutte contre les cancers ; Lien avec les missions de l’Institut …

**Publication et valorisation des résultats** **:** Les résultats de l’étude feront l’objet de publications scientifiques dans des revues à comité de lecture. Un rapport d’analyse sera également produit et transmis au Health data hub comme pour toutes exploitation portant sur des données issues du SNDS.

**Origine / Source des données :**

**Nombre de sujets nécessaires/taille de la population :** Soit sur base de données – exhaustif ; soit sur calcul

**Nombre de centres participants :** Si étude nécessitant un recueil de données ad’hoc

**Nombre moyen d’inclusions par centre :** Si étude nécessitant un recueil de données ad’hoc

**Critères d’inclusion / non-inclusion / exclusion (cohorte étudiée) :**

**Critère de jugement principal / secondaire :**

**Type d’étude et Méthodologie :**

**Calendrier de l’étude (en Mois – sans date) :**

**Période d’inclusion :**

**Information individuelle des patients :** oui/non ; préciser les choix et justifier la demande de dérogation

# équipe projet

|  |  |  |
| --- | --- | --- |
| Equipe | Service | Rôle dans le projet |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

## Responsables de traitement et service chargé de la mise en œuvre des traitements

*Rédaction conjointe Institut/Partenaire*

*Le responsable de traitement est le promoteur de l’étude. Il détermine les finalités et les moyens d’un traitement, c’est à dire l’objectif et la façon de le réaliser. Une étude peut avoir plusieurs responsables de traitement.*

*Le service chargé de la mise en œuvre des traitements est l’organisme ayant la charge et la responsabilité d’effectuer les analyses pour le compte du responsable de traitement.*

Responsable de traitement : *Organisme, Responsable légal, adresse*

Service chargé de la mise en œuvre des traitements : *Organisme, Responsable légal, adresse, nom du Service le cas échéant*

## Équipe coordinatrice

*Rédaction conjointe Institut/Partenaire. L’objectif est de présenter les équipes qui pilotent de façon opérationnelle l’étude*

*Préciser les liens entre les différentes équipes, le conventionnement prévu afin que le lecteur sache qui fait quoi*

Équipe 1 :

Département Données et évaluation en santé

Direction de l’Observation, des Sciences des données, et de l’Évaluation

Institut national du cancer

## Partenaires

## Comité d’experts

## Délégué à la protection des données (DPO)

La DPO est Laurence Loupiac, service juridique, Institut national du cancer.

Ajouter le DPO du porteur de l’étude

## Organisme financeur et contractualisation

Une convention de responsabilité conjointe ou de sous traitance (supprimer la mention inutile) sera signée entre les différentes parties une fois l’ensemble des autorisations obtenues. La convention précise le rôle et les responsabilités de chaque organisme (responsable de traitement, service de la mise en œuvre des traitements, équipe collaboratrice …) ainsi que les modalités de partage et d’accès aux données.

# contexte de l’étude, recherche ou évaluation

*À compléter par le partenaire avec références bibliographiques intégrées sous format Zotero®, EndNote® ou équivalent*

# Objectifs de l’étude

*À compléter par le partenaire*

## Objectif principal de l’étude

## Objectifs secondaires

# Respect de l’éthique

Cette étude est un traitement de données à caractère personnel ayant pour finalité la réalisation de recherches, études et évaluations dans le domaine de la santé ne répondant pas à la définition des recherches impliquant la personne humaine telles que définies à l'article L. 1121-1 du CSP (code de la santé publique). A ce titre, aucune intervention n’est effectuée sur la personne et les données exploitées ont été collectées par ailleurs.

Si passage au Cesrees et à la Cnil (sinon paragraphe à supprimer)

Préalablement à la réalisation de l’étude, le protocole sera soumis au Comité scientifique et éthique de la plateforme de données en cancérologie, et le cas échéant au Comité éthique et scientifique sur les recherches, études et évaluations dans le domaine de la santé (Cesrees) et à la Commission nationale de l’informatique et des libertés (Cnil).

Si conforme à la décision unique de l’Institut (sinon paragraphe à supprimer)

Préalablement à la réalisation de l’étude, le protocole sera soumis au Comité scientifique et éthique de la plateforme de données en cancérologie. Le projet relevant de l’autorisation unique délivrée par la CNIL à l’Institut il ne sera pas soumis au Comité éthique et scientifique sur les recherches, études et évaluations dans le domaine de la santé (Cesress) et à la Commission nationale de l’informatique et des libertés (Cnil).

*A supprimer si on ne mobilise pas les données du SNDS :*

Conformément aux dispositions relatives aux SNDS (Système national des données de santé) mentionnées au V de l’article L. 1461-1 du CSP, les traitements portant sur ces données n’auront pas pour objet une des deux finalités interdites suivantes :

1° La promotion des produits mentionnés au II de l'article L. 5311-1 en direction des professionnels de santé ou d'établissements de santé ;

2° L'exclusion de garanties des contrats d'assurance et la modification de cotisations ou de primes d'assurance d'un individu ou d'un groupe d'individus présentant un même risque.

Conformément aux dispositions de la loi « informatique et libertés », la présentation des résultats du traitement de données ne pourra en aucun cas permettre l'identification directe ou indirecte des personnes concernées par la recherche.

Les personnes seront informées de l’étude (conformément aux dispositions précisées dans le présent document). Ceci comprendra notamment la publicité de l’étude sur le site de l’Institut national du cancer (<https://lesdonnees.e-cancer.fr>).

# justification d’intérêt public de l’étude, la recherche ou l’évaluation

*À compléter par le partenaire avec références bibliographiques intégrées sous format Zotero®ou EndNote*

*Éléments potentiels : Fréquence de la pathologie ; Sévérité de la pathologie ; Lien avec la Stratégie décennale de lutte contre les cancers ; Lien avec les missions de l’Institut …*

*Ne pas oublier de mettre ce que cela apporte pour les personnes malades, ayant été atteintes d’un cancer ou à risque de cancer, leur proche …*

# Publication des résultats et valorisation

*Rédaction conjointe Institut/Partenaire*

L’Institut national du cancer publie régulièrement les résultats des études conduites à partir de la plateforme de données en cancérologie en s’appuyant d’une part sur son site internet (<https://e-cancer.fr>) et d’autre part sur sa collection de rapports.

Les résultats de ces travaux font également l’objet de publications dans des revues scientifiques internationales à comité de lecture et de communications dans les congrès appropriés.

Si conforme à la décision unique de l’Institut (sinon paragraphe à supprimer)

L’Institut propose également d’appliquer les dispositions prévues par la décision unique de la CNIL n 2019-083 du 20 juin 2019, à savoir :

* un enregistrement préalable de l’étude auprès du Health Data Hub – de la Plateforme de données de Santé (HDH) ;
* à la fin de l’étude, la méthode et les résultats obtenus seront être communiqués au HDH en vue de leur publication par ce dernier selon des modalités définies par ses soins.

Si passage au Cesrees et à la Cnil et portant sur les données du SNDS de la PDC (sinon paragraphe à supprimer)

L’étude étant menée sur des données issues du SNDS, les obligations de transparence prévues à l’article L.1461-3 du Code de la santé s’appliquent, à savoir :

* un enregistrement préalable de l’étude auprès du Health Data Hub – de la Plateforme de données de Santé (HDH) ;
* à la fin de l’étude, la méthode et les résultats obtenus seront être communiqués au HDH en vue de leur publication par ce dernier selon des modalités définies par ses soins.

L‘étude étant conjointement réalisée, toutes les publications feront référence à l’étude XXX (REF° et à la plateforme de données en cancérologie (REF = celle Bousquet PJ).

Les règles de publications suivront les normes dites de Vancouver, en veillant à associer l’ensemble des équipes participant au projet.

# Type d’étude

*À compléter par le partenaire*

# Population concernée ou Cohorte

*À compléter par le partenaire*

*Prendre en compte les notions de périodes (cf. consignes en début de document)*

*Dans le cas d’étude avec des objectifs multiples, s’assurer que pour chacun des objectifs les points suivants sont valables (faire plusieurs sections si nécessaires)*

|  |  |
| --- | --- |
| Période d’inclusion (ou de ciblage)  | Du au  |
| Période nécessaire à la qualification des critères d’inclusion | Du au  |
| Période de suivi  | Du au  |
| Période nécessaire à la qualification de co-facteurs et des critères de suivi | Du au  |
| Période d’extraction des données | Du au  |

## Description de la population concernés par l’étude

*A compléter par le partenaire : Si plusieurs sources de données sont mobilisées, préciser la ou les sources qui serviront pour l’identification des personnes et faire un renvoi vers la section correspondant aux sources de données*

### Critères d’inclusion

*Préciser âge et régime de protection sociale*

### Critères de non inclusion

### Critères d’exclusion

## Taille de la population – Nombre de sujets nécessaires – Effectifs prévisionnels

*Rédaction conjointe Institut/partenaire : Si l’étude est basée sur une hypothèse, préciser le nombre de personnes, s’il s’agit d’un échantillonnage, indiquer le nombre de personnes et le taux d’échantillonnage, si l’étude est exhaustive, indiquer le nombre de personnes ciblées.*

# Gestion des données

*Rédigé par l’Institut si uniquement plateforme de données en cancérologie ; sinon préciser les sources complémentaires.*

## Origine des données de santé à caractère personnel (Source des données)

*Préciser toutes les sources et apporter un descriptif par sources. S’assurer que le descriptif se suffise à lui-même pour qu’une personne tierce puisse comprendre quelles sont les données mobilisées et leur provenance. Considérer chaque source indépendamment.*

### « Plateforme des données en cancérologie »

La plateforme de données en cancérologie est présentée en annexe. Elle comprend à ce jour les données issues du SNDS portant sur des personnes atteintes d’un cancer ou à risque élevé de cancer. Ces données seront progressivement complétées par les données issues des registres des cancers, du dossier communicant de cancérologie, des résumés de passage aux urgences et des dépistages organisés. Des données issues d’études conduites par les industries de santé pourront également être intégrées à la plateforme de données en cancérologie. Les premières personnes ont été incluses en 2010. Chaque personne est suivie pendant 20 ans.

