

# **Guide sur les études de faisabilité pour des projets de recherche** dans le cadre d'utilisations secondaires de données de santé

---

# Sommaire

REMERCIEMENTS	3
INTRODUCTION	4
À qui s'adresse ce guide ?	4
Quels sont les objectifs de ce guide ?	5
Comment est-il organisé ?	5
<b>1. DÉFINITION DE L'ÉTUDE DE FAISABILITÉ</b>	<b>6</b>
<b>2. CONTEXTE RÉGLEMENTAIRE ET ÉLÉMENTS CONTRACTUELS</b>	<b>7</b>
<b>2.1.</b> Le rôle des acteurs et leur qualification au regard du RGPD	7
<b>2.2.</b> Les obligations des acteurs	11
<b>2.3.</b> Éléments de contractualisation entre le détenteur de données et l'industriel	12
<b>3. DÉROULÉ D'UNE ÉTUDE DE FAISABILITÉ</b>	<b>14</b>
<b>Étape 1.</b> Demande de réalisation d'une étude de faisabilité	14
<b>Étape 2.</b> Faisabilité du projet de recherche	16
<b>Étape 3.</b> Restitution des résultats de l'étude de faisabilité	17
ANNEXES	21

# Remerciements

Ce guide a été élaboré par le Leem dans le cadre des travaux du Groupe de Travail « Faisabilité ».

## **Équipe Leem :**

- Nathalie Manaud (Leem) ;
- Mélody Monthéard (Leem) ;
- Marine Tabary (Leem).

## **Membres du Groupe de Travail « Faisabilité » :**

- Stéphanie Duret (Novartis) ;
- Mélina Gilberg (Amgen France) ;
- Yoann Gorrichon (Abbvie) ;
- Yoann Lelarge (Johnson & Johnson) ;
- Laura Luciani (Merck) ;
- Valérie Machuron (Roche) ;
- Gaëlle Nachbaur (GSK) ;
- Pauline Prince (Novartis) ;
- Pierre-Alexandre Squara (Pfizer).

## **Membres du Comité Données personnelles de santé :**

- Sarah Amar (Sanofi)
- Selima Ellouze (DPO Incyte)
- Antoine Perpoil (Amgen France)

Cette démarche s'est enrichie grâce à la consultation et à la relecture attentive de **détenteurs de données, ainsi que d'autres acteurs de l'écosystème des données de santé** intervenant dans le cadre des études de faisabilité, qui ont accepté de partager leur expertise et/ou leurs retours d'expérience :

- The Lymphoma Academic Research Organisation (LYSARC) ;
- Association RESPIC pour la Recherche dans les établissements privés d'intérêt collectif ;
- Banque Nationale de Données Maladies Rares (BNDMR) ;
- France Cohortes ;
- Lifen ;
- Clinique des données et Département Innovation CHU de Nantes ;
- Assistance Publique – Hôpitaux de Paris (AP-HP) ;
- Paris Saclay Cancer Cluster (PSCC) ;
- Unicancer ;
- Ospi ;
- Health Data Hub (HDH).

**Nous remercions l'ensemble des contributeurs pour leur expertise et leurs partages d'expériences qui ont permis d'élaborer le présent guide.**

# Introduction

**Les études de faisabilité constituent une étape importante pour développer des projets de recherche en santé. Elles permettent d'évaluer la capacité d'une base de données à répondre aux objectifs de tout projet de recherche.**

Un état des lieux réalisé par le Leem en 2024, visant à évaluer l'attractivité de la France pour l'usage secondaire des données de santé, a été réalisé au sein des entreprises du médicament et de leurs filiales en Allemagne, en Espagne, au Canada et dans les Pays scandinaves. Celui-ci a notamment révélé que la France apparaissait comme le pays le plus défavorable pour les études de faisabilité préalables à l'initiation d'un projet, à l'inverse de l'Espagne (*cf. annexe 5*). Aussi, pour tirer le meilleur parti des atouts dont la France dispose par ailleurs, et afin de renforcer son attractivité en stimulant et encourageant la réalisation de projets, il a été proposé de travailler à la mise en place d'un cadre méthodologique permettant de faciliter la réalisation d'études de faisabilité.

Les bénéfices clés d'une étude de faisabilité sont nombreux :

- 1. Optimisation de la collaboration entre industriels et détenteurs de données**, grâce à des échanges structurés et une compréhension mutuelle des contraintes et des objectifs ;
- 2. Réduction du risque de non-aboutissement d'un projet de recherche en santé**, en s'assurant que la question de recherche est réellement adressable par les données disponibles ;
- 3. Respect des échéanciers et du cadre réglementaire applicable**, grâce à une anticipation des obstacles et une planification réaliste ;
- 4. Prise de décision éclairée** avant tout engagement contractuel, en réduisant les risques liés à des choix insuffisamment documentés ;

- 5. L'étude de faisabilité est un investissement** faible comparé au coût potentiel de l'échec d'un projet de recherche sur données.

Dans un contexte où la valorisation de l'utilisation secondaire des données de santé devient un enjeu majeur pour l'innovation thérapeutique, il est essentiel de disposer d'un cadre méthodologique partagé facilitant les collaborations entre industriels et détenteurs de données.

## À QUI S'ADRESSE CE GUIDE ?

Ce guide méthodologique sur les pratiques de faisabilité pour les études impliquant des données de santé s'adresse aux détenteurs de données (producteurs ou gestionnaires de bases de données) et aux industriels souhaitant mettre en place **des études pour des projets de recherche utilisant des données de santé existantes et nécessitant d'en apprécier la faisabilité**.

Le terme « **industriel** » désigne les entreprises du secteur de la santé, notamment pharmaceutiques, qui souhaitent porter un projet de recherche s'appuyant sur des données de santé existantes pour en évaluer la faisabilité avant mise en œuvre.

Le terme « **détenteur de données** » désigne les établissements ou structures (hôpitaux, centres de recherche, autres organismes publics ou privés, etc.) qui collectent, structurent et gèrent des données de santé. Ces détenteurs de données sont en mesure de mobiliser leurs bases pour répondre à une demande d'étude de faisabilité.

La pratique de faisabilité est de plus en plus répandue au sein des détenteurs de données. Une enquête menée par le Comité Stratégique de Filière Industries et Technologies de Santé (CSF) auprès de 21 détenteurs de données — incluant 17 registres/cohortes — révèle que **86 % des répondants réalisent des études de faisabilité**<sup>(1)</sup>. Cependant, les process et méthodes appliqués sont hétérogènes.

Ce guide a pour objectif d'accompagner les acteurs qui réalisent des études de faisabilité à partir des différentes sources de données de santé secondaires, notamment :

- **Les registres et cohortes (et entrepôts de données de santé thématiques)** collectant des données de santé pour permettre dans un second temps de répondre à des objectifs scientifiques précis (usage secondaire), avec des protocoles de collecte et des variables prédéfinies (questionnaires, *Case Report Forms* (CRF), etc.).
- **Les entrepôts de données de santé hospitaliers (EDSH)** centralisant notamment des données issues des systèmes d'information hospitaliers couvrant de multiples pathologies et situations cliniques. Ces données n'ont pas été initialement collectées à des fins de recherche mais pour des besoins de soins et de gestion.

*À noter que la base principale du SNDS est exclue du périmètre couvert dans ce guide.*

Bien que des initiatives soient en cours de développement, notamment dans le cadre des objectifs de l'Espace Européen des Données de Santé, il n'existe à ce jour aucune ressource officielle réunissant (ou donnant accès) à l'ensemble des bases de données de santé disponibles en France, avec leurs caractéristiques détaillées.

Ce manque de visibilité peut rendre le travail de prospection long, souvent basé sur des travaux de veille continue et sur les réseaux professionnels des équipes de recherche. Il convient de noter que cette première étape de prospection n'est pas systématique : dans certains cas, des bases de données sont déjà pré-identifiées, recommandées par la HAS.

Des initiatives émergent progressivement pour structurer cette cartographie<sup>(2)</sup>, avec l'ambition de créer à terme un référencement national des bases de données exploitables dans le cadre de l'évaluation de l'innovation en santé.

## QUELS SONT LES OBJECTIFS DE CE GUIDE ?

Le guide vise à :

- faciliter la compréhension mutuelle entre industriels et détenteurs de données ;
- apporter une aide méthodologique et des points de repère pratiques pour la réalisation des études de faisabilité ;
- accompagner la réalisation des études de faisabilité, en proposant un parcours structuré et une maîtrise des délais.

*Les recommandations décrites dans le présent guide n'engagent pas la responsabilité du Leem, de ses adhérents ni des détenteurs de données impliqués dans la démarche d'élaboration du guide.*

## COMMENT EST-IL ORGANISÉ ?

Ce guide est organisé en 3 grandes parties :

### 1. Définition de l'étude de faisabilité

Il s'agit dans cette première partie de fournir une définition claire et partagée des études de faisabilité pour des projets de recherche utilisant des données de santé existantes.

