HEALTH
DATA HUB

ANALYSE D'IMPACT RELATIVE A LA PROTECTION DES DONNEES DANS LE CADRE DE LA CONSTITUTION DE L'ENTREPÔT DE DONNÉES DE SANTÉ EMC2

Sommaire

1. Contexte **4**
1.1. Vue d'ensemble 4
1.1.1. Quel est le traitement qui fait l'objet de l'étude ? 4
1.1.2. Quelles sont les responsabilités liées aux traitements ? 6
1.1.3. Quels sont les référentiels applicables ? 12
1.2. Données, processus et supports 12
1.2.1. Quelles sont les données traitées ? Description des données, destinataires et durées de conservation 12
1.2.2. Quel est le cycle de vie des données ? 22
1.2.3. Quels sont les supports de données ? 28
2. Principes fondamentaux **30**
2.1. Mesures garantissant la proportionnalité et la nécessité du traitement 30
2.1.1. Finalités et fondements 30
2.1.2. Les données collectées sont-elles adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données) ? 34
2.1.3. Les données sont-elles exactes et tenues à jour ? 38
2.1.4. Quelle est la durée de conservation des données ? 39
2.2. Mesures protectrices des droits des personnes concernées 41
2.2.1. Comment les personnes concernées sont-elles informées à propos du traitement ? 41
2.2.2. Si applicable, comment le consentement des personnes concernées est-il obtenu ? 46
2.2.3. Comment les personnes concernées peuvent-elles exercer leurs droits d'accès et droit à la portabilité ? 46
2.2.4. Comment les personnes concernées peuvent-elles exercer leurs droits de rectification et droit à l'effacement (droit à l'oubli) ? 49
2.2.5. Comment les personnes concernées peuvent-elles exercer leurs droits de limitation et droit d'opposition ? 52
2.2.6. Les obligations des sous-traitants sont-elles clairement définies et contractualisées ? 54
2.2.7. En cas de transfert de données en dehors de l'Union européenne, les données sont-elles protégées de manière équivalente ? 55
3. Etude des risques liés à la sécurité des données **58**
3.1. Mesures existantes ou prévues 58
3.1.1. Mesures contribuant à traiter des risques liés à la sécurité des données 58
3.1.1.1. Chiffrement 58
3.1.1.2. Anonymisation 59

3.1.1.3. Cloisonnement des données (par rapport au reste du SI) 60
3.1.1.4. Contrôle des accès logiques 61
3.1.1.5. Journalisation 62
3.1.1.6. Contrôle d'intégrité 63
3.1.1.7. Archivage 63
3.1.1.8. Sécurité des documents papiers 64
3.1.1.9. Pseudonymisation 64
3.1.2. Mesures générales de sécurité 66
3.1.2.1. Sécurité de l'exploitation 66
3.1.2.2. Lutte contre les logiciels malveillants 66
3.1.2.3. Sauvegardes 67
3.1.2.4. Maintenance 67
3.1.2.5. Sécurité des canaux informatiques (réseaux) 68
3.1.2.6. Sécurité physique 68
3.1.2.7. Traçabilité 69
3.1.2.8. Sécurité du matériel 69
3.1.2.9. Eloignement des sources de risque 70
3.1.2.10. Protection contre les sources de risque non humaines 70
3.1.3. Mesures organisationnelles (gouvernance) 71
3.1.3.1. Organisation 71
3.1.3.2. Politique (gestion des règles) 71
3.1.3.3. Gérer les risques 71
3.1.3.4. Intégrer la protection de la vie privée dans les projets 72
3.1.3.5. Gestion des incidents de sécurité et les violations de données 72
3.1.3.6. Gestion des personnels 73
3.1.3.7. Relation avec les tiers 73
3.1.3.8. Superviser la protection de la vie privée 73
3.2. Accès illégitime à des données 73
3.3. Modification non désirée de données 76
3.4. Disparition de données 77
4. Formalisation de la validation 79

1. Contexte

1.1. Vue d'ensemble

En 2021, le Health Data Hub (ci-après "HDH") a décidé de mettre en œuvre l'entrepôt de données de santé (EDS) EMC2 (Entrepôt Multi-Centrique Chaîné), à savoir une base de données multicentrique, standardisée et chainée avec la base principale du SNDS (Système National des Données de Santé).