## Catégories de données recueillies (Ciblage des données pertinentes)

*Préciser les informations d’intérêt et supprimer les catégories inutilisées*

Les catégories de données utilisées sont les suivantes :

* Informations relatives à l’état civil de la personne (sexe, mois et année de naissance)
* Informations relatives au statut vital et au décès (mois et années, lieu et causes médicales du décès) ;
* Informations relatives au lieu de vie, aux habitudes de vie, à la défavorisation sociale (Indice de défavorisation sociale - Fdep) ou pouvant contribuer à l’étude des inégalités sociales et géographiques et d’ordre économique ;
* Informations relatives à la vie professionnelle
* Informations médicales
* Informations relatives à des prélèvements biologiques identifiants (données génétiques)
* Information médicosociales et médico-économiques relatives au parcours en santé (incluant la prévention, le dépistage, le diagnostic, les soins hospitaliers, les soins de ville, le suivi et l’après cancer);
* Informations médico-sociales relatives à la situation des personnes en situation de handicap ;
* Informations médico-administratives relatives aux arrêts de travail et aux prestations en espèces, aux organismes d’assurance maladie obligatoire et, s’il y a lieu, aux organismes d’assurance maladie complémentaire intervenant dans la prise en charge financière du bénéficiaire des soins et prestations
* Informations relatives aux professionnels (mois et année de naissance, sexe, lieux, modes d’exercices…) et services de santé et sociaux (type, lieux…) intervenant dans le parcours de santé des personnes.

*Partie à préciser*

### Justifie l’utilisation de variables dites sensibles :

Mois, année de naissance : Cette information permet d’une part l’appariement des personnes entre les différentes sources, et d’autre part à définir l’âge au moment de la consommation des soins

Date de soins : Cette information permet d’une part l’appariement des personnes entre les différentes sources, et d’autre part de définir les délais entre les différents soins, et la survie

Mois, année de décès : cette information permet d’identifier la censure d’une personne (délai de suivi), d’identifier les consommations de soins qui sont proches du décès et les délais entre les soins et le décès, et de définir l’âge au décès

Commune de résidence : Cette information permet l’appariement des personnes entre les différentes sources et d’attribuer un indice de défavorisation (FDep) aux personnes.

Commune de décès : Cette information permet l’appariement des personnes entre les différentes sources

### Spécificité des données de la Plateforme de données en cancérologie par rapport au SNDS (à supprimer si on ne mobilise pas les données du SNDS)

Des variables ou indicateurs complémentaires, ainsi que des tables de données et des référentiels, déduits des informations du SNDS ont été créés par l’Institut afin de mieux caractériser les personnes, leur cancer et leurs consommations de soins. En effet ces informations se trouvent dispersées, évoluent dans le temps, et leur recherche est consommatrice de temps et source d’erreur. Afin d’éviter aux utilisateurs de devoir refaire ces exploitations à partir des données sources, leur faire gagner du temps d’exploitation et pour garantir une homogénéité de traitement de la donnée, l’Institut extrait, qualifie, hiérarchise, date et historise les informations issues de ces différentes sources et les restitue dans des tables mises à disposition des utilisateurs de la plateforme.

Ces tables sont mises à jour chaque année lors de l’actualisation des données. Elles s’appuient sur les consommations de soins de ville (incluses dans le DCIR – datamart de consommation inter-régime), à l’hôpital (PMSI MCO – médecine, chirurgie, obstétrique, HAD – Hospitalisation à domicile, SSR – Soins de suite et de réadaptation, ou RIM-P Recueil d'information en psychiatrie) et les ALD (affection longue durée) …

Les variables ou indicateurs définis par l’Institut sont relatifs :

* Aux bénéficiaires : sexe, mois-année de naissance, date de décès, régime d’assurance maladie, bénéficiaire de la CMU-C/CSS (Couverture mutuelle universelle complémentaire et Complémentaire santé solidaire), code commune de résidence, bénéficiaire d’une ALD pour un cancer…
* Aux comorbidités dont le score de Charlson…
* Aux consommations de soin en rapport avec le cancer : chirurgie dont exérèse, hospitalisations pour/avec chimiothérapie, hospitalisations pour/avec radiothérapie, acte de radiothérapie en secteur libéral, consommation de médicament anti-cancéreux en liste en sus, en rétrocession, en ville, soins palliatifs en MCO, SSR ou HAD, hospitalisation post-chirurgie en SSR, acte spécifique d’anatomopathologie en ville…
* A la caractérisation du cancer du bénéficiaire pour chaque année depuis l’inclusion de la personne dans la cohorte : type de cancer (in situ, invasif, envahissement ganglionnaire ou métastatique), localisation du cancer (appareil, organe), sources ayant permis d’identification du cancer (ALD, PMSI MCO, HAD ou SSR), mode de prise en charge en MCO (chirurgie, chimiothérapie, radiothérapie). Cela permet de repérer les cas incidents d’une année donnée. Une confrontation des différents codes d’une même personne permet de corriger certains codages tels que de probables métastases codées comme de tumeurs primitives ou des localisations non précises.

L’utilisateur garde la possibilité de travailler sur les données avant correction et sur les données sources ayant permis la création des variables ou indicateurs complémentaires.

# Analyses des données

*Rédaction conjointe Institut/Partenaire*

*Si des algorithmes doivent être utilisés, préciser ces algorithmes ; s’ils doivent être développés, l’indiquer*

## Critère de jugement principal

## Critère de jugement secondaire

# Limites de l’étude

# Calendrier prévisionnel, faisabilité du projet

*Rédaction conjointe Institut/Partenaire*

# Information des personnes

## Information individuelle

*A compléter si une information individuelle est prévue, sinon supprimer la section. Attention dans ce cas à une éventuelle demande de dérogation à l’information des personnes (notamment si celle-ci n’a pas été effectuée par ailleurs).*

*Attention : considérer chaque source de données indépendamment*

*La dérogation à l’information des personnes est propre à chaque source*

*La lettre d’information d’une source peut, si elle le mentionne, permettre l’accès à un ensemble de source (et ainsi assurer l’information des personnes pour la présente étude)*

## Information collective

Dans la mesure où l’Institut ne dispose pas de l’identité et de l’adresse des personnes concernées et compte tenu du nombre de personnes concernées, de l’incertitude sur leur statut vital et de l’ancienneté d’une partie des données, une information individuelle de l’ensemble des personnes concernées exigerait un effort disproportionné. Dès lors, conformément aux autorisations CNIL de la plateforme référencées en annexe et conformément à l’article 14 paragraphe 5 point b) du RGPD, l’Institut met à disposition sur son site internet une information collective relative à la Plateforme ainsi que des informations sur les projets de recherche ultérieurs menés à partir des données qu’elle contient et comportant l’ensemble des mentions prévues par l’article 14 du RGPD.

## Droit des personnes

*A compléter en fonction des sources de données*

Conformément au Règlement général sur la protection des données 2016/679 (RGPD) et à la loi informatique modifiée, les personnes concernées disposent d’un droit d’accès, droit d’opposition, de rectification, d’effacement et d’un droit à la limitation du traitement de leurs données dans les conditions fixées par le RGPD. Les personnes concernées disposent aussi du droit de prévoir des directives sur le sort de leurs données après leur décès et de formuler une réclamation auprès de la CNIL si elles estiment que leurs droits n’ont pas été respectés.

Conformément à la délibération n°2023-040, l’Institut national du cancer ne disposant pas de l’identité des usagers, les personnes concernées qui souhaiteraient exercer leurs droits seront sollicitées par l’Institut après leur demande afin de fournir des informations complémentaires afin que l’Institut tente une réidentification au sein de la PDC. En cas de refus de fournir ces informations complémentaires ou en cas d’échec de la réidentification, les personnes concernées pourront exercer leurs droits en contactant directement les responsables de traitement des organismes à l’origine de la collecte des données concernées (la CNAM pour le SNDS, registres des cancers ou structures de gestion des dépistages). Une procédure a été développée afin de leur permettre de faire valoir leurs droits, présentée sur le site internet de l’Institut.

Les coordonnées de l’Institut, de son représentant légal et de sa déléguée à la protection des données sont disponibles sur le site <https://www.e-cancer.fr/Politique-de-protection-des-donnees-a-caractere-personnel>. Les usagers ont également la possibilité de contacter par écrit le service juridique de l’Institut national du cancer, 52 avenue André Morizet, 92 100 Boulogne Billancourt afin de lui faire part de leur demande ou par mail à l’adresse suivante dpo@institutcancer.fr.

Si un usager estime, après avoir contacté les différents organismes à l’origine de la collecte des données concernées ou l’Institut, que ses droits ne sont pas respectés, il peut adresser une réclamation en ligne à la CNIL ou par voie postale.