### 2. Contexte réglementaire et éléments contractuels

Cette partie décrit le cadre réglementaire dans lequel s'inscrivent les études de faisabilité pour des projets de recherche utilisant des données de santé existantes, et traite des modalités contractuelles entre industriels et détenteurs de données. Ces éléments constituent des repères et non des prescriptions opposables. Il appartient aux parties prenantes de les adapter à leur contexte spécifique et aux besoins de chaque étude.

### 3. Déroulé d'une étude de faisabilité

Cette partie a pour objectif de permettre une compréhension des besoins et attentes des entreprises du médicament. Il s'agit de mettre à disposition l'ensemble des informations nécessaires à la préparation des premiers échanges entre industriels et détenteurs de données, afin de fluidifier la réalisation de ces études de faisabilité. Cette partie présente également le résultat attendu d'une étude de faisabilité et propose un contenu structuré.

(1) CSF, poster ISPOR—The International Society for Pharmacoeconomics and Outcomes Research "From Data to Decisions: Enhancing Real-World Data Quality: A Survey of Health Data Producers in France"

(2) Cf. le [Portail Épidémiologique France](#), le [recensement CNIL des EDS](#), le [catalogue du Health Data Hub \(HDH\)](#), le [recensement de la Haute Autorité de Santé \(HAS\)](#).

# 1 Définition de l'étude de faisabilité

Cette première partie du guide a pour objectif de clarifier ce qu'est une étude de faisabilité pour des projets de recherche utilisant des données de santé existantes. La définition proposée n'est pas prescriptive. Bien que le guide vise à harmoniser les démarches, chaque structure peut conserver ses propres pratiques. Il ne s'agit donc pas d'imposer une méthode unique, mais de proposer un socle commun entre industriels et détenteurs de données.

## Objectifs principaux de la démarche de l'étude de faisabilité

L'étude de faisabilité repose sur une analyse approfondie de la base de données de santé pré-sélectionnée dont les objectifs principaux sont notamment :

- la vérification de la capacité à extraire ou structurer les variables d'intérêt **indispensables** à la mise en œuvre d'une question de recherche pré-identifiée ;
- l'évaluation de la **représentativité des données de santé**, afin de garantir la validité des résultats et leur transposabilité aux populations cibles ;
- l'analyse de la **qualité des données de santé** (exhaustivité, exactitude, cohérence, disponibilité des variables), pour s'assurer de leur exploitabilité dans le cadre du projet de recherche<sup>(3)</sup>.

## Modalités et implications pratiques

L'étude de faisabilité est une étape exploratoire qui repose sur une **analyse structurée, approfondie et documentée**, visant à évaluer l'adéquation entre les caractéristiques de la base de données de santé et les exigences du projet de recherche. Cette démarche peut inclure l'examen de la qualité des variables, de leur disponibilité, de la représentativité de la population cible, de la temporalité des données, etc.

### **QU'EST-CE QU'UNE ÉTUDE DE FAISABILITÉ ?**

L'étude de faisabilité pour des projets de recherche utilisant des données de santé existantes est une **démarche structurée menée en amont, visant à vérifier dans quelle mesure une base de données existante est adaptée et suffisante pour répondre aux objectifs du projet, dans les conditions scientifiques, réglementaires et techniques envisagées**. Toutefois, elle ne doit pas se substituer au projet de recherche en répondant à la question scientifique pour laquelle la faisabilité a été menée.

L'étude de faisabilité **peut également impliquer un coût financier et requérir une contractualisation**, notamment lorsque des ressources humaines sont mobilisées pour la complétion des données, des statistiques, etc. Elle implique souvent un engagement formel entre les parties, avec des livrables attendus et des délais à respecter.

## Portée et limites de l'étude de faisabilité

L'étude de faisabilité permet *in fine* **d'identifier les éventuelles inadéquations entre les objectifs du projet de recherche en santé et les caractéristiques des bases de données sources** (contenu, structure, qualité). Elle aide à prendre une décision éclairée avant tout engagement contractuel pour la réalisation d'un projet de recherche, réduisant ainsi le risque de non-aboutissement ou de collaborations non fructueuses.

<sup>(3)</sup> Dans le contexte des EDSH, pour les données non structurées, il peut être difficile de communiquer des métriques de qualité robustes au stade la faisabilité.

# 2 Contexte réglementaire et éléments contractuels

Parce qu'elle nécessite l'analyse des données contenues dans la base initiale, l'étude de faisabilité implique nécessairement un traitement de données personnelles. Une première question à se poser est donc : **qui en est responsable ?** Est-ce le détenteur de la base de données, l'industriel du médicament porteur d'un projet de recherche en santé, ou les deux conjointement ? En parallèle de l'identification du responsable de traitement, il convient d'identifier quel(s) acteur(s) interviennent en qualité de sous-traitant(s) de données personnelles.

Cette qualification est déterminante, car elle permet d'identifier quelles obligations au titre du RGPD et au titre de la loi Informatique et Libertés s'appliquent à chacun des acteurs.

Cette partie du guide présente les principaux éléments à prendre en compte pour sécuriser la réalisation des études de faisabilité :

- **le rôle des acteurs et leur qualification au regard du RGPD** (responsable (conjoint) de traitement ou sous-traitant) ;
- **les éventuelles formalités réglementaires à accomplir.**

Ces aspects réglementaires peuvent, le cas échéant, être complétés par un encadrement contractuel, afin de formaliser un accord entre les parties définissant clairement les responsabilités et les conditions de réalisation de l'étude de faisabilité.

## 2.1. Le rôle des acteurs et leur qualification au regard du RGPD

La première étape consiste à identifier l'entité ou les entités agissant en qualité de responsable(s) de traitement dans le cadre de l'étude de faisabilité. Cette qualification est essentielle, car c'est au responsable de traitement qu'incombent une grande partie des exigences prévues par le RGPD, ainsi que l'accomplissement des formalités auprès de la CNIL au titre de la loi Informatique et Libertés<sup>(4)</sup>.

Conformément au RGPD, le **responsable du traitement** est la personne physique ou morale [...] qui détermine, seule ou conjointement, les finalités et les moyens du traitement. Le sous-traitant est la personne physique ou morale [...] qui traite les données personnelles pour le compte du responsable de traitement. Ces définitions s'appliquent pleinement aux études de faisabilité.

**L'analyse qui suit propose des repères pour aider les acteurs impliqués dans une étude de faisabilité**

(détenteurs de données, industriels ou tout autre acteur intervenant dans la réalisation d'une étude de faisabilité) à **identifier leurs rôles respectifs au regard du RGPD** (responsable du traitement, responsable conjoint ou sous-traitant). Il est rappelé que la qualification des parties au regard du RGPD s'analyse nécessairement au cas par cas, selon les interactions entre ces acteurs et leur degré d'implication dans le traitement de données (quelle entité est à l'initiative de l'étude de faisabilité et en définit les caractéristiques essentielles ? quelle entité en tire un bénéfice en premier lieu ? etc.).

---

(4) Ainsi que le cas échéant, au titre du code de la santé publique s'agissant des obligations de l'industriel responsable d'un traitement impliquant des données issues du SNDS élargi (dont ont par exemple vocation à faire partie les entrepôts hospitaliers). Dans ce cadre, les articles L. 1461-1 et suivants du CSP s'appliquent aux responsables de ces traitements et leur imposent des exigences en matière de finalités interdites, de transparence et de sécurité via le référentiel de sécurité du SNDS.

### CAS N°1 - Le plus courant

## Le détenteur de données est seul responsable du traitement

Il arrive fréquemment que les détenteurs de données établissent, de manière indépendante et autonome, des indicateurs relatifs à leur base de données. Les traitements de données correspondants sont effectués par le détenteur, sans intervention ni influence de l'industriel. Dès lors qu'un traitement de données est bien réalisé dans le cadre de la phase préliminaire à un projet de recherche et que le jeu de données constitué peut présenter un intérêt pour un nombre indéfini d'entités, dans le cadre d'études de faisabilité, le détenteur de données est réputé avoir un intérêt direct dans la réalisation de ce traitement.

**Dans la mesure où le détenteur définit seul les finalités et les caractéristiques essentielles du traitement, il agit en qualité de seul responsable de traitement.**

En pratique, les indicateurs établis par les détenteurs de données sous leur propre responsabilité peuvent être mis à disposition des industriels (soit sur demande, soit publiquement), sous forme de catalogues de métadonnées n'incluant aucune donnée personnelle. L'industriel qui consulte ces données – aux fins d'évaluer la faisabilité de sa recherche – ne traite donc aucune donnée personnelle et ne joue aucun rôle au titre du RGPD (ni responsable de traitement, ni sous-traitant).