Dans cette perspective, le HDH a créé un partenariat avec 7 acteurs dont :

- quatre établissements de santé (ci-dessous "établissements de santé" ou "centres") ;
- deux structures en charge d'apporter leur expertise ;
- une structure d'appui au pilotage, identifiée comme partenaire.

La constitution de cet EDS répond à des besoins exprimés par l'écosystème en termes de mutualisation des données afin de constituer des bases de données de taille critique. Il permet en outre de répondre à un besoin crucial de l'Agence Européenne du Médicament (ci-après "EMA"), formalisé au travers d'un appel à projet visant à constituer une base utilisable pour conduire des études de pharmaco-épidémiologie. Ayant remporté cet appel d'offres, le HDH bénéficie d'un financement de l'EMA afin de financer une partie des activités menées dans le cadre de ce projet.

Les données alimentant cet entrepôt pourront ensuite être réutilisées dans le cadre de recherches, études et évaluations dans le domaine de la santé après réalisation des formalités nécessaires au titre des articles 66 et 72 et suivants de la loi Informatique et Libertés modifiée. Ces formalités peuvent être aussi bien un avis du CESREES et une autorisation de la CNIL que la conformité à la MR004.

1.1.1. Quel est le traitement qui fait l'objet de l'étude ?

Le traitement réside donc dans la mise en œuvre d'un entrepôt de données de santé (ci-après "EDS"), alimenté à la fois par le Système National des Données de Santé et par les bases de données des quatre établissements de santé.

La constitution d'un entrepôt de données hospitalières multicentriques est un enjeu majeur pour le développement de la recherche "en vie réelle" en France. Au sein du HDH, l'entrepôt EMC2 regroupe des données cliniques et paracliniques liées à la prise en charge intra-hospitalière des patients, enrichies des données de la base principale du SNDS, ouvrant de nouvelles perspectives pour la recherche en santé.

La base principale du SNDS est un système de bases de données médico-administratives parmi les plus volumineuses et exhaustives du monde, liées aux remboursements des actes et des soins des bénéficiaires de l'ensemble des régimes d'assurance maladie obligatoire. Il présente un intérêt majeur pour la recherche et l'innovation en santé en France notamment dans les domaines épidémiologique, pharmaco-épidémiologique et médico-économique.
Néanmoins, l'absence de données cliniques ou de données sur les résultats des examens complémentaires au sein de la base principale du SNDS limite la portée de nombreuses études. L'entrepôt EMC2 vise à combler - au moins en partie - ce manque en appariant les données cliniques et paracliniques fournies par quatre établissements de santé français avec

les données individuelles de la base principale du SNDS pour les patients concernés, complétées d'un échantillon de la base principale du SNDS afin de pouvoir mener des analyses comparatives en population générale.

La mise en œuvre de l'entrepôt EMC2 a pour finalité de permettre la réutilisation des données qu'il contient à des fins de recherche, d'étude et d'évaluation dans le domaine de la santé. Tout responsable de traitement réalisant les formalités adéquates et en premier lieu les établissements hospitaliers partenaires, les équipes de recherche internes à l'EMA pourront réaliser ces études. Il s'agira en particulier de la conduite d'études d'intérêt public de différentes natures telles que :

- l'observation et l'évaluation de la prise en charge des patients,
- la caractérisation des populations de patients,
- l'évaluation de l'utilisation / des pratiques, de l'efficacité et de la sécurité en vie réelle (hôpital et ville) des produits de santé (en particulier médicaments et dispositifs médicaux (DM) inscrits au remboursement ou en accès précoce),
- l'évaluation médico-économique (efficience et 'impact budgétaire) des produits de santé administrés (médicaments) ou utilisés (DM) en vie réelle,
- le développement et l'évaluation d'actions de prévention pour les pathologies d'intérêt,
- la conception et la validation d'outils d'aide à l'interprétation des signaux, au diagnostic ou à la prise en charge préventive ou curative.