# Sécurité et circuit des données

## Gestion du risque d’identification des personnes ou de ré-identification des personnes

### Pseudonymisation (code personne)

Pour chaque source de données de la plateforme de données en cancérologie, un code personne unique (pseudonyme) est attribué à chaque personne par l’organisme responsable de la source de données. Ce code est différent du numéro d’identification utilisé par l’organisme dans ses propres bases de données. L’organisme conserve une liste permettant de relier le numéro d’identification et le pseudonyme. Elle est complétée à chaque nouvelle livraison de données à l’Institut. Cette liste n’est pas transmise à l’Institut.

Le pseudonyme unique (code personne) est composé de trois parties codées sur 16 caractères alphanumériques :

* 1ère partie : « xxx- » pour identifier la source ;
* 2ème partie : année de première identification de la personne codée sur 4 caractères ;
* 3ème partie : numéro d’identifiants fonctionnant sur un principe d’incrémentation codé sur 8 caractères ;

Exemple : 34G-201200001413

## Hébergement et sécurité de la base de données (Support de données)

Le système d’information de la plateforme de données en cancérologie a été homologué (décision n°2023-27 du 3 avril 2023 de l’autorité d’homologation de l’Institut national du cancer portant homologation du système d’information intitulé « Plateforme des données en cancérologie »). Toute personne accédant à la plateforme signe un document l’engageant, conformément notamment à l’article 32 du Règlement général sur la protection des données à prendre toutes précautions utiles dans le cadre de ses missions afin de protéger la confidentialité des informations auxquelles il a accès, et en particulier d’empêcher qu’elles ne soient modifiées, endommagées ou communiquées à des personnes non expressément autorisées à recevoir ces informations.

La plateforme est hébergée chez un hébergeur de données de santé (HDS) intégrant la sphère « hébergement souverain européen » (Circulaire n° 6282-SG du 5 juillet 2021 relative à la doctrine d’utilisation de l’informatique en nuage par l’État) et SNDS.

## Circuit des données à caractères personnels et modalité de protection de leur confidentialité

Les analyses sont réalisées sur la plateforme de donnés en cancérologie de Institut.

Aucun export de données autres que des données agrégées sous forme de résultats ne pourra être réalisé.

*Section à adapter ou supprimer en cas de chaînage entre différentes sources d’information*

### Circuit des données

L’intégration des données d’accès précoce dans la PDC suit le schéma suivant :

1. Extraction des données à partir du système d’information (SI) du producteur XXXXX
2. Transmission à l’Institut selon le protocole défini en 14.3.2
3. Appariement à la plateforme de données en cancérologie
4. Validation de l’appariement

Si pour des raisons de mauvaise qualité des données l’appariement n’est pas validé, une nouvelle extraction peut être effectuée (retour à l’étape 1) et ainsi de suite jusqu’à obtention d’un appariement satisfaisant ou décision de l’équipe projet de finaliser le projet avec le groupe de personnes ayant pu être appariées.

### Transmission des données

Le transfert de données doit obligatoirement porter sur des données préalablement chiffrées et suivre les recommandations en vigueur sur le niveau de chiffrement. Actuellement le système de chiffrement asymétrique utilisé est celui de GNU Privacy Guard (GPG). Il est accessible via des outils comme « Kleopatra » présent sur le site GPG4WIN. Le chiffrement hybride peut être utilisé pour de gros volumes de données.

Les destinataires des fichiers de données sont nommément identifiés. Le déchiffrement des données sera réalisé par le département des systèmes d’information de l’Institut, puis mis à disposition des équipes traitant les données.

Le support de transfert est un disque dur externe ou un serveur FTP sécurisé ou un outil sécurisé de transfert de données en ligne selon la volumétrie des données.

Une sauvegarde des données sources sera réalisée avant chaque intégration.

Le détail des étapes du transfert est fourni en annexe.

### Intégration des données

Une fois mises à disposition dans un espace de production, les données seront nettoyées et intégrées en base en lecture seule pour les utilisateurs. Ces étapes incluent une procédure de recette (tests d’acceptation) portant sur la qualité des données (vérification des volumétries, contrôles de cohérence, recherche de valeurs aberrantes…).

Les données une fois intégrées et nettoyées seront sauvegardées.

## Transfert des données de l’étude

Les données de l’étude ne seront pas transférées (pas de transfert en dehors de l’Union Européenne).

# Bibliographie

# Annexes

#### Présentation de la plateforme de données en cancérologie

|  |  |
| --- | --- |
|  http://interne.e-cancer.fr/images/stories/INCaNet/INCaNet/nouveau-logo-inca_2015.jpg | **La Plateforme de données en cancérologie** |
| **Direction de l’Observation, des Sciences des données et de l’Évaluation** |

La plateforme de données permet de regrouper les données issues de différents producteurs de données, tout en sécurisant les outils de productions dans le respect des nouvelles exigences règlementaires, et en mettant à disposition les outils nécessaires à leur exploitation.

Actuellement l’Institut dispose d’un entrepôt de données homologué Système national des données de santé (SNDS) et regroupant des données du SNDS, et des registres des cancers

L’Institut est par ailleurs autorisé à enrichir la plateforme avec :

* des données issues des [dépistages organisés](https://www.e-cancer.fr/Professionnels-de-sante/Depistage-et-detection-precoce) du cancer du sein, du cancer colorectal et du cancer du col de l’utérus issues des Centres régionaux de coordination des dépistages des cancers (CRCDC). Ces données seront intégrées à la Plateforme lorsqu’un système uniformisé de gestion des données aura été créé.
* Des données issues des dossiers médicaux composant le dossier communiquant de cancérologie (comptes-rendus d’anatomocytopathologie, comptes-rendus de génétique moléculaire, fiches RCP, Programme personnalisé de soins, Programme personnalisé après cancer) ;
* Des données issues de la base Oscour : les résumés de passage aux urgences ;
* Des données d’études pour lesquelles l’Institut est responsable ou responsable conjoint du traitement des données ou les données d’études réalisés par des industriels de santé dans le cadre du projet de [Filière intelligence artificielle et cancer](https://www.e-cancer.fr/Institut-national-du-cancer/Filiere-Intelligence-Artificielle-et-Cancers).

La plateforme permet de :

* Respecter le référentiel de sécurité du SNDS ;
* Sécuriser la production de données auprès de certains acteurs comme les registres ou des bases clinico-biologiques en leur mettant à disposition des outils répondant aux référentiels de sécurité ;
* Favoriser le déploiement des outils de production des documents interopérables : fiches de Réunion de Concertation Pluridisciplinaire, Parcours Personnalisé de Soins, Compte-rendu d’anatomopathologie, compte-rendu de génétique moléculaire… ;
* Créer une interface d’analyse et de restitution des analyses et résultats ;
* Proposer pour un certain nombre d’acteurs un environnement sécurisé de recherche (« bulle sécurisé »).

La plateforme joue ainsi un rôle d’agrégateur de données et de facilitateur pour de nouvelles exploitations, en s’appuyant sur une documentation partagée, l’identification et la définition de référentiels et de thésaurus.

Elle vise à répondre aux questionnements des acteurs de la cancérologie en leur offrant une plateforme de données, probablement unique au monde, afin d’en faire un outil de recherche publique et privée, pour les acteurs partenaires et académiques, de la recherche et du soin. De manière complémentaire, des jeux de données traitées et qualifiées pourront être mis à disposition dans le Hub date santé, plateforme de ré-exploitation plus générale de données de santé non identifiantes.

Trois objectifs :

* Créer une plateforme de données en oncologie compatible avec le Health data hub ;
* Faire émerger des projets innovants par et pour la recherche publique ou privée ;
* Renforcer la cohérence et valoriser l’ensemble de la démarche, en particulier à l’international.

**La cohorte cancer**

**Une extraction du SNDS…**

La cohorte cancer est une extraction du SNDS pour toute personne identifiée comme ayant ou ayant eu une tumeur solide ou une hémopathie maligne (cas incident ou prévalent) depuis 2010. Elle compte à ce jour 9,5 millions de personnes. Une personne entre dans la cohorte dès lors qu’elle a une hospitalisation en secteur MCO (médecine, chirurgie, obstétrique) considérée en lien avec le cancer, une ALD (affection longue durée) pour cancer, de la radiothérapie en secteur libéral (la radiothérapie publique donne lieu à une hospitalisation et sera identifiée via la PMSI – programme de médicalisation des systèmes d’information), le remboursement d’un médicament anticancéreux en ville ou en rétrocession ou certains actes CCAM (classification commune des actes médicaux) d’anatomopathologie [1].

Toutes les consommations de soins quelles qu’elles soient des personnes incluses sont extraites de l’année précédant l’inclusion jusqu’aux 20 années suivantes. La durée du suivi a pour but d’étudier :

* La survie et les récidives à long terme pour les cancers de bon pronostic
* Les effets secondaires à long terme des traitements
* La survenue de seconds cancers
* Le devenir à l’âge adulte des personnes atteintes de cancer pendant l’enfance.

Ainsi, la cohorte apporte de la connaissance pour permettre une meilleure planification des soins, l’évaluation des prises en charge et l’évaluation de politiques de santé. Elle apporte en effet une vision à la fois transversale selon différents axes (période, type de cancer, âge, zone géographique …) et longitudinale et permet un suivi régulier des activités, globale ou par axe, par la production de tableaux de bord thématiques : hospitalisation, chimiothérapie, radiothérapie, chirurgie, populations spécifiques (enfants, personnes âgées, cancers rares …).

Elle permet aussi l’analyse des parcours de soins [2-4], le recours à de nouveaux médicaments à condition qu’ils soient inscrits sur la liste en sus, en ATU (autorisation temporaire d’utilisation), en AAP ou AAC (autorisation d’accès précoce ou compassionnel), en rétrocession ou délivrés en officine.