### EXEMPLES

■ *Le détenteur d'une base de données établit des indicateurs standards, tels que le nombre de patients inclus dans la base, le taux de complétude des données, la répartition hommes/femmes, etc. Ces données sont mises à disposition des porteurs de projets, sous la forme d'un catalogue de métadonnées n'incluant aucune donnée personnelle.*

■ *Le détenteur d'une base de données met en place des « requêtes types », que les porteurs de projets peuvent solliciter de manière ad hoc. Les traitements de données correspondants sont standards et prédéfinis par le détenteur de données, le porteur de projet ne peut pas les personnaliser. Le résultat de ces requêtes est mis à disposition de l'industriel à sa demande, sous forme de données agrégées n'incluant aucune donnée personnelle ou ré-identifiante.*

Lorsque le détenteur de données agit seul en qualité de responsable du traitement, il est de fait seul responsable du respect des dispositions du RGPD (notamment le respect du principe de licéité, minimisation, l'obligation d'informer les personnes concernées, etc.) et de la loi Informatique et Libertés, i.e., l'accomplissement des formalités préalables requises. Sur ce dernier point, **la conformité du traitement réalisé par le détenteur de données implique que les traitements de données aux fins d'évaluation de la faisabilité (indicateurs prédéfinis, requêtes simples, etc.) aient été anticipés dans le cadre de la formalité initiale effectuée par le détenteur pour la mise en œuvre de sa base de données.**

Un contrat ou clause relative à la protection des données pourra s'avérer utile et est recommandé(e), même si un tel contrat ou clause n'est pas obligatoire (cf. § 2.3 *Éléments de contractualisation*).

**Ce cas constitue la configuration la plus efficiente pour la réalisation d'une étude de faisabilité dans des délais optimisés, tout en assurant le respect des exigences réglementaires.**



## CAS N°2 - Frontière entre faisabilité et projet de recherche

### L'industriel est seul responsable du traitement

Les indicateurs ou requêtes simples mises en place par les détenteurs de données peuvent s'avérer insuffisant(e)s pour permettre aux porteurs de projets d'évaluer la faisabilité de leur projet de recherche. Ces derniers peuvent alors solliciter une analyse « personnalisée » des données disponibles dans la base, en définissant les variables à prendre en compte, les critères à évaluer ou la méthodologie générale de l'étude de faisabilité. Le détenteur de données procèdera alors à un traitement de données à la demande de l'industriel, pour répondre aux besoins spécifiques de ce dernier. Dans ce cas, l'industriel agira en qualité de responsable de traitement et le détenteur de données agira en qualité de sous-traitant.

#### Il est important de rappeler les éléments suivants :

1. L'accès aux données personnelles n'est pas un critère de qualification du responsable de traitement. Le fait que l'industriel n'accède pas aux données directement ou indirectement identifiantes contenues dans la base ne remet pas en question sa qualification en tant que responsable de traitement ;
2. La connaissance exhaustive des variables disponibles dans une base n'est pas nécessaire pour qualifier une entité de responsable de traitement. Ainsi, l'absence de maîtrise complète de la base n'empêche pas un industriel d'être qualifié de responsable de traitement ;
3. Le fait que le détenteur de données agisse en qualité de sous-traitant dans le cadre de l'étude de faisabilité ne remet pas en question son rôle de responsable de traitement de sa base de données, ni le contrôle qu'il exerce sur l'accès aux données et leur traitement. Ainsi, un détenteur de données :
  - peut conseiller le porteur de projet dans la conduite de son étude de faisabilité ;
  - peut refuser de traiter les données sur instruction de l'industriel, s'il pense que ces instructions sont contraires aux dispositions du RGPD ou toute autre réglementation applicable ;
  - décide s'il communique à l'industriel les résultats de l'étude de faisabilité sous forme de données personnelles pseudonymisées ou sous forme de données non personnelles<sup>(5)</sup>.

#### EXEMPLE

■ *Les indicateurs standards et les « requêtes types » mis(es) en place par un détenteur de données ne permettent pas à un industriel d'évaluer la faisabilité de son projet de recherche. L'industriel demande d'effectuer une requête ad hoc personnalisée, plus complexe, auprès du détenteur de données. Le détenteur de données n'avait pas anticipé ce type de requête au préalable et accepte de la réaliser à la demande et pour le compte de l'industriel.*

Dans ce contexte, les deux acteurs interviennent dans le traitement des données avec des rôles distincts (responsable de traitement ou sous-traitant), et chacun demeure tenu de respecter le RGPD. Seul l'industriel devra en revanche réaliser les formalités prévues par la loi Informatique et Libertés. Un contrat adapté devra être mis en place liant le sous-traitant à l'égard du responsable du traitement (cf. § 2.3 *Éléments de contractualisation*). Un contrat de prestation pouvant inclure une clause relative à la protection des données pourra s'avérer utile et est recommandé(e), même si un tel contrat ou clause n'est pas obligatoire (cf. § 2.3 *Éléments de contractualisation*).

(5) Voir infra  
« 2.3 Éléments de contractualisation entre le détenteur de données et l'industrie »

### CAS N°3 - Collaboration étroite entre détenteur et industriel

#### **Le détenteur de données et l'industriel sont responsables conjoints du traitement**

Le détenteur de données et l'industriel peuvent collaborer étroitement pour la conduite d'une étude de faisabilité et définir, conjointement, les finalités et les moyens du traitement de données associé. Dans ce cas, le détenteur et l'industriel agiront en qualité de responsables conjoints de traitement.

##### **EXEMPLES**

■ *Un industriel sollicite une requête ad hoc auprès d'un détenteur de données. Ce dernier n'avait pas anticipé ce type de requête et identifie un intérêt propre à mener cette analyse pour son propre compte. Les parties construisent alors une requête de manière conjointe, qui répond à leurs besoins respectifs.*

■ *Le détenteur d'une base de données et un industriel souhaitent mener conjointement une recherche en santé (en tant que responsables conjoints) et souhaitent en évaluer la faisabilité au préalable. L'étude de faisabilité résulte d'une initiative commune des deux parties, répond à leurs intérêts communs et est construite conjointement par les deux.*

Le détenteur et l'industriel qui agissent en qualité de responsables conjoints du traitement seront conjointement responsables de la conformité de l'étude de faisabilité aux dispositions du RGPD<sup>(6)</sup> et de la loi Informatique et Libertés. Un contrat spécifique devra être établi, définissant de manière transparente leurs obligations respectives aux fins d'assurer le respect des exigences RGPD (cf. § 2.3 *Éléments de contractualisation*).

### CAS N°4

#### **Autres cas**

D'autres cas peuvent impliquer la participation d'un acteur qui n'est ni le détenteur de données, ni l'industriel, mais qui intervient tout de même dans le cadre d'une étude de faisabilité.

C'est par exemple le cas d'une CRO qui agit pour le compte d'un industriel. Cela est également le cas d'un prestataire mettant à disposition un outil d'intelligence artificielle permettant la structuration des données et facilitant l'évaluation de la faisabilité — par exemple, une solution d'IA dédiée à l'analyse de données permettant d'identifier les patients, de pseudonymiser les informations personnelles, ou encore d'extraire et de traiter les données).

Ces situations nécessitent une analyse au cas par cas : cet acteur peut être qualifié de sous-traitant du détenteur et/ou de l'industriel, ou, dans certains cas, de responsable conjoint du traitement avec l'un d'eux ou avec les deux. En revanche, *a priori*, cet acteur ne devrait pas pouvoir être considéré comme seul responsable du traitement qu'implique l'étude de faisabilité.

(6) Article 26 RGPD  
« Lorsque deux responsables du traitement ou plus déterminent conjointement les finalités et les moyens du traitement, ils sont les responsables conjoints du traitement. [...] »  
(cf. CNIL, Le règlement général sur la protection des données).

## 2.2. Les obligations des acteurs

La qualification des acteurs – responsable (conjoint) de traitement ou sous-traitant – permet d'en déduire leurs obligations respectives au regard de la réglementation relative à la protection des données. Cette réglementation repose sur deux textes principaux : le **RGPD** et la **loi Informatique et Libertés**.

### ■ Les exigences du RGPD

Tant le responsable du traitement que le sous-traitant sont soumis aux dispositions du RGPD.

**Pour le sous-traitant**, les principales obligations sont les suivantes<sup>(7)</sup> :

- il agit pour le compte et sur instructions documentées du responsable du traitement ;
- il met en œuvre des mesures techniques et organisationnelles appropriées pour garantir la sécurité des données ;
- il assiste le responsable du traitement dans le respect de ses obligations, notamment en matière de sécurité, de réalisation des analyses d'impact sur la protection des données (le cas échéant), de gestion des demandes d'exercice des droits et des violations de données ;
- etc.