La liste précédente des études ne se veut pas exhaustive : elle a été déterminée au regard des problématiques d'importance identifiées par les établissements hospitaliers partenaires ainsi que par l'EMA. L'entrepôt EMC2 sera ouvert à toute réutilisation des données dans une finalité d'intérêt public et pour lequel son périmètre est pertinent : contribuer à la recherche, aux études, à l'évaluation et à l'innovation dans les domaines de la santé et de la prise en charge médico-sociale.

Par rapport à des études équivalentes conduites sur la base principale du SNDS seule, l'entrepôt EMC2 permettra des analyses sur des sous-groupes de patients non identifiables à partir des seules données du PMSI mais repérables à partir des critères cliniques ou paracliniques transmis par les 4 établissements de santé. Par exemple, pour de nombreuses pathologies, le PMSI (via la caractérisation CIM-10 des diagnostics lors des séjours hospitaliers) ne permet pas de différencier les patients selon la sévérité ou le stade de la maladie (insuffisance cardiaque, cancer, etc.). Les données cliniques ou les résultats d'examens fournis par les 4 établissements de santé de l'entrepôt EMC2 permettront d'identifier et de caractériser des sous-groupes de patients jusque là non identifiables et donc de réaliser des analyses stratifiées selon différents phénotypes ou stades de sévérité, et ainsi d'affiner les études.

Pour constituer cet EDS, le HDH a déposé une demande d'autorisation sur le fondement de l'article 66-III de la loi du 6 janvier 1978 modifiée, qui soumet à autorisation les traitements comportant des données relatives à la santé, justifiés par l'intérêt public et non-conformes à un référentiel.

La conformité du présent EDS a été réalisée sur la base de la délibération n° 2021-123 du 2 novembre 2021 portant rectification du 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 adopté le 7 octobre 2021. Eu égard, entre autres, au traitement de

données du SNDS, l'EDS présente certains écarts avec le référentiel susvisé, lesquels sont détaillés dans un document distinct et justifient la demande d'autorisation.

La plateforme technologique qui sera utilisée pour la constitution de l'EDS [redacted] Elle a été homologuée le 5 novembre 2020 et cette homologation a fait l'objet d'une révision le 30 novembre 2021.

La présente AIPD couvre tous les traitements qui seront réalisés sur la plateforme technologique pour permettre au HDH de constituer une base de données multicentrique, standardisée et chainée avec la base principale du SNDS. Cet entrepôt de données de santé permettra, par exemple, à l'EMA de traiter des questions relatives à la pharmacovigilance et à la pharmaco-épidémiologie, et aux établissements de santé partenaires et autres porteurs de projet de conduire des études revêtant un caractère d'intérêt public.

Les traitements dont la responsabilité incombe au HDH sont les suivants :

- Traitement 1 - [redacted]
- Traitement 2 - [redacted]
- Traitement 3 - [redacted]
  - Sous-finalité 1 :
  - Sous-finalité 2 :
  - Sous-finalité 3 :
- Traitement 4 - [redacted]
  - Sous-finalité 1 :
  - Sous-finalité 2 :

1.1.2. Quelles sont les responsabilités liées aux traitements ?

Les parties prenantes qui auront des rôles et responsabilités au sens du RGPD et qui interviennent dans le traitement des données personnelles sont les suivantes :

- Le HDH, en tant que responsable de traitement ;
- Les quatre établissements de santé (HCL, CLB, CHRU de Nancy, FHSJ), en tant que sous-traitants ;
- La société Microsoft, en tant que sous-traitant ;
- La société CDC Arkhinéo, en tant que sous-traitant.