**… Enrichie**

Des variables complémentaires, déduites des informations du SNDS ont été créées afin de mieux caractériser à la fois le cancer - type (in situ, invasif, envahissement ganglionnaire ou métastatique) ou localisation (appareil, organe) - la personne et sa consommation de soins par année. En effet ces informations se trouvent dispersées, évoluent dans le temps, et leur recherche est consommatrice de temps et source d’erreur. Afin d’éviter aux utilisateurs de devoir refaire ces exploitations à partir des données sources et pour garantir une homogénéité de traitement de la donnée, l’Institut extrait, qualifie, hiérarchise, date et historise les informations issues de ces différentes sources et les restitue dans des tables mises à disposition des utilisateurs de la plateforme.

L’appariement avec d’autres bases de données a pour objectif d’adjoindre des informations qui font actuellement défaut (cliniques, anatomopathologiques, biologiques, génomiques…). Dans un premier temps certaines de ces données seront issues des bases des registres des cancers et du dépistage avant d’envisager d’autres sources complémentaires.

**Alimentation**

Chaque nouvelle année N, une fois par an (en année N+1), les identifiants de l’année N correspondant aux critères d’inclusion sont recherchés. La CNAM procède à l’extraction des consommations de soin de ces personnes :

* Consommations des années N-1 et N pour les nouveaux entrants dans la cohorte ;
* Consommations de l’année N pour les personnes déjà présentes dans la cohorte,

et fournit les données à l’Institut selon les modalités définies par la convention qui lie les 2 organismes.

**Les données des registres des cancers**

Les registres des cancers ont des missions nationales de veille et de recherche. Ils colligent des caractéristiques fines des cancers et couvrent la population adulte à hauteur de 20% et à 100% celle des enfants. Les registres sont évalués par le Comité d'Évaluation des Registres (CER) dans le cadre de la procédure de qualification qui porte sur leur qualité, leur intérêt pour la santé publique et la recherche et leur opportunité par rapport à la politique nationale développée dans leur domaine. 28 registres ont intégré un partenariat financier et scientifique dont fait partie l’Institut.

**Les données du dépistage**

Lorsque le système national des dépistages aura été constitué, les données nationales des dépistages organisés seront progressivement intégrées (cancers du sein, colorectal et du col de l’utérus). Les centres régionaux de coordination des dépistages des cancers (CRCDC) couvrent tout le territoire national. Ils colligent des informations relatives à la participation au dépistage, le niveau de risque des personnes, le diagnostic d’une tumeur détectée ainsi que le(s) premier(s) traitement(s). Le pilotage opérationnel et technique des programmes de dépistage, ainsi que l’évaluation organisationnelle, est assuré par l’Institut.

**Les données du dossier communicant de cancérologie**

Le dossier communicant de cancérologie vise à définir les principaux documents retraçant le parcours des personnes atteintes ou ayant été atteintes d’un cancer. Cela comprend les comptes-rendus d’anatomo-cytopathologie, de génétique moléculaire, la fiche de réunion de concertation pluridisciplinaire, le programme personnalisé de soins et le programme de l’après-cancer. Ces documents vont progressivement venir enrichir la plateforme de données en cancérologie et permettront d’apporter des informations essentielles sur la maladie comme son stade au diagnostic, les propositions thérapeutiques et les thérapeutiques décidées …

**Les résumés de passage aux urgences**

La base Oscour est une base de données développée par Santé publique France et disposant des informations relatives aux passages aux urgences. En accédant à ces informations au sein de la plateforme de données en cancérologie il devient possible de mieux investiguer les raisons des recours aux urgences et de mieux évaluer l’impact des thérapeutiques (avec notamment l’étude des effets secondaires ou indésirables).

**Les données des industries de santé**

L’industrie de la santé produit de nombreuses études, que cela soit lors du développement des thérapeutiques (avec notamment les études appelées essais de phase 1-2-3), au moment de la mise sur le marché des thérapeutiques (avec notamment les études relatives aux accès précoces) et dans le suivi de l’utilisation des thérapeutiques. En nouant un partenariat avec les industries de santé, l’Institut national vise à l’ouverture de ces données, en permettant leur dépôt et exploitation au sein de la plateforme de données en cancérologie. Ces données apportent des précisions notamment sur la maladie et sur les traitements proposés.

**Appariement**

Dans la cohorte cancer, les identifiants sont pseudonymisés avec des pseudonymes générés par la CNAM différents de ceux du SNDS. Seule la CNAM détient la correspondance entre les pseudonymes de la cohorte et du SNDS.

Les données des registres ainsi que celles du dépistage seront appariées aux données de la cohorte selon une méthode indirecte déterministe ou probabiliste. Dans les deux cas, l’Institut dispose d’autorisation CEREES - CNIL pour chainage de ces données à la cohorte cancer.

**Sécurité**

Le système d’information de la plateforme de données en cancérologie a été homologué (décision n°2023-27 du 3 avril 2023 de l’autorité d’homologation de l’Institut national du cancer portant homologation du système d’information intitulé « Plateforme des données en cancérologie »)

Toute personne accédant à la plateforme signe un document l’engageant, conformément notamment à l’article 32 du Règlement général sur la protection des données à prendre toutes précautions utiles dans le cadre de ses missions afin de protéger la confidentialité des informations auxquelles il a accès, et en particulier d’empêcher qu’elles ne soient modifiées, endommagées ou communiquées à des personnes non expressément autorisées à recevoir ces informations.

**Information individuelle des personnes**

Dans la mesure où l'Institut ne dispose pas de l'identité et de l'adresse des personnes concernées et compte tenu du nombre de personnes concernées (environ 7 millions), de l'incertitude sur leur statut vital et de l'ancienneté d'une partie des données, une information individuelle de l'ensemble des personnes concernées exigerait un effort disproportionné.

Dès lors, conformément aux autorisations CNIL référencées ci-dessous et conformément à l'article 14 paragraphe 5 point b) du RGPD (Règlement général sur la protection des données), l’Institut met à disposition sur son site internet une information collective relative à la Plateforme ainsi que des informations sur les projets de recherche ultérieurs menés à partir des données qu'elle contient et comportant l'ensemble des mentions prévues par l'article 14 du RGPD.

Une information collective est proposée sur le site internet de l’Institut : <https://lesdonnees.e-cancer.fr>

**Autorisations CNIL**

Délibération de la Commission nationale de l’informatique et des libertés (CNIL) n°2023-0404 en date du 20 avril 2023 autorisant l’Institut national du cancer à modifier un traitement de données à caractère personnel ayant pour finalité la constitution d’un entrepôt de données de santé, dénommé « Plateforme de données en cancérologie », nécessaire à la mise en œuvre de ses missions de service public,

Délibération de la Commission nationale de l’informatique et des libertés (CNIL) N°2019-083 en date du 20 juin 2019 autorisant l’Institut national du cancer à mettre en œuvre des traitements automatisés à des fins de recherche, d'étude et d'évaluation basés sur la « Plateforme de données en cancérologie ».

**Bibliographie (exemples de travaux)**

1-Bousquet PJ, Lefeuvre D, Tuppin P *et al*. Cancer care and public health policy evaluations in France: Usefulness of the national cancer cohort. *PLoS One*. 2018 Oct 31;13(10):e0206448.

2-Lefeuvre D, Le Bihan-Benjamin C, Pauporte I *et al*. French Medico-Administrative Data to Identify the Care Pathways of Women With Breast Cancer. Clin Breast Cancer. 2017 Jul;17(4):e191-e7

3-Lefeuvre D, Catajar N, Le Bihan-Benjamin C *et al*. Breast cancer screening: Impact on care pathways. Cancer Medicine. 2019;00:1–9.

4-Janah A, Gauthier LR, Morin L *et al*. Access to palliative care for cancer patients between diagnosis and death: a national cohort study. Clinical Epidemiology 2019 Jun 5;11:443-455

5-Janah A, Le Bihan-Benjamin C, Mancini J, *et al.* Access to inpatient palliative care among cancer patients in France: an analysis based on the national cancer cohort. BMC Health Serv Res. 2020 Aug 26;20(1):798.

6-Lailler G, Memoli V, Le Bihan Benjamin C, *et al.*. Five-Year Adjuvant Endocrine Therapy Adherence Trajectories Among Women With Breast Cancer: A Nationwide French Study Using Administrative Data. Clin Breast Cancer. 2021 *in press*

*7-Dumas E, Laot L, Coussy F et al.*[*The French Early Breast Cancer Cohort (FRESH): A Resource for Breast Cancer Research and Evaluations of Oncology Practices Based on the French National Healthcare System Database (SNDS).*](https://pubmed.ncbi.nlm.nih.gov/35681651/) *Cancers (Basel). 2022 May 27;14(11):2671. doi: 10.3390/cancers14112671.PMID: 35681651*

*8-Houzard S, Courtois E, Le Bihan Benjamin C et al* [*Monitoring Breast Cancer Care Quality at National and Local Level Using the French National Cancer Cohort.*](https://pubmed.ncbi.nlm.nih.gov/35750594/) *Clin Breast Cancer. 2022 Oct;22(7):e832-e841. doi: 10.1016/j.clbc.2022.05.006. Epub 2022 May 21.PMID: 35750594*

*9-Dumas E, Grandal Rejo B, Gougis P et al Concomitant medication, comorbidity and survival 1 in patients with breast cancer Nature Communications 2024 in press*

 **Pour plus d’informations**

Une information collective est proposée sur le site internet de l’Institut : <https://www.cancer.fr/>



|  |  |
| --- | --- |
|  |  |

#### Aspects réglementaires

*Rédigé par l’Institut*

*Si plusieurs sources de données, mettre les autorisations de chacun des sources.*

*La plateforme de données en cancérologie est un entrepôt de données autorisé par la Cnil : délibération de la Commission nationale de l’informatique et des libertés (CNIL) N\_2019-082 en date du 20 juin 2019 autorisant l’Institut national du cancer à constituer un entrepôt de données de santé, dénommé « Plateforme de données en cancérologie », nécessaire à la mise en œuvre de ses missions de service public.*

### Avis du Comité scientifique et éthique de la Plateforme de données en cancérologie, du CESREES et autorisation de la CNIL

*A adapter si passage ou non Cesrees / Cnil*

Préalablement à la conduite de l’étude, un avis / une autorisation sera obtenue du Comité scientifique et éthique de la Plateforme de données en cancérologie, du Comité Éthique et Scientifique pour les Recherches, les Études et les Évaluations dans le domaine de la santé (Cesrees) et de la Commission nationale de l’informatique et des libertés (Cnil).