**Pour le responsable du traitement**, les obligations incluent notamment le respect des principes fondamentaux du RGPD :

- licéité, i.e., disposer d'une base légale et d'une exception au titre des articles 6 et 9 du RGPD pour traiter les données dans le cadre de l'étude de faisabilité ;
- transparence, i.e., informer les personnes concernées du traitement de leurs données aux fins de l'étude de faisabilité, en leur fournissant l'ensemble des éléments d'information prévus à l'article 13 ou 14 du RGPD, d'une façon concise, transparente, compréhensible et aisément accessible<sup>(8)</sup> ;
- minimisation, i.e., traiter les seules données strictement nécessaires pour répondre à la finalité poursuivie, à savoir évaluer la faisabilité d'une recherche ;
- limitation de la durée de conservation ;
- etc.

### ■ Les exigences de la loi Informatique et Libertés

La loi Informatique et Libertés dispose que les traitements de données personnelles de santé, dont les études de faisabilité font partie, font l'objet de formalités préalables<sup>(9)</sup> auprès de la CNIL. Ces formalités relèvent de l'un des cas de figure ci-dessous :

- déclaration de conformité à un cadre de référence établi par la CNIL (par exemple le référentiel RS-002 relatif aux entrepôts de données de santé ou la méthodologie de référence MR-004)<sup>(10)</sup> ;
- à défaut de cadre de référence existant ou en cas de non-conformité au cadre de référence applicable, autorisation du traitement par la CNIL.

L'accomplissement des formalités préalables est une obligation qui incombe au responsable de traitement, qu'il s'agisse du détenteur de données, de l'industriel ou de tout autre acteur.

**Les exemples ci-dessous illustrent les formalités qui doivent être accomplies en fonction de(s) acteur(s) impliqué(s), dans les cas listés au paragraphe 2.1.**

#### CAS N°1

##### **Le détenteur de données est seul responsable du traitement**

Le détenteur de données s'assure que l'évaluation de la faisabilité est bien intégrée dans les finalités couvertes par la formalité initiale accomplies lors de la constitution de sa base de données.

Ces formalités peuvent être, selon les cas, une déclaration de conformité au référentiel RS-002 (si la base initiale est un entrepôt de données de santé), une déclaration de conformité à la MR-004 (si la base de données initiale est une cohorte par exemple) ou encore une autorisation de la CNIL.

(7) Les obligations du sous-traitant sont prévues à l'article 28 du RGPD.

(8) Cette information peut par exemple être effectuée via la remise d'une note d'information individuelle ou via un portail de transparence.

(9) Voir notamment les fiches disponibles sur le site de la CNIL ou du HDH concernant les formalités réglementaires, ainsi que l'information des personnes.

(10) La déclaration de conformité à un cadre de référence est effectuée une fois pour toutes, et non pour chaque projet ou traitement.

## CAS N°2

### L'industriel est seul responsable du traitement

L'industriel étant responsable de traitement, il doit s'assurer de réaliser les formalités préalables requises pour l'étude de faisabilité.

Dans la mesure où l'étude de faisabilité est menée sur des données existantes et répond à la définition de « recherche n'impliquant pas la personne humaine », les formalités peuvent être, selon les cas :

- une déclaration de conformité à la MR-004 ;
- une autorisation de la CNIL après avis du CESREES<sup>(11)</sup>.

Lorsque l'étude de faisabilité est indissociable du projet de recherche en tant que tel, les formalités applicables seront celles requises par le projet de recherche.

Il peut s'agir, selon les cas, d'une déclaration de conformité à la MR-001 (après avis d'un CPP), à la MR-003 (après avis d'un CPP), à la MR-004 ou une autorisation de la CNIL (après avis du CESREES).

## CAS N°3

### Le détenteur et l'industriel sont responsables conjoints du traitement

Les formalités applicables sont similaires à celles décrites au cas 2 ci-dessus, et relèvent de la responsabilité conjointe du détenteur de données et de l'industriel.

En cas de conformité à un cadre de référence, chacune des parties devra procéder à une déclaration de conformité auprès de la CNIL (si cette démarche n'avait encore jamais été réalisée auparavant). Dans le cas contraire, une demande d'autorisation devra être soumise à la CNIL, au nom des deux responsables conjoints de traitement.

## 2.3. Éléments de contractualisation entre le détenteur de données et l'industriel

La contractualisation entre le détenteur de données et l'industriel du médicament dépend du niveau d'engagement attendu, de la nature de la prestation et du rôle des parties au regard du RGPD.

Lorsque la prestation reste limitée (par exemple, la fourniture d'indicateurs standards) et que les parties ne sont pas dans une relation de responsables conjoints de traitement ou de responsable de traitement/sous-traitant, la signature d'un **devis détaillant les conditions de réalisation et de rémunération peut suffire à encadrer la collaboration**.

*A contrario*, une étude de faisabilité menée à la demande spécifique d'un industriel et qui nécessite de mobiliser des ressources conséquentes au sein du détenteur de données, sera *a priori* régie par un contrat de prestation de services.

Compte-tenu de la diversité des situations possibles et des procédures internes à chaque acteur, il convient de consulter le département juridique pour déterminer l'encadrement contractuel approprié. De même, lorsqu'un **contrat de prestation de services** est établi, le détail de son contenu est élaboré au cas par cas.

(11) Comité éthique et scientifique pour les recherches, études et évaluations dans le domaine de la santé

Il est possible de distinguer les principales situations suivantes :

#### CAS N°1

### Le détenteur de données est seul responsable du traitement

Aucune clause relative à la protection des données n'est requise au titre du RGPD. Toutefois, il est recommandé que les parties établissent clairement leurs rôles respectifs (détenteur = responsable de traitement et industriel = NA) et les obligations qui incombent au détenteur de données.

Le devis ou contrat liant le détenteur de données et l'industriel pourra par exemple inclure une clause qui établit :

- que le détenteur de données reconnaît agir en qualité de seul responsable du traitement dans le cadre de l'étude de faisabilité,
- que le détenteur s'engage, à ce titre, à respecter l'ensemble des obligations prévues par le RGPD et la loi Informatique et Libertés, notamment en termes de licéité, de transparence et de formalités préalables.

#### CAS N°2

### L'industriel est seul responsable du traitement et le détenteur des données est sous-traitant

Le traitement de données par un sous-traitant doit être régi par un contrat, qui définit l'objet et la durée du traitement, la nature et la finalité du traitement, le type de données à caractère personnel et les catégories de personnes concernées, ainsi que les obligations et les droits du responsable du traitement (industriel) et du sous-traitant (détenteur de données), conformément à l'article 28 du RGPD<sup>(13)</sup>. L'encadrement contractuel de la sous-traitance est requis indépendamment de l'existence d'un contrat de prestation de services. Lorsqu'un contrat de prestation de services existe, le contrat requis au titre de l'article 28 du RGPD peut y être intégré (sous forme de clause ou d'annexe).

À titre d'exemple, le contrat de prestation de services peut inclure les clauses suivantes :

- **Objet de la prestation** : une définition claire de l'étude de faisabilité et de ses finalités ;
- **Livrables attendus** : par exemple le fait que les résultats doivent être fournis exclusivement sous forme agrégée, sans données individuelles directement ou indirectement identifiantes<sup>(12)</sup> ;
- **Propriété intellectuelle** : par exemple le fait que les livrables et analyses demeurent la propriété exclusive du détenteur de données ;
- **Délais de réalisation** : il est recommandé que les résultats de faisabilité soient transmis dans un délai inférieur à trois mois, en fonction des démarches réglementaires requises et du temps de contractualisation, le cas échéant ;
- **Modalités de financement** : les conditions de rémunération doivent être précisées, en cohérence avec le travail réalisé et la grille tarifaire du détenteur de données ;
- **Confidentialité** : toute communication de l'industriel ou du détenteur de données doit se faire dans le respect de l'accord de confidentialité défini ;
- **Protection des données personnelles** : la signature d'une clause relative à la protection des données personnelles dépend exclusivement du rôle de chaque acteur au regard du RGPD (responsable [conjoint] de traitement ou sous-traitant).

#### CAS N°3

### Le détenteur de données et l'industriel sont responsables conjoints du traitement

Lorsque les parties agissent en qualité de responsables conjoints, elles doivent définir de manière transparente leurs obligations respectives aux fins d'assurer le respect des exigences du RGPD par voie d'accord entre elles, conformément à l'article 26 du RGPD<sup>(14)</sup>. L'encadrement de la responsabilité conjointe est requis indépendamment de l'existence d'un contrat de prestation de services. Lorsqu'un contrat de prestation de services existe, l'accord requis au titre de l'article 26 du RGPD peut y être intégré (sous forme de clause ou d'annexe).