La composition et l’avis du Comité scientifique et éthique de la plateforme de données en cancérologie sont rapportés en annexe.

### Accès aux données et confidentialité

Conformément aux dispositions relatives à la plateforme de données, l’accès aux données est règlementé et uniquement possible aux personnes individuellement identifiées et habilitées par l’Institut national du cancer.

### Audit

L’ensemble des actions conduites sur la plateforme de données est tracé. Des audits pourront être réalisés par les autorités compétentes afin de s’assurer que les exploitations effectuées sont en accord avec le présent protocole et du respect des règles d’utilisation des données de la plateforme. Des audits des extractions agrégées pourront également être effectués.

### Archivage des documents et Rapport final de l’étude

Les principaux documents relatifs à l’étude font l’objet d’un plan archivage. Ceci concerne les autorisations et évaluations, le rapport d’étude ainsi que l’ensemble des exportations (données agrégées et résultats d’analyses).

De mêmes, l’Institut conservera les documents et rapports nécessaires à la réalisation des actions de transparence telles qu’édictées par la règlementation. L’étude sera enregistrée sur le répertoire des études conduites à partir des données du SNDS mis en place par le Health data hub. Le rapport d’étude final sera remis au HDH dans les 2 ans suivant la fin de l’étude. Un rapport sera également rendu public, soit sous la forme d’une publication, soit sous la forme du rapport d’étude transmis au HDH, en cas de non publication des résultats dans les temps impartis.

### Délibération de la CNIL - Délibération 2023-040 du 20/04/2023

La Commission nationale de l'informatique et des libertés,

Saisie par l’Institut national du cancer d’une demande de modification de l’autorisation relative à un traitement automatisé de données à caractère personnel ayant pour finalité la constitution d'un entrepôt de données de santé visant à étudier les trajectoires des personnes atteintes d'un cancer ;

Vu le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l’égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données ou RGPD) ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l’informatique, aux fichiers et aux libertés, notamment ses articles 44-3° et 66-III ;

Vu la délibération n° 2019-082 du 20 juin 2019 autorisant l’Institut national du cancer (INCa) à mettre en œuvre un traitement automatisé de données à caractère personnel ayant pour finalité la constitution d’un entrepôt de données de santé visant à étudier les trajectoires des personnes atteintes d’un cancer, intitulé : "Plateforme de données en cancérologie" ;

Vu la délibération n° 2021-118 du 7 octobre 2021 portant adoption d’un référentiel relatif aux traitements de données à caractère personnel mis en œuvre à des fins de création d’entrepôts de données dans le domaine de la santé ;

Vu le dossier et ses compléments ;

Sur la proposition de Mme Valérie PEUGEOT, commissaire, et après avoir entendu les observations de M. Benjamin TOUZANNE, commissaire du Gouvernement.

**Formule les observations suivantes :**

|  |  |
| --- | --- |
| **Responsable du traitement** | **L’Institut national du cancer (INCa).** |
| **Modifications du traitement de données** | **Les modifications portent sur :*** **les finalités du traitement ;**
* **les catégories de personnes concernées par le traitement ;**
* **les bases de données existantes rassemblées au sein de la PDC (source des données) ;**
* **les catégories de données traitées ;**
* **les catégories de destinataires ;**
* **les modalités d’information des personnes ;**
* **les mesures de sécurité.**
 |
| **Points de non-conformité au référentiel concerné** | **La Commission prend acte de ce que le dossier de demande mentionne que le traitement envisagé est conforme aux dispositions du référentiel "entrepôt de données dans le domaine de la santé", à l’exception :*** **de certaines finalités poursuivies ;**
* **des sources des données collectées ;**
* **des catégories de données concernées ;**
* **des durées de conservation ;**
* **des modalités d’information des personnes ;**
* **des droits des personnes ;**
* **des mesures de sécurité.**

**En dehors de ces exceptions, le traitement devra respecter le cadre prévu par le référentiel "entrepôt de données dans le domaine de la santé".** |
| **Sur la finalité et la base légale du traitement** | **Les modifications envisagées ont vocation à étendre les finalités poursuivies par la Plateforme de cancérologie (PDC). Outre les finalités mentionnées dans la délibération n° 2019-082, le traitement permettra à l’INCa de :*** **définir et suivre le stratégie décennale de lutte contre les cancers ;**
* **réaliser des statistiques publiques en lien avec cette stratégie ;**
* **réaliser des expertises à la demande des ministères, de la Cour des comptes et du Parlement sur les questions relatives à la cancérologie ;**
* **mettre à disposition les données de la PDC à des tiers pour la réalisation de recherches, études ou évaluations dans le domaine de la cancérologie et pour répondre à la stratégie nationale des filières "industries et technologies de santé".**

**Les finalités du traitement sont déterminées, explicites et légitimes, conformément aux dispositions de l’article 5.1.b du RGPD.****La Commission prend acte de ce que l’INCa s’engage à :*** **ne pas traiter ou mettre à disposition les données de l’entrepôt à des fins de promotion des produits de santé en direction des professionnels de santé ou d’établissements de santé, ou à des fins d’exclusion de garanties des contrats d’assurance et de modification de cotisations ou primes d’assurance d'un individu ou d'un groupe d'individus présentant un même risque ;**
* **ne pas permettre la réutilisation des données de la PDC dans le cadre de recherche impliquant la personne humaine telles que définies aux articles L. 1121-1 à L. 1128-12 du code de la santé publique (CSP) ;**
* **publier les résultats des études conduites à partir de la PDC.**

**La Commission estime qu’il y a lieu de faire application des dispositions de l’article 44-3° et 66-III de la loi n° 78-17 du 6 janvier 1978 modifiée relative à l’informatique, aux fichiers et aux libertés (ci-après loi "informatique et libertés" modifiée), qui soumettent à son autorisation les traitements comportant des données relatives à la santé et justifiés, comme en l’espèce, par l’intérêt public.****Les utilisations futures des données contenues dans cet entrepôt s’inscriront dans le cadre des dispositions des articles 66 et 72 et suivants de la loi "informatique et libertés" modifiée, qui imposent que chaque traitement soit justifié par l’intérêt public et fasse l’objet de formalités propres.** |
| **Sur la base légale du traitement** | **Le traitement a pour base légale l’exécution d’une mission d’intérêt public dont est investi le responsable de traitement, au sens de l’article 6.1.e du RGPD.****Le traitement de données sensibles est nécessaire :*** **aux fins de recherche scientifique, s’agissant notamment de la mise à disposition des données de la PDC à des tiers pour la réalisation de recherches, études ou évaluations dans le domaine de la cancérologie (art. 9.2.j du RGPD) ;**
* **pour un motif d’intérêt public important pour les autres objectifs poursuivis par la PDC, sous réserve que les traitements ultérieurs soient réalisés sous la responsabilité de l’INCa (art. 9.2.i du RGPD).**
 |
| **Sur les données traitées** | **L’entrepôt regroupera les données issues :*** **des composantes de la base principale du système national des données de santé (SNDS), dans la limite de l’accès permanent dont dispose l’INCa en application des articles R. 1461-11 et suivants du CSP et sous réserve de leur disponibilité ;**
* **des registres des cancers ;**
* **d’études pour lesquelles l’INCa est responsable ou responsable conjoint de traitement ;**
* **des résumés de passage aux urgences (base Oscour) ;**
* **de recherches, études ou évaluations réalisées par les industriels de santé dans le cadre du projet de la filière intelligence artificielle et cancer ;**
* **des dossiers médicaux composant le dossier communicant de cancérologie ;**
* **des centres régionaux de coordination des dépistages des cancers.**

**La profondeur historique des données traitées sera de dix-neuf ans, en plus de l’année en cours.*****Les données relatives aux patients :*****Seules les données des personnes étant ou ayant été atteintes d’un cancer, des personnes rentrant dans un dispositif de prévention, de dépistage ou des personnes suspectées d’être atteintes d’un cancer seront rassemblées au sein de la PDC. Les personnes suspectées d’être atteintes d’un cancer sont des personnes présentant des risques élevés, soit en raison d’un test de dépistage positif soit parce qu’elles présentent une prédisposition génétique.****Les données suivantes seront traitées dans la PDC :*** **données d’identification : prénom, nom, coordonnées (téléphoniques, électroniques, postales), numéro d’identifiant permanent du patient (IPP), numéro d’identifiant de l’épisode de soin (IEP), identifiant national de santé (INS), date de naissance complète, sexe et genre ;**
* **données de santé : informations médicales, médicosociales et médico-économiques relatives au parcours de santé, informations médico-sociales relatives à la situation des personnes en situation de handicap, informations relatives aux arrêts de travail et aux prestations en espèce, informations relatives au statut vital et aux causes de décès ;**
* **données relatives à la vie personnelle : informations relatives au lieu de vie, aux habitudes de vie, à la vie professionnelle, à la défavorisation sociale ;**
* **données relatives aux organismes d’assurance maladie obligatoire et aux organismes d’assurance maladie complémentaire, limitées à celles contenues dans la base principale du SNDS.**

**Les données d’identification (prénom, nom, IEP, EPP, INS), les coordonnées (postales, électroniques, téléphoniques) et la date de naissance complète proviendront exclusivement des dossiers médicaux composant le dossier communicant en cancérologie. Les données d’identification et les coordonnées des personnes seront supprimées dès la pseudonymisation à bref délai effectuée dans la PDC.****La Commission prend acte de ce que l’INS fera l’objet d’une opération de pseudonymisation à bref délai dès le versement des données des dossiers communicant en cancérologie dans la PDC.*****Les données relatives aux professionnels de santé :**** **données d’identification : nom, prénoms, numéro ADELI, numéro du répertoire partagé des professionnels de santé (RPPS) ;**
* **données relatives à la vie professionnelle : fonction, lieu et mode d’exercice.**

**Les données dont le traitement est envisagé sont adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités du traitement, conformément aux dispositions de l’article 5.1.c du RGPD.** |
| **Sur les accédants et les destinataires** | **Seul le personnel désigné et habilité de l’INCa peut accéder aux données dans le cadre de la présente autorisation.****Sont destinataires des données :*** **le personnel désigné et habilité des sous-traitants intervenant pour le compte de l’INCa ;**
* **le personnel des partenaires de l’INCa qui conduisent un projet de recherche, d’étude ou d’évaluation dans le domaine la cancérologie, sous réserve qu’un contrat de responsabilité conjointe, au sens de l’article 26 du RGPD, ou de sous-traitance, au sens de l’article 28 du RGPD, ait été conclu avec l’INCa.**

**La Commission prend acte de ce que les catégories de destinataires et d’accédants sont soumis au secret professionnel régi par les dispositions des articles 226-13 et 226-14 du code pénal.** |
| **Information des personnes** | **La Commission prend acte des engagements de l’INCa selon lesquels :*** **une liste des bases de données existantes rassemblées au sein de la PDC sera publiée sur le site web de l’INCa. Cette liste devra être mise à jour régulièrement dès la conclusion ou la fin d’un contrat de mise à disposition de données afin de fournir une information transparente aux personnes concernant les bases de données rassemblées au sein de la PDC ;**
* **un renvoi vers les supports d’information des responsables de traitement des bases de données rassemblées au sein de la PDC (fournisseurs de données) sera réalisé depuis le site web de l’INCa ;**
* **une ligne téléphonique dédiée sera mise en œuvre par l’INCa pour informer les personnes concernant la PDC.**

**La Commission demande que :*** **l’information relative à la mise en œuvre de la PDC soit accessible sur les sites web des responsables de traitement des bases de données rassemblées au sein de la plateforme ;**
* **un délai d’un mois soit respecté entre la délivrance de l’information aux personnes concernées ou leurs représentants légaux ou la mise à jour de la liste des bases de données diffusées sur le site web de l’INCa et le versement effectif des données dans la PDC ;**
* **il lui soit communiqué tous les trois ans un rapport sur le fonctionnement de la PDC et sur les recherches réalisées à partir des données qu’il contient.**

**La Commission rappelle qu’il appartient à l’INCa, en sa qualité de responsable de traitement de la PDC, de transmettre aux responsables de traitement des bases de données rassemblées au sein de la plateforme une note d’information à destination des patients, de leurs représentants légaux et des professionnels de santé conforme aux dispositions du RGPD et de s’assurer que cette information leur sera remise. Les supports d’information devront comporter l’ensemble des mentions prévues par l’article 14 du RGPD.****La Commission rappelle également que, lorsque le responsable de traitement souhaite adopter une approche à plusieurs niveaux pour informer les personnes, la première modalité d’information devrait contenir, conformément aux lignes directrices sur la transparence au sens du RGPD du 29 novembre 2017 du groupe de travail "de l’Article 29" (G29), les informations relatives à la finalité du traitement, l’identité du responsable de traitement, une description des droits des personnes concernées, ainsi que toute information sur le traitement qui aurait une forte incidence sur la personne concernée. En l’espèce la Commission demande que l’ensemble des supports d’information indique que l’exercice des droits par les personnes, notamment leur droit d’opposition, n’aura pas de conséquence sur leur prise en charge médicale et mentionne un renvoi vers la page web dédiée à l’information relative à la PDC.****Il est rappelé que l’information relative à la constitution et aux modifications de la PDC ne se substitue pas à :*** **l’information incombant aux fournisseurs de données au titre des traitements de données à caractère personnel qu’ils mettent en œuvre en leur qualité de responsable de traitement, notamment s’agissant du traitement relatif à la tenue et la gestion des dossiers médicaux, des traitements relatif à la constitution d’une base de données pseudonymisées en vue de leur réutilisation à des fins de recherche, d’étude ou d’évaluation et des traitements mis en œuvre à des fins de recherche, d’étude ou d’évaluation ;**
* **l’information incombant à tout responsable de traitement qui procède à la réutilisation des données à caractère personnel de la PDC notamment à des fins de recherche, d’étude ou d’évaluation.**

**L’article L. 1461-3 du CSP subordonne l’accès aux données du SNDS et de ses composantes à la communication à la Plateforme des données de santé (PDS) de plusieurs éléments par les responsables de traitement, avant et après les études. Les projets de recherche, d’étude ou d’évaluation réalisés à partir des données de la PDC devront être enregistrés au sein du répertoire public de la PDS. La Commission recommande que la PDC soit également inscrite au répertoire public de la PDS.****S’agissant des personnes dont les données sont exclusivement issues du SNDS et des personnes déjà incluses dans la PDC dont le suivi médical n’est plus en cours :****En application de l'article 69 de la loi "informatique et libertés" et de l’article 14.5.b du RGPD, l'obligation d'information individuelle de la personne concernée peut faire l'objet d'exceptions, notamment dans l'hypothèse où la fourniture d'une telle information se révélerait impossible, exigerait des efforts disproportionnés ou compromettrait gravement la réalisation des objectifs du traitement. En pareils cas, le responsable de traitement prend des mesures appropriées pour protéger les droits et libertés ainsi que les intérêts légitimes de la personne concernée, y compris en rendant les informations publiquement disponibles.****En l'espèce, il est fait exception au principe d'information individuelle et des mesures appropriées seront mises en œuvre, notamment par :*** **la diffusion sur le site web de l’INCa d’une note d’information relative à la PDC ;**
* **la diffusion sur les pages des réseaux sociaux de l’INCa d’une note d’information relative à la PDC ;**
* **l’insertion dans les guides pratiques "guides patients", disponibles sur le site web de l’INCa, d’une note d’information dédiée à la PDC ;**
* **la diffusion via les lettres d’informations (newsletter) des associations de patients ou l’affichage au sein de ces associations d’une note d’information dédiée à la PDC.**

**S’agissant des personnes dont les données sont issues des centres régionaux de coordination des dépistages des cancers :****Ces personnes recevront un premier niveau d’information renvoyant au site web de la PDC, intégré au courrier d’invitation au dépistage organisé du cancer du sein, du cancer colorectal et du cancer du col de l’utérus.****S’agissant des personnes dont les données sont issues d’un registre des cancers :****Ces personnes recevront un premier niveau d’information renvoyant au site web de la PDC, intégré à la note d’information relative au registre du cancer concerné.****S’agissant des personnes dont les données sont issues du dossier communicant en cancérologie :****Ces personnes recevront un premier niveau d’information renvoyant au site web de la PDC, qui sera intégré :*** **au livret d’accueil mentionné à l’article L. 1112-2 du CSP ;**
* **aux documents remis à la personne lors de l’annonce d’un diagnostic ("mallette patient").**

**La Commission rappelle que chacun de ces premiers niveaux d’information devra comporter les informations mentionnées par le G29 dans ses lignes directrices sur la transparence.****Sous les réserves susmentionnées, ces modalités sont conformes au principe de transparence et aux exigences d’information prévues par les articles 12 et suivants du RGPD.** |
| **Droits des personnes** | **La Commission relève que, les finalités poursuivies par la PDC ne nécessitant pas le traitement de données directement identifiantes, les données de la PDC seront pseudonymisées par les responsables de traitements des bases de données rassemblées au sein de la PDC ou, s’agissant des données issues des dossiers médicaux composant le dossier communicant en cancérologie, dès leur versement.****Les personnes pourront exercer leurs droits (accès, rectification, opposition ou effacement) en contactant directement les responsables de traitement des organismes à l’origine de la collecte des données concernées. L’INCa s’est engagé à faciliter l’exercice de ces droits en renvoyant vers ces différentes structures. Les personnes pourront contacter l’INCa par voie électronique ou postale.****La Commission rappelle que si les personnes concernées fournissent des informations complémentaires permettant leur ré-identification, l’exercice de leurs droits devra être rendu possible directement auprès de l’INCa, conformément à l’article 11 du RGPD.** |
| **Mesures de sécurité** | **Le responsable de traitement a réalisé et transmis à l’appui de la demande d’autorisation une analyse d’impact relative à la protection des données spécifique à la PDC. Une homologation de cet entrepôt a également été réalisée par l’autorité d’homologation le 5 octobre 2022 conformément à l’arrêté du 22 mars 2017 relatif au référentiel de sécurité applicable au SNDS, pour une durée d’un an, sous réserve de la mise en œuvre du plan d’action qu’elle a défini. Cette décision d’homologation n’est valable que jusqu’au 5 octobre 2023 et devra donc être renouvelée avant cette date.****Les mesures de sécurité planifiées ou mises en place dans la PDC ont été comparées avec les exigences de sécurité mentionnées dans le référentiel "entrepôt de données dans le domaine de la santé" par le responsable de traitement. Des écarts avec les exigences de sécurité de ce référentiel ont été relevés concernant :*** **la pseudonymisation des données de santé ;**
* **les identifiants patient pseudonymisés qui pourront être fournis par la CNAM pour les données relatives au SNDS ou générés par la PDC.**