(12) Les données mises à disposition doivent en effet être anonymes, au sens de l'avis du G29 sur les techniques d'anonymisation. Pour plus d'informations, le HDH propose un [guide d'évaluation du caractère anonyme d'un jeu de données](#), ainsi que des [bonnes pratiques d'ouverture de bases en open data](#).

(13) Cf. CNIL, [Le règlement général sur la protection des données](#).

(14) Cf. CNIL, [Le règlement général sur la protection des données](#).

# 3 Déroulé d'une étude de faisabilité

L'étude de faisabilité commence par la formulation d'une demande par l'industriel et se termine par la formalisation des résultats issus de l'évaluation de la faisabilité de la base de données. Cette partie présente les grandes étapes d'une étude de faisabilité, accompagnées de bonnes pratiques pour sécuriser et optimiser la démarche.



## Étape 1 : Demande de réalisation d'une étude de faisabilité

Les premiers échanges dans le cadre d'une étude de faisabilité permettent de **clarifier les objectifs scientifiques**, de préciser les besoins en données de santé et les contraintes temporelles, et d'évaluer la capacité de la base à répondre aux exigences du projet. Au-delà des aspects techniques, ces discussions abordent également les **modalités de collaboration** : conditions financières, partenariats éventuels, contraintes réglementaires et dispositions contractuelles à anticiper.

Afin de **faciliter la préparation des premiers échanges entre industriels et détenteurs de données dans le cadre d'une étude de faisabilité**, une checklist de documents et informations utiles à partager a été élaborée. Elle synthétise les points clés à anticiper et aborder par chacune des parties.

Les éléments figurant dans les checklists ci-dessous n'ont pas vocation à être obligatoirement partagés. Il revient aux parties prenantes de déterminer les documents nécessaires à la réa-

lisation de l'étude de faisabilité. Les checklists permettent avant tout de dresser un panorama indicatif des informations susceptibles d'être partagées et d'en faciliter le recueil.

## CHECKLIST STANDARDISÉE DES INFORMATIONS ET DOCUMENTATIONS À PARTAGER PAR LE DÉTENTEUR DE DONNÉES AVEC L'INDUSTRIEL

### Aspects réglementaires

- ☐ Signature d'un accord de confidentialité (si proposé ou requis)
- ☐ Disponibilité des autorisations réglementaires
- ☐ Existence d'une notice d'information patient
- ☐ Disponibilité du consentement patient (lorsque requis)
- ☐ Existence d'un portail de transparence pour les patients

### Informations générales sur la base de données de santé

- ☐ Transmission de la fiche synthèse de la base - Présentation générale de la base (*contexte de constitution, statut et gouvernance de la base, publications, modalités de financement, ...*)
- ☐ Identification d'un interlocuteur unique
- ☐ Communication de la liste des partenariats non-confidentiels existants, si applicable
- ☐ Liste d'appariements à d'autres bases de données envisagés et/ou réalisés
- ☐ Informations sur les modalités d'accès aux données

### Description de la base de données de santé

- ☐ Transmission du synopsis / protocole de collecte (*liste des variables collectées, critères d'éligibilité et d'exclusion, etc.*), si applicable
- ☐ Transmission de la version finale du (e)CRF (*Case Report Form*) ou du dictionnaire de données
- ☐ Transmission du formulaire de renseignement « Exploration bases de données » renseigné (*cf. Annexe 3*)
- ☐ Indication sur la structuration des données (*structurées / non structurées / semi-structurées*)

### Qualité des données

- ☐ Informations sur la maintenance de la base de données, son contrôle qualité et sa profondeur
- ☐ Informations sur les ressources disponibles pour le monitoring de la base et la capacité à mener des analyses

### Contraintes calendaires

- ☐ Communication de contraintes calendaires éventuelles - Délais inhérents à la réalisation de l'étude
- ☐ Communication de délais inhérents à des comités scientifiques et éthiques ou de pilotage (périodicité de revue des projets) - Circuit d'approbation d'un éventuel projet
- ☐ Délai estimé de conduite d'une étude de faisabilité

### Aspects financiers

- ☐ Transmission d'une grille tarifaire pour les études de faisabilité
- ☐ Conditions du partenariat en cas de projet de recherche



## CHECKLIST STANDARDISÉE DES INFORMATIONS ET DOCUMENTATIONS À PARTAGER PAR L'INDUSTRIEL AVEC LE DÉTENTEUR DE DONNÉES

Contexte du projet	
<input type="checkbox"/>	Description du contexte et des objectifs de l'étude de faisabilité
<input type="checkbox"/>	Description du projet de recherche : study concept/synopsis
Données attendues	
<input type="checkbox"/>	Définition de la population cible (étudiée)
<input type="checkbox"/>	Événements du parcours de soins et de l'évolution de la pathologie des patients à documenter (principal et secondaires)
<input type="checkbox"/>	Variables clés recherchées et type de données attendues
<input type="checkbox"/>	Temporalité souhaitée - Période d'intérêt pour les données
Contraintes spécifiques	
<input type="checkbox"/>	Contraintes techniques ou réglementaires spécifiques
<input type="checkbox"/>	Délais de réalisation souhaités pour l'étude de faisabilité - Contraintes calendaires
<input type="checkbox"/>	Délais de réalisation souhaités pour le projet de recherche - Contraintes calendaires
<input type="checkbox"/>	Besoin d'appui méthodologique ou technique

## Étape 2. Faisabilité du projet de recherche

Cette étape vise à déterminer si la base de données est réellement exploitable pour répondre aux objectifs du projet de recherche, en tenant compte de la qualité des données, des capacités techniques, des ressources mobilisables et des contraintes opérationnelles.

Il s'agit d'une phase d'**évaluation technique**, permettant de tester concrètement l'utilisabilité de la base au regard de l'usage envisagé. L'objectif est d'apprécier la pertinence des données, au-delà de la simple acceptation de seuils ou de la facilité d'utilisation.

À ce stade, la question de recherche est bien définie et l'industriel est en mesure de :

- **définir avec précision les caractéristiques des données nécessaires pour la conduite du projet de recherche** : population étudiée, méthode d'échantillonnage, groupe de traitement/exposition, groupe comparateur, durée de suivi, délai des données, etc. ;

- **définir les variables d'intérêt spécifiques au projet**, dans le but de faciliter la compréhension du domaine thérapeutique ;
- **pré-spécifier, lorsque possible, l'importance de la qualité des données** en anticipant les méthodes d'analyse (par exemple, calculs de taille d'échantillon, utilisation de critères temporels (par exemple, calculs de taille d'échantillon, utilisation de critères temporels, période d'observation nécessaire pour mesurer un événement, cohérence entre les dates des différentes variables, analyses de sensibilité, ajustements statistiques pour les erreurs de mesure).

La faisabilité technique nécessite généralement de mobiliser des **moyens humains** et des **compétences spécifiques** pour assurer plusieurs étapes clés. Une analyse de faisabilité approfondie est généralement rémunérée, sauf lorsque le détenteur de données précise que les informations demandées sont déjà disponibles et que leur génération n'entraîne aucun surcoût.



## Qualification de la source de données

La qualification de la source peut être réalisée en évaluant si la qualité des données des variables d'intérêt est adéquate pour l'analyse prévue.

L'évaluation de l'adéquation à l'usage peut être réalisée à partir des métriques et métadonnées disponibles pour une source globale ou sur le sous-ensemble final de données sélectionné pour l'étude (par exemple, extraction spécifique, sous-population ciblée, agrégation de sources).

Cette évaluation peut inclure :

- La mise à disposition par le détenteur de données d'un jeu de données d'exemple anonymisé ou fictif, un « bac à sable », permettant à l'industriel d'apprécier les caractéristiques intrinsèques du jeu de données (structure, complétude, granularité, actualisation, etc.) et d'évaluer leur adéquation avec les types d'analyse envisagés ;
- La réalisation de statistiques exploratoires (analyses descriptives et de complétude, etc.) par le détenteur sur des jeux de données ;
- L'analyse de la capacité de la base à répondre aux critères méthodologiques du projet, tels que la fréquence des événements, le suivi longitudinal, la disponibilité des variables clés et leur complétude, etc. ;
- L'identification des limitations techniques éventuelles, comme, l'hétérogénéité des sources de données, etc.

## Étape 3. Restitution des résultats de l'étude de faisabilité

À l'issue de la démarche, cette étape consiste à formaliser les résultats obtenus et à les transmettre à l'industriel. Elle permet de statuer sur la poursuite ou non du projet de recherche.