**Au regard de la volumétrie et de la sensibilité des données traitées, la pseudonymisation des données de santé devra s’effectuer au sein de la PDC dans un environnement d’intégration comprenant des mesures de sécurité techniques et organisationnelles renforcées, notamment afin de garantir un cloisonnement réseau, cryptographique et système spécifique à l’environnement d’intégration, par rapport au reste de l’entrepôt de données de santé. Cette pseudonymisation devra s’effectuer à bref délai et les données directement identifiantes et les coordonnées devront être supprimées dès la pseudonymisation des données et des documents réalisée.****La Commission rappelle que les différents identifiants pseudonymisés relatifs à chacune des différentes bases hébergées, l’identifiant pseudonymisé dérivé de l’INS, ainsi que l’identifiant principal, devront être générés conformément aux exigences du référentiel précité par une fonction de hachage cryptographique résistante aux attaques par force brute ou un générateur de nombres pseudo-aléatoires cryptographiquement sûr.****La table de liaison dont le but est de stocker les différents identifiants pseudonymisés des bases hébergées ainsi que l’identifiant principal devra être cloisonnée du reste de l’entrepôt avec un chiffrement dédié à cette seule table. Des profils d’accès spécifiques devront être mis en œuvre afin de garantir une traçabilité renforcée lors de l’accès à ces données.****Les données pseudonymisées seront chiffrées au repos par des algorithmes à l’état de l’art au sein d’un hébergeur certifié pour l’hébergement de données de santé, localisé en France et non soumis à des règlementations extra européennes.****Les mesures de sécurité, qui devront être opérationnelles lors de la mise en œuvre du traitement, devront répondre aux exigences prévues par les articles 5.1.f et 32 du RGPD, compte tenu des risques identifiés par le responsable de traitement. Il appartiendra au responsable de traitement de procéder à une réévaluation régulière des risques pour les personnes concernées et une mise à jour, le cas échéant, de ces mesures de sécurité.** |
| **Transferts de données** | **La présente autorisation ne vaut pas autorisation de transfert en dehors de l’Union européenne.** |
| **Sur la durée de conservation** | **Les données d’identification et les coordonnées des personnes provenant des dossiers médicaux composant le dossier communicant en cancérologie seront supprimées dès leur versement dans la PDC. L’INS des personnes sera supprimé dès l’opération de pseudonymisation à bref délai réalisée.****Les données à caractère personnel issues du SNDS sont conservées pour une durée conforme aux dispositions de l’article R. 1461-13. I. 1° du CSP.****Les données à caractère personnel autres que celles issues du SNDS seront conservées pendant vingt ans à compter de la dernière inclusion de données concernant la personne dans l’entrepôt.****La Commission considère que ces durées de conservation des données n’excèdent pas la durée nécessaire aux finalités pour lesquelles elles sont collectées et traitées, conformément aux dispositions de l’article 5.1.e du RGPD.** |
| **Observations complémentaires** | **Les autres conditions de mise en œuvre de la PDC restent inchangées.** |

**AUTORISE l’Institut national du cancer** à mettre en œuvre, pour une durée de dix ans, le traitement décrit ci-dessus.

La Présidente

Marie-Laure DENIS

### Délibération de la CNIL - Délibération 2019-083 du 20/06/2019 (à supprimer si l’étude n’entre pas dans le cadre de la décision unique )

Commission Nationale de l’Informatique et des Libertés

Nature de la délibération : Décision unique

Etat juridique : En vigueur

Date de publication sur Légifrance : Samedi 11 juillet 2020

**Délibération n° 2019-083 du 20 juin 2019 portant décision unique et autorisant l’Institut National du Cancer (INCa) à mettre en œuvre des traitements automatisés à des fins de recherche, d’étude et d’évaluation basés sur la « Plateforme de données en cancérologie » de l’INCa (Saisine n° 918314)**

La Commission nationale de l'informatique et des libertés,

Saisie par l’Institut National du Cancer d’une demande d’autorisation relative à un traitement automatisé de données à caractère personnel ayant pour finalité la réalisation de traitements à des fins de recherche, étude ou évaluation basés sur la Plateforme de données en cancérologie de l’INCa autorisée conformément à la délibération n°2019-082 du 20 juin 2019 ;

Vu la convention n° 108 du Conseil de l’Europe pour la protection des personnes à l’égard du traitement automatisé des données à caractère personnel ;

Vu le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l’égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l’informatique, aux fichiers et aux libertés, notamment ses articles 6 II, 44 3°, 66, 72 et suivants ;

Vu la loi n°2016-41 du 26 janvier 2016 de modernisation de notre système de santé, notamment son article 193 ;

Vu le Code de la santé publique, notamment son article L.1415-2 ;

Vu le décret n° 2019-536 du 29 mai 2019 pris pour l’application de la loi n° 78-17 du 6 janvier 1978 relative à l’informatique, aux fichiers et aux libertés ;

Vu l’avis favorable du comité d’expertise pour les recherches, les études et les évaluations dans le domaine de la santé (CEREES) du 18 juillet 2018 ;

Après avoir entendu Mme Valérie PEUGEOT, commissaire, en son rapport, et Mme Nacima BELKACEM, commissaire du Gouvernement, en ses observations,

**Formule les observations suivantes**

L’INCa dispose d’une autorisation pour la constitution d’un entrepôt, dénommé Plateforme de données en cancérologie (délibération n°2019-082), nécessaire à la mise en œuvre de ses missions de service public. Cet entrepôt se compose de données issues du Système national des données de santé (SNDS), de données issues des registres du cancer et de données issues des structures de gestion de dépistage du cancer.

Au titre de ses missions, l’INCa envisage de mettre en œuvre annuellement une centaine de traitements sur la base des données figurant dans cet entrepôt.

Pour ce faire, elle a déposé auprès de la Commission une demande d’autorisation fondée sur l’article 66 de la loi n° 78-17 du 6 janvier 1978 modifiée (ci-après loi informatique et libertés ).

Les traitements décrits relèvent du régime de l’autorisation de traitement à des fins de recherche, d’étude ou d’évaluation. La Commission a estimé opportun, au vu des éléments présentés dans le dossier de demande, et notamment de la volumétrie des traitements envisagés, d’en autoriser la mise en œuvre sur le fondement des dispositions des articles 66-IV, 72 et suivants de la loi informatique et libertés , dans le cadre d’une décision unique.

**Sur le responsable de traitement**

L’Institut National du Cancer (INCa), dont les missions sont précisées à l’article L. 1415-2 du code de la santé publique, est l’agence d'expertise sanitaire et scientifique en cancérologie de l’État chargée de coordonner les actions de lutte contre le cancer.

Ainsi, il a notamment pour mission l’observation et l’évaluation du dispositif de lutte contre le cancer, l’information des professionnels et le public sur l'ensemble des problèmes relatifs au cancer et la réalisation, à la demande des ministres intéressés, de toute expertise sur les questions relatives à la cancérologie et à la lutte contre le cancer.

**Sur l’application des dispositions liées au SNDS**

La Commission rappelle que l’ensemble des dispositions législatives et réglementaires relatives au SNDS est applicable en l’espèce et notamment l’interdiction d’utiliser ces données pour les finalités décrites à l’article L. 1461-1-V du code la santé publique.

**Sur la licéité des traitements et les conditions permettant de traiter des données concernant la santé**

Les traitements mis en œuvre par l’INCa s’inscrivent dans le cadre de ses missions de service public telles que précisées ci-dessus et sont nécessaires à l’exécution d’une mission d’intérêt public.

Ils sont, à ce titre, licites au regard des articles 6, paragraphe 1 point e) et remplissent une condition permettant le traitement des données concernant la santé au regard de 9 paragraphe 2 point j) du Règlement général sur la protection des données (ci-après RGPD).

**Sur la finalité des traitements et leur caractère d’intérêt public**

Les traitements nécessitant un accès aux données de la Plateforme de données en cancérologie ont pour finalité la réalisation de recherches, d’études et évaluations dans le cadre de l’exercice des missions de service public de l’INCa.

Pour ce faire, seront réalisées en particulier des études et évaluations :

* l’élaboration de tableaux de bord portant sur l’ensemble de la cancérologie (par exemple : dépenses relatives aux médicaments anticancéreux en milieu hospitalier ou en ville, activités de chimiothérapie en fonction du type de cancer ou de l’établissement de soins) ;
* des études relatives aux trajectoires de soins (suivi d’un groupe de personnes pour identifier la survenue d’un évènement ou observer les séquences de soins) ;
* des études demandées par des ministères.

La Commission considère que la finalité des traitements est déterminée, explicite et légitime, conformément à l’article 5 paragraphe 1 point b) du RGPD.

Par ailleurs, elle estime que ces traitements présentent une finalité d’intérêt public, conformément à l’article 66-I de la loi informatique et libertés .