La restitution prend généralement la forme d'un **rapport synthétique** qui peut inclure :

- une présentation générale de la gouvernance de la base ;
- l'évaluation de la complétude de la base et de la qualité des données ;
- l'évaluation des aspects pratiques et organisationnels liés à la faisabilité.

### ■ Présentation générale de la gouvernance de la base

Cette section rappelle l'objectif initial de la constitution de la base dans laquelle s'inscrit l'étude de faisabilité. Elle souligne l'importance de comprendre le fonctionnement général de la base, les objectifs de la constitution de la base, la méthode de collecte (manuelle, automatisée, via IA)

et les acteurs impliqués (techniciens, cliniciens...). Elle met également en évidence la diversité des sources et des centres qui alimentent la base, la répartition des lieux de collecte pour garantir la représentativité et limiter les biais, ainsi que la capacité de chaînage avec d'autres bases, comme le SNDS.

Pour assurer la transparence, la section présente les références et le cadre réglementaire : exemples de publications associées, modalités d'accès et conformité aux exigences CNIL, CPP et portail de transparence. Elle décrit aussi les ressources mobilisées pour le *data management* (outils, équipes, procédures de contrôle qualité) et les éventuelles reconnaissances officielles (HAS). Enfin, elle aborde l'archivage et la traçabilité, confirmant que les données extraites à des fins analytiques sont conservées et accessibles pour des analyses supplémentaires ou des répliques, conformément aux règles définies dans la PIA (*Privacy Impact Assessment*).

## ■ L'évaluation de la taille, de la complétude de la base et de la qualité des données

Cette section vise à présenter des composantes essentielles permettant de garantir la **puissance statistique** et la **validité des résultats**.

Le tableau ci-dessous synthétise les critères essentiels à examiner pour évaluer la faisabilité à partir d'une base de données. Il regroupe les éléments clés permettant de garantir la pertinence

scientifique et la puissance statistique (population, variables, durée, représentativité) et la qualité et complétude des données, qui assure la fiabilité des analyses (taux de complétude, fraîcheur, cohérence, suivi).

Ces éléments constituent des repères pratiques pour décider du « Go/No-Go » et anticiper les ajustements nécessaires avant le lancement du projet.

Dimension	Éléments clés à vérifier
<b>Population cible</b>	<ul style="list-style-type: none"> <li>→ Estimation du nombre de patients inclus selon les critères d'inclusion/exclusion</li> <li>→ Représentativité : évaluation des risques de biais liés à la source de données et à la couverture (géographique, types de centres, caractéristiques des patients)</li> </ul>
<b>Variables d'intérêt</b>	<ul style="list-style-type: none"> <li>→ Nature et disponibilité (diagnostic, traitement, événements)</li> <li>→ Taux de complétude globale (ex : complétude <math>\geq 90\%</math>) et taux de complétude par variable (ex : complétude <math>\geq 90\%</math>)</li> <li>→ Limites connues de la base (estimation d'un % de données manquantes sur les données de suivi)</li> <li>→ Type (structurées / semi-structurées / non structurées)</li> </ul>
<b>Durée &amp; suivi</b>	<ul style="list-style-type: none"> <li>→ Période d'observation / de collecte</li> <li>→ Suivi longitudinal ou non</li> <li>→ Date de début (et fin) d'inclusion des patients, si applicable</li> <li>→ Fréquence du suivi (événements ou consultations)</li> <li>→ Statut vital</li> </ul>
<b>Qualité des données</b>	<ul style="list-style-type: none"> <li>→ % de données manquantes</li> <li>→ Définition des « perdus de vue » et leur taux de présence dans la base, etc.</li> <li>→ Contrôles qualité et cohérence</li> </ul>
<b>Temporalité &amp; fraîcheur</b>	<ul style="list-style-type: none"> <li>→ Date de dernière mise à jour des données</li> <li>→ Fréquence des mises à jour de la base</li> <li>→ Délai (moyen) avant saisie dans la base pour l'inclusion d'un patient, si applicable</li> <li>→ Délai (moyen) avant saisie pour les visites de suivi des patients, si applicable</li> </ul>

## ■ Évaluation de la mise en œuvre opérationnelle et de l'impact financier du projet de recherche envisagé

Au-delà des critères scientifiques, l'étude de faisabilité peut inclure une évaluation détaillée de la mise en œuvre opérationnelle et de l'impact financier du projet de recherche.

L'étude de faisabilité vise à fournir une analyse approfondie de la capacité réelle d'une base de données à répondre aux besoins d'un projet de recherche. Elle doit dépasser les informations générales issues du CRF ou des échanges préliminaires, en apportant une réponse structurée et contextualisée.

## Proposition opérationnelle et financière

### Cette proposition doit inclure :

- **Chiffrage du coût induit** par la poursuite du projet avec la base de données analysée, en cas de conclusion positive de l'étude de faisabilité.
- **Estimation des délais de réalisation** du projet de recherche, en prenant en compte les démarches réglementaires et contractuelles ainsi que l'obtention des éventuelles validations internes de la source de données.
- **Évaluation des ressources nécessaires**, comprenant :
  1. **Charge de travail estimée** : ressources humaines et techniques requises pour extraire, traiter et analyser les données conformément aux standards attendus.
  2. **Mobilisation des équipes** : capacité du détenteur de données à accompagner le projet (temps lié à la saisie ou à la complétion des données, expertises à mobiliser selon les méthodes d'analyse prévues, disponibilité des équipes et des outils).

Sur la base des résultats de l'étude de faisabilité, **trois issues sont possibles** :

1. **Faisabilité confirmée** : le projet peut être poursuivi avec la base analysée. Les parties peuvent alors engager la phase de contractualisation du projet.
2. **Faisabilité conditionnelle** : le projet est réalisable moyennant certaines adaptations (modifications de la méthode ou du protocole si nécessaire afin d'exploiter les données disponibles, enrichissement des données, **mobilisation nécessaire de ressources humaines supplémentaires pour la complétion de la base**, etc.).
3. **Non-faisabilité** : la source choisie ne correspond pas à la question de recherche ou ne permet pas de répondre aux objectifs du projet.

Il convient de noter que la réalisation d'une étude de faisabilité n'engage en aucun cas l'industriel à poursuivre le projet de recherche, quels que soient les résultats obtenus. La faisabilité constitue une étape en amont dont les conclusions peuvent orienter la décision finale sans la conditionner. Toutefois, une étude de faisabilité rigoureuse et bien documentée peut constituer un levier stratégique apportant les garanties supplémentaires pour la poursuite d'un projet.



### BOITE À OUTILS

- ✓ **Checklist standardisée de faisabilité** (cf. *Tableau 1 : Checklist standardisée des informations et documentations à partager par le détenteur de données avec l'industriel* et *Tableau 2 : Checklist standardisée des informations et documentations à partager par l'industriel avec le détenteur de données*) ;
- ✓ **Formulaire de renseignement « Exploration bases de données »** (cf. *Annexe 3*) ;
- ✓ **Modèle type d'étude de faisabilité** (cf. *Annexe 4*), pour laquelle la faisabilité a été menée.

# Conclusion

Ce guide méthodologique propose un cadre partagé pour optimiser la réalisation des études de faisabilité. Il constitue un outil opérationnel conçu pour renforcer les collaborations entre les industriels et les détenteurs de données, au service d'un écosystème des données de santé attractif et favorisant la vitalité des autres acteurs.

Au-delà de son rôle méthodologique, ce guide traduit une démarche collective visant à favoriser les partenariats entre industriels et détenteurs de données, dans un esprit de transparence et d'efficacité. Cette harmonisation des pratiques à l'échelle nationale s'inscrit dans une dynamique plus large : la construction de l'Espace Européen des Données de Santé. Adopté le 11 février 2025 avec une entrée en application à partir de mars 2027, ce cadre réglementaire européen amplifie les enjeux portés par ce guide en ce qu'il accompagnera l'accès transfrontalier aux données de santé pour la recherche et l'innovation.

En effet, l'Espace Européen des Données de Santé implique :

- une harmonisation des pratiques à l'échelle européenne ;
- des exigences d'interopérabilité et de standardisation des données ;
- un renforcement des garanties en matière de sécurité et de confidentialité ;
- la mise à disposition d'un catalogue de données, reflétant la pertinence et la qualité des bases.

Les principes méthodologiques proposés dans ce guide constituent une préparation à cette transformation européenne. Ils devront évoluer à moyen/long termes pour prendre en compte cette nouvelle dimension tout en conservant leur vocation première : faciliter des collaborations et des études de qualité.

# Annexes

## Annexe 1 GLOSSAIRE

### **Cohortes, registres et entrepôts de données de santé thématiques**

Les registres et cohortes (et entrepôts de données de santé thématiques) collectent des données de santé pour permettre de répondre à des objectifs scientifiques précis, avec des protocoles de collecte et des variables prédéfinies (Questionnaires, *Case Report Forms* [CRF]).

### **Détenteurs de données**

Le terme « détenteur de données » désigne les établissements ou structures (hôpitaux, centres de recherche, autres organismes publics ou privés, etc.) qui collectent, structurent et gèrent des données de santé. Ces détenteurs de données sont en mesure de mobiliser leurs bases pour répondre à une demande d'étude de faisabilité. Les données pourraient également être mobilisées pour le projet de recherche en santé pour lequel la faisabilité aurait été réalisée.

### **Étude de faisabilité**

L'étude de faisabilité pour des projets de recherche utilisant des données de santé existantes est une démarche structurée menée en amont, visant à vérifier dans quelle mesure une base de données existante est adaptée et suffisante pour répondre aux objectifs du projet, dans les conditions scientifiques, réglementaires et techniques envisagées. Toutefois, elle ne doit pas se substituer au projet de recherche en répondant à la question scientifique pour laquelle la faisabilité a été menée.

### **Entrepôts de données de santé hospitaliers**

Les entrepôts de données de santé hospitaliers centralisent notamment des données issues des systèmes d'information hospitaliers couvrant de multiples pathologies et situations cliniques. Ces données n'ont pas été initialement collectées à des fins de recherche mais pour des besoins de soins et de gestion.

### **Industriel**

Le terme « industriel » désigne les entreprises du secteur de la santé, notamment pharmaceutiques, qui souhaitent porter un projet de recherche s'appuyant sur des données de santé pour en évaluer la faisabilité avant mise en œuvre.

### **MR-004**

Méthodologie de référence de la CNIL qui encadre les traitements de données à caractère personnel réalisés dans le cadre de recherches n'impliquant pas la personne humaine (RNIPH), ainsi que des études et évaluations dans le domaine de la santé.

### **Projet de recherche en santé**

Un projet de recherche en santé mené par un industriel du médicament est une initiative scientifique visant à générer des connaissances nouvelles ou à valider des hypothèses dans le domaine médical ou pharmaceutique, généralement en lien avec le développement ou l'évaluation de médicaments, dispositifs médicaux ou stratégies thérapeutiques. Ces projets peuvent s'appuyer sur des données de santé secondaires, c'est-à-dire des données initialement collectées pour d'autres finalités (soins, facturation, suivi administratif) et réutilisées à des fins de recherche.

### **Référentiel de santé RS-002**

Référentiel publié par la CNIL qui encadre les entrepôts de données de santé. Il définit les exigences juridiques, organisationnelles et techniques permettant de mettre en place et d'exploiter un entrepôt de données de santé sans passer par une demande d'autorisation spécifique, à condition de respecter l'ensemble des dispositions qu'il prévoit.

### **Responsable de traitement**

Personne physique ou morale qui détermine les finalités et les moyens du traitement des données utilisées.

### **Sous-traitant**

Personne physique ou morale qui traite des données pour le compte d'un autre organisme (« le responsable de traitement »), dans le cadre d'un service ou d'une prestation.

### **Utilisation secondaire de données de santé**

Lorsqu'elles sont réutilisées à des fins de recherche, d'innovation, de santé publique, d'élaboration des politiques ou de médecine personnalisée, les données de santé relèvent d'une utilisation dite « secondaire ».

### **Variable d'intérêt**

Une variable d'intérêt dans un projet de recherche en santé désigne un élément mesurable ou observable qui est sélectionné parce qu'il est pertinent pour répondre à la question scientifique ou aux objectifs de l'étude.

## Annexe 2

# LISTE DES ABRÉVIATIONS ET ACRONYMES

- **AP-HP** : Assistance publique – Hôpitaux de Paris
- **BNDMR** : Banque nationale de données maladies rares
- **CESREES** : Comité éthique et scientifique pour les recherches, études et évaluations dans le domaine de la santé
- **CHU** : Centre hospitalier universitaire
- **CNIL** : Commission nationale de l'informatique et des libertés
- **CPP** : Comité de protection des personnes
- **CRF** : *Case Report Form*
- **CRO** : *Contract Research Organization* (organisation de recherche contractuelle)
- **CSF** : Comité stratégique de filière industries et technologies de santé
- **CSP** : Code de la santé publique
- **EDSH** : Entrepôts de données de santé hospitaliers
- **HAS** : Haute Autorité de Santé
- **HDH** : *Health Data Hub*
- **IA** : Intelligence artificielle
- **LYSARC** : *The Lymphoma Academic Research Organisation*
- **MR-004** : Méthodologie de référence MR-004
- **PIA** : *Privacy Impact Assessment*
- **PSCC** : Paris Saclay Cancer Cluster (PSCC)
- **RESPIC** : Association RESPIC pour la recherche dans les établissements privés d'intérêt collectif
- **RGPD** : Règlement général sur la protection des données
- **RS-002** : Référentiel de santé RS-002
- **SNDS** : Système national des données de santé

## Annexe 3

# FORMULAIRE DE RENSEIGNEMENT

## « EXPLORATION BASE DE DONNÉES »

Ce formulaire est proposé à titre d'exemple et constitue un simple modèle illustrant les grandes étapes du processus. Il n'a pas vocation à être utilisé de manière systématique. Son objectif est d'être adapté et personnalisé en fonction des besoins spécifiques de chaque étude de faisabilité pour laquelle il serait utilisé.

Catégorie	Questions	Réponses apportées par le détenteur de données
ADMINISTRATIF	Quel est le nom de la base de données ?	
	Qui est le contact référent de cette base de données ? (Nom/e-mail)	
	Pour quel objectif/pour quelle finalité la base de données a-t-elle été constituée ?	
	La création de la base a-t-elle fait l'objet d'une validation par une autorité compétente ?	
	Quelle est le nom de l'entité juridique responsable de la base de données ?	
	Quelles sont les modalités d'information patient pour la constitution de cette base de données ? Et pour la réalisation de recherches ultérieures ?	
	Quelle est le nom de l'entité juridique responsable de la base de données ?	
	Comment est financée cette base de données actuellement ?	
GOUVERNANCE	Y a-t-il d'autres partenaires liés à cette base de données ? Si oui, quel est leur rôle ?	
	Quelles sont les modalités d'accès pour un industriel ?	
	Quelle est la gouvernance de cette base de données ?	
	Des activités sont-elles sous-traitées à un prestataire ? Si oui, lesquelles ?	
CONTENU	À quelle date a débuté la collecte des données de cette base ?	
	Cette base est-elle multicentrique ? Quels centres y participent ?	
	Quel est le type de collecte ? Rétrospective, prospective ou les deux ? Sources et outils utilisés ?	
	Quels sont les critères d'inclusion et d'exclusion des pa-tients ?	
	Quel est le volume de données de cette base ?	
	Quelles sont les variables collectées dans cette base de données ?	
	La constitution de cette base est-elle toujours en cours ? Quelle est sa fréquence de mise à jour ?	
	La base est-elle chaînée à d'autres bases ou a-t-elle pro-jet de l'être ? Si oui, lesquelles ?	
PUBLICATIONS	Existe-t-il des publications en lien avec cette base de données ? Si oui, pourriez-vous les préciser avec des références disponibles ?	

## Annexe 4

# MODÈLE TYPE D'ÉTUDE DE FAISABILITÉ

Ce modèle est proposé à titre indicatif et illustre les principales composantes d'une étude de faisabilité. Il n'a pas vocation à être appliqué de manière systématique. Son objectif est d'être adapté et personnalisé en fonction des spécificités de chaque étude de faisabilité pour laquelle il serait utilisé.

### ÉVALUATION DE LA FAISABILITÉ

#### Présentation des parties impliquées

Entité	Représentée par	Nom/Prénom	Adresse	Courriel
[Industriel]	[Fonction]	[x]	[x]	[x]
[Détenant de données]	[Fonction]	[x]	[x]	[x]

#### Présentation générale de la base

Ce tableau permet d'obtenir une vue d'ensemble relativement complète de la base de données. La liste des informations peut être adaptée selon le contexte.

Identification	
Nom de la base	[x]
Objectifs de constitution	[x]
...	
Protocole et collecte	
Protocole	[x]
Critères (inclusion/exclusion)	[x]
Méthode de collecte (manuelle, automatisée, IA)	[x]
Acteurs impliqués	[x]
...	
Sources et représentativité	
Origine des données	[x]
Répartition des lieux de collecte	[x]
Types de centres	[x]
Caractéristiques des patients	[x]
Capacité de chaînage (ex. SNDS)	[x]
...	