**Sur les catégories de données traitées**

La Commission rappelle que le responsable de traitement ne doit traiter, dans le cadre de la présente décision unique, que les données strictement nécessaires et pertinentes au regard des objectifs des traitements et contenues dans laPlateforme de données en cancérologie.

Les catégories de données traitées susceptibles d’être traitées sont celles dont le traitement est autorisé dans le cadre de la délibération n° 2019-082, à savoir :

* les données issues du SNDS pour lesquelles l’INCa dispose, au titre des articles R. 1461-12 et R.1461-14 du code de la santé publique, d’un accès permanent ;
* les données issues des registres du cancer ;
* les données issues des structures de gestion de dépistages du cancer.

La Commission rappelle que, conformément à l'article 30 du RGPD, le responsable de traitement devra tenir à jour, au sein du registre des activités de traitement la liste des traitements mis en œuvre dans le cadre de la présente décision unique.

Le caractère adéquat, pertinent et limité à ce qui est nécessaire au regard des finalités pour lesquelles les données sont traitées, la zone géographique concernée et la profondeur historique des données consultées devront être justifiés pour chaque traitement mis en œuvre dans le cadre de la présente décision unique dans le registre des activités de traitement.

**Sur la durée de conservation des données**

La durée de la présente décision unique, durant laquelle des traitements à des fins d’étude, recherche ou évaluation réalisés à partir des données de l’entrepôt peuvent être mis en œuvre, est fixée à six ans.

La durée de conservation de chaque traitement devra être limitée à la durée nécessaire à la mise en œuvre de ce traitement et ne pourra excéder deux ans après la dernière publication des résultats ou en l’absence de publication, excéder la date de signature du rapport final de l’étude.

**Sur la publication des résultats**

La Commission rappelle que, lorsque le résultat du traitement de données est rendu public, l’identification directe ou indirecte des personnes concernées doit être impossible, conformément à l’article 68 de la loi informatique et libertés.

Les résultats des études réalisées dans le cadre de la présente décision unique pourront notamment être adressés au public, aux pouvoirs publics ou faire l’objet de publications scientifiques.

**Sur les catégories de destinataires des données**

Seul l’INCa a accès aux données dans le cadre de la présente décision unique. L’INCa tient à jour des documents indiquant la ou les personnes compétentes en son sein pour délivrer l'habilitation à accéder aux données, la liste des personnes habilitées à accéder à ces données, leurs profils d'accès respectifs et les modalités d'attribution, de gestion et de contrôle des habilitations. Seules les personnes habilitées par le responsable de traitement peuvent avoir accès aux données.

Ces catégories de personnes sont soumises au secret professionnel dans les conditions définies par les articles 226-13 et 226-14 du code pénal.

La qualification des personnes habilitées et leurs droits d'accès doivent être régulièrement réévalués, conformément aux modalités décrites dans la procédure d’habilitation établie par l’INCa.

**Sur l'information et les droits des personnes**

Les personnes concernées sont informées des traitements et peuvent exercer leurs droits conformément aux modalités décrites dans la délibération précitée.

En tout état de cause, l’ INCa s’engage à mettre à disposition sur son site internet une information collective relative à l’entrepôt ainsi que des informations sur les projets de recherche ultérieurs menés à partir des données qu’il contient et comportant l’ensemble des mentions prévues par l’article 14 du RGPD pour chaque projet de recherche.

**Sur la sécurité des données et la traçabilité des actions**

Seuls des traitements réalisés dans l’environnement homologué de l’INCa conformément à la délibération précitée peuvent être réalisés dans le cadre de la présente décision unique.

**Principe de transparence**

La mise à disposition des données du SNDS et de ses composantes est conçue de façon à rendre compte de leur utilisation à la société civile. À cette fin, l’article L. 1461-3 du code de la santé publique subordonne l’accès aux données du SNDS et de ses composantes à la communication à l’INDS de plusieurs éléments par le responsable de traitement, avant et après les études.

Ainsi, le responsable de traitement s’engage à enregistrer auprès du répertoire public tenu par l’INDS les études réalisées dans le cadre de cette décision unique. Cet enregistrement, à effectuer avant le début des traitements par le responsable de traitement ou la personne agissant pour son compte, s’accompagne de la transmission à l’INDS d’un dossier comportant :

* le protocole, incluant la justification de l’intérêt public, ainsi qu’un résumé, selon le modèle mis à disposition par l’INDS ;
* la déclaration d’intérêts du responsable du traitement, en rapport avec l'objet des traitements.

À la fin des études, la méthode et les résultats obtenus devront être communiqués à l’INDS en vue de leur publication. L’enregistrement des traitements et la transmission des résultats sont effectués conformément aux modalités définies par l’INDS.

Enfin, la Commission relève que l’INCa met à disposition sur son site internet une page dédiée aux données du cancer, dans laquelle figurent les résultats des évaluations réalisées dans le domaine de l’épidémiologie, la prévention et les facteurs de risque, le dépistage, la prise en charge, la recherche et l'innovation, et la vie avec le cancer (http://lesdonnees.e-cancer.fr).

Dans ces conditions, la Commission autorise, conformément à la présente délibération, l’INCa à mettre en œuvre les traitements décrits ci-dessus pendant une durée de six ans, à compter de la disponibilité effective des données au sein de la Plateforme de données en cancérologie, avec obligation de remise d’un bilan à la Commission à l’issue de ce délai.

La Commission rappelle que tout traitement ne remplissant pas les conditions prévues par la présente décision unique devra faire l’objet de formalités spécifiques auprès d’elle.

La décision DE-2011-058 autorisant l’INCa à mettre en œuvre un traitement de données de santé à caractère personnel ayant pour finalité de constituer des tableaux de bord et d’étudier la trajectoire des patients atteints d'un cancer à partir des données du SNIIRAM (demande d’autorisation n° 911297) est abrogée.

La Présidente

Marie-Laure DENIS

### Avis du Comité scientifique et éthique de la plateforme de données en cancérologie

**Comité Scientifique et Éthique de la Plateforme de Données en Cancérologie**

**Composition**

* **Djadi-Prat Juliette**, *présidente du Comité*, médecin de santé publique, responsable adjointe de l’unité de recherche clinique, Hôpital européen Georges Pompidou Paris
* **Poli Marie-Joséphine**, *vice-présidente du Comité*, représentant des usagers, présidente régionale (Corse) de France-assos-santé
* **Bossi-Malafosse Jeanne**, juriste, responsable du département « Données personnelles » et co-responsable du département « Sciences du vivant », Cabinet Delsol
* **Delpuech Claude**, docteur en science physique, membre du Comité consultatif national d’éthique (CCNE), responsable du Département magnétoencéphalographie du Centre d'Etude et de Recherche Multimodal Et Pluridisciplinaire en imagerie du vivant (CERMEP – imagerie du vivant)
* **Empereur Fabienne**, médecin de santé publique, médecin coordonnateur du réseau ONCOPL (réseau régional de cancérologie des pays de la Loire)
* **Franck Denis**, onco-radiothérapeute, président de l’UNHPC (union nationale hospitalière de cancérologie), Clinique Pasteur Toulouse
* **Haguenoer Ken**, médecin de santé publique, directeur médical du Centre régional de coordination des dépistages des cancers Centre-Val de Loire
* **Linda Pierre**, otorhinolaryngologue, Hôpital Schuman Metz
* **Martin Anne-Laure**, pharmacienne, directrice Stratégie data Unicancer
* **Ménegaux Florence**, médecin de santé publique, Centre de recherche en Epidémiologie et Santé des Populations (CESP) – Inserm U1018
* **Molinié Florence**, médecin de santé publique, directrice du registre des cancers de Loire-Atlantique/Vendée
* **Quesnot Stéphane,** médecin directeur de l’information médicale, Direction de l’offre de soins - MGEN
* **Rachas Antoine**, médecin de santé publique, Département des études sur les pathologies et les patients (DEPP) CNAM
* **Rouleau** **Etienne**, biologiste spécialisé en oncogénétique, chef du Service génétique des tumeurs – Institut Gustave Roussy
* **Rousseau Agnès**, secrétaire générale comité du Val d'Oise de la Ligue contre les cancers
* **Troussard Xavier**, hématologue, président de la FHF Cancer, directeur du Registre des hémopathies malignes de Basse Normandie, Centre hospitalo-universitaire de Caen
* **Villata Séréna**, docteur en science informatique, directrice scientifique adjointe Laboratoire d'informatique, signaux et systèmes (I3S) Sophia Antipolis

**Avis**

Quatre types d’avis sont émis par le Comité scientifique et éthique :

**☐** Avis favorable *Le projet est validé par le Comité tel que proposé*

**☐** Avis favorable sous conditions : les réponses aux remarques permettront de délivrer un avis favorable *Le projet est validé par le Comité, quelques points peuvent être éclaircis. Les réponses à ces points ne nécessitent pas un nouveau passage du projet en séance*

**☐** Avis réservé : le dossier sera revu en commission après réponses aux réserves *Le projet doit être étayé et repasser en séance*

**☐** Avis défavorable *Le projet ne peut être conduit dans sa forme actuelle (présence d’éléments bloquants). Si les porteurs veulent poursuivre le projet, ils doivent recommencer l’ensemble du processus de soumission.*

L’évaluation du projet par le Comité scientifique et éthique porte notamment sur :

* Le plan éthique
* L’information des personnes
* Le plan scientifique et méthodologique

Des remarques peuvent compléter l’avis et sont apportées à titre indicatif par le Comité.

1. Le promoteur est le responsable de traitement : à adapter [↑](#footnote-ref-1)
2. Chargé de la mise en œuvre des traitements : à adapter [↑](#footnote-ref-2)