Références et cadre réglementaire	
Publications associées	[x]
Consentement	[x]
Autorisations (CNIL, CPP)	[x]
...	
Ressources et qualité	
Publications associées	[x]
Consentement	[x]
Autorisations (CNIL, CPP)	[x]
...	
Ressources et qualité	
Outils	[x]
Équipes	[x]
Procédures de contrôle qualité	[x]
Analyses statistiques	[x]
Présence dans recensements officiels	[x]
...	
Archivage et traçabilité	
Règles de conservation	[x]
Mention des règles de conservation dans la PIA ( <i>Privacy Impact Assessment</i> )	[x]
Archivage des données extraites	[x]
...	



## Question de recherche, objectifs et besoins de l'industriel

Suite à l'identification des besoins de l'industriel pour son projet de recherche, le tableau ci-dessous présente les principaux éléments évalués.

Catégorie	Éléments recherchés (qualitatifs/quantitatifs)	Capacité de la base à répondre aux besoins
<b>Population cible</b>	<ul style="list-style-type: none"> <li>– Estimation du nombre de patients selon les critères définis</li> <li>– Taille de la population cible</li> <li>– Représentativité</li> </ul>	[x]
<b>Variables d'intérêt</b>	<ul style="list-style-type: none"> <li>– Nature et disponibilité des variables</li> <li>– Taux de complétude des données clés (diagnostics, traitements, événements cliniques)</li> </ul>	[x]
<b>Type de données</b>	<ul style="list-style-type: none"> <li>– Structurées vs non structurées</li> <li>– Autres caractéristiques</li> </ul>	[x]
<b>Durée et suivi</b>	<ul style="list-style-type: none"> <li>– Durée suffisante pour répondre à la question de recherche</li> <li>– Fréquence des événements et consultations</li> </ul>	[x]
<b>Taille et représentativité de la cohorte</b>	<ul style="list-style-type: none"> <li>– Évaluation des risques de biais liés à la source et à la couverture</li> <li>– Répartition géographique</li> <li>– Types de centres</li> <li>– Caractéristiques des patients</li> </ul>	[x]

## Évaluation de la complétude et de la qualité des données

Le tableau ci-dessous présente l'évaluation de la complétude et de la qualité de la base de données, en s'appuyant sur les critères définis lors de la démarche de faisabilité.

Catégorie	Éléments évalués	Définition des attentes (quantitatif/qualitatif)	Complétude / Qualité
<b>Qualité des données</b>	<ul style="list-style-type: none"> <li>– Taux de complétude des variables clés (diagnostics, traitements, événements, décès)</li> <li>– Limites connues (données manquantes, défauts de structuration)</li> <li>– Définition et taux des patients « perdus de vue »</li> <li>– Connaissance du statut vital des patients</li> </ul>	[x]	[ % ou Oui/Non]
<b>Temporalité</b>	<ul style="list-style-type: none"> <li>– Période de collecte</li> <li>– Début d'inclusion</li> <li>– Durée moyenne de suivi</li> <li>– Fréquence des visites</li> <li>– Suivi longitudinal ou non</li> <li>– Possibilité de collecte rétrospective</li> </ul>	[x]	[Oui/Non]
<b>Actualité des données</b>	<ul style="list-style-type: none"> <li>– Fraîcheur des données</li> <li>– Dernière mise à jour</li> <li>– Fréquence des mises à jour</li> <li>– Délais moyens de saisie (inclusion et suivi)</li> </ul>	[x]	[ % ou Oui/Non]

## Annexe 4 (suite)

### MODÈLE TYPE D'ÉTUDE DE FAISABILITÉ

#### Évaluation des aspects pratiques et organisationnels de la faisabilité

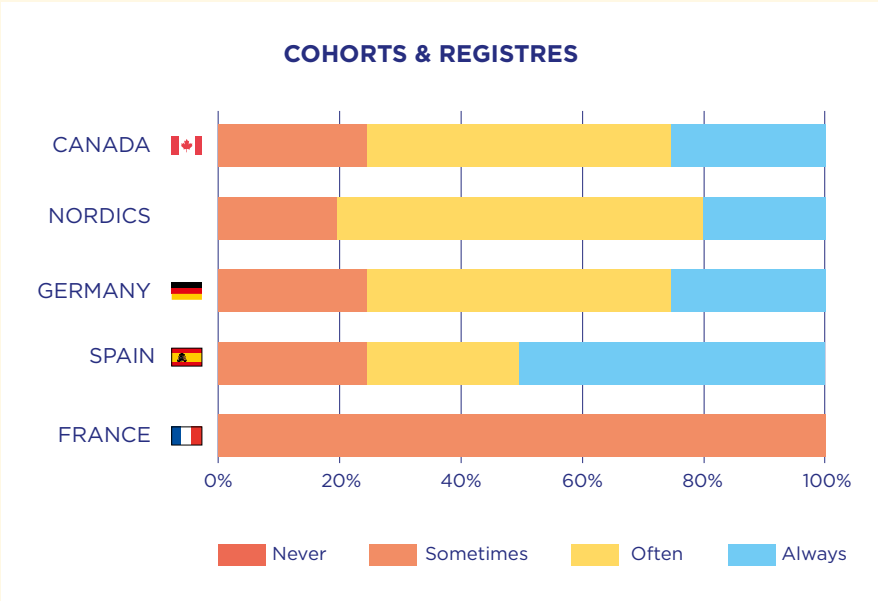
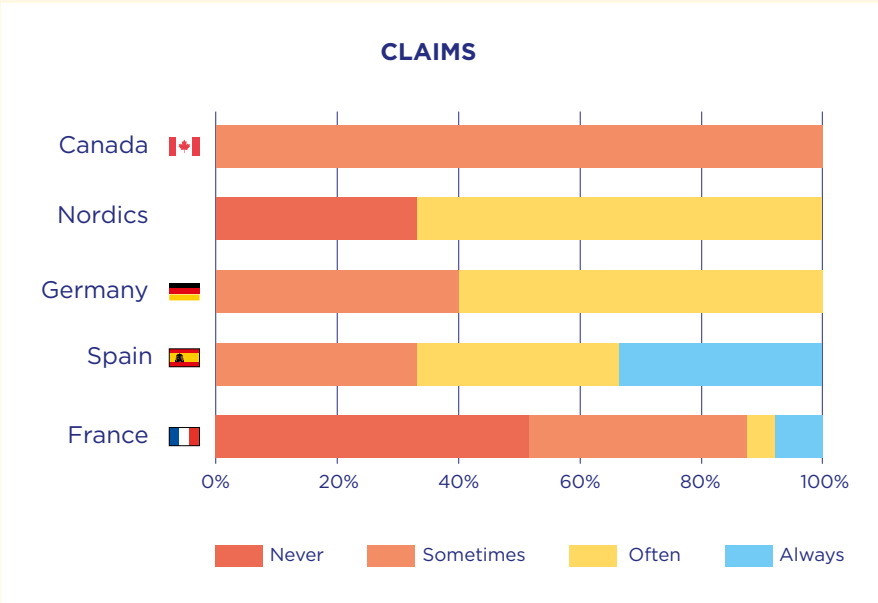
Le tableau ci-dessous présente l'évaluation des aspects pratiques et organisationnels liés à la faisabilité du projet, y compris la possibilité de réaliser un *Dry run* statistique.

Catégorie	Éléments évalués	Évaluation / Réponse
<b>Dry run statistique</b>	– Possibilité d'effectuer un <i>Dry run</i> statistique	[x]
<b>Charge de travail</b>	– Estimation du volume de travail (heures/jours) – Ressources humaines nécessaires (ARC, data managers, statisticiens) – Ressources techniques (logiciels, serveurs, outils d'analyse)	[x]
<b>Mobilisation des équipes</b>	– Capacité à mobiliser des équipes pour compléter les données manquantes – Temps estimé pour saisie complémentaire – Capacité à mobiliser les expertises nécessaires (statistique, data science)	[x]
<b>Contraintes organisationnelles</b>	– Disponibilité des équipes dans le calendrier projet – Coordination entre industriel et détenteur de données	[x]
<b>Coût potentiel</b>	– Impact financier lié à la complétude ou à la récupération des données manquantes	[Estimation]

# FRÉQUENCE DE MISE EN ŒUVRE D'ÉTUDES DE FAISABILITÉ SELON LES PAYS EN 2024

## LES RAISONS DE LA MOINDRE ATTRACTIVITÉ DE LA FRANCE ÉTUDES DE FAISABILITÉ

La France apparaît comme le pays le plus défavorable pour les études de faisabilité préalable à l'initiation d'un projet, à l'inverse de l'Espagne.



# leem

[www.leem.org](http://www.leem.org)

## Retrouvez-nous

 Leem

 LeemFrance

 LeemFrance

Pour toute question sur le guide, contacter Mélody Monthéard :  
[mmontheard@leem.org](mailto:mmontheard@leem.org)