➢ **Responsable de traitement de l'EDS**

L'EDS EMC2 est mis en œuvre par le HDH, en tant que responsable de traitement :

Nom de l'établissement	Type de structure	Représenté par (nom et fonction)	Coordonnées du RT	Coordonnées du DPO	Résumé du rôle dans l'EDS
Health Data Hub	Groupement d'intérêt public	Stéphanie COMBES, Directrice	9 rue Georges Pitard 75015 Paris	dpd@health-data-hub.fr	Assure en tant que responsable de traitement, l'organisation, le pilotage et la

					mise à disposition des données de l'EDS

Le HDH est responsable de traitement de l'EDS, dans la mesure où il en a déterminé les finalités et les moyens :

- Concernant les finalités, la mise en œuvre de l'EDS s'est opérée à l'initiative du HDH qui s'est rapproché de quatre établissements de santé partenaires. La constitution de l'EDS répond en effet à des besoins de l'écosystème remontés au HDH : besoin de mutualisation des données afin de constituer des bases de données de taille critique pour la recherche, l'optimisation du parcours de soin ou le pilotage de l'activité hospitalière, besoin de compléter les données administratives avec des données cliniques, besoin d'accéder à des données standardisées dans des formats internationaux permettant de répondre à des études pour le compte d'autorités réglementaires européennes. Enfin, l'EDS multicentrique s'est d'abord construit autour d'une famille de cas d'usage identifiée comme parmi les plus stratégiques, à savoir l'utilisation des données de vie réelle pour l'évaluation et la surveillance de la sécurité et de l'efficacité des technologies de santé, et en particulier du médicament (pharmacovigilance), en situation réelle de soins, permettant ainsi de répondre et de remporter un appel d'offres de l'EMA. Cette thématique a pris d'autant plus d'importance avec la crise sanitaire.
- Sur les moyens, le HDH définit le périmètre des données, le périmètre des personnes concernées (détaillé dans la partie 3.2 du présent document), les destinataires des données (les porteurs de projet qui auront effectué les formalités préalables adéquates) et la durée de conservation. Par ailleurs, l'EDS sera hébergé sur la plateforme technologique du HDH.

➢ **Sous-traitants de l'EDS**

Six acteurs sont identifiés comme sous-traitants du traitement de mise en œuvre de l'entrepôt :

- Les 4 établissements de santé chargés de fournir les données dans l'EDS EMC2 (HCL, CLB, CHRU de Nancy, FHSJ) ;
- Microsoft [redacted] Cette plateforme est hébergée dans les centres de données de Microsoft Azure situés en Zone France dans la région « France Centre » (région parisienne).

- CDC Arkhinéo, chargé d'archiver électroniquement les données de journalisation des utilisateurs de la plateforme technologique du HDH.

Dans le cadre de la mise en œuvre de l'entrepôt, les HCL, le CLB, le CHRU de Nancy et la FHSJ, en leur qualité de fournisseurs de données, Microsoft, en sa qualité d'hébergeur des données, et CDC Arkhinéo, en tant qu'archiveur des traces d'activité, traiteront des données à caractère personnel pour le compte, sur instruction et sous l'autorité du responsable de traitement. Ces sous-traitants inscriront les traitements réalisés au sein de leur registre des activités de traitement conformément à l'article 30.2 du RGPD.
Conformément à l'article 28 du RGPD, la description des traitements, les obligations de ces sous-traitants y compris en matière de sécurité et de gestion des violations de données sont formalisées au sein de contrats.

Si l'on suit le cycle de vie de la donnée depuis les établissements de santé jusqu'à la réalisation d'un projet sur la plateforme technologique du HDH, les différentes étapes peuvent être résumées de la manière suivante :

1. Chez le responsable de données : préparation d'une copie des données en vue de leur ingestion et stockage sur la plateforme technologique du HDH
2. Collecte et conservation des données pseudonymisées par le HDH dans l'EDS :
    a. mise en place d'un flux d'ingestion sécurisé pour le transfert de la copie des données
    b. masquage des identifiants
    c. stockage des données
3. Préparation et exposition des données au porteur de projet sur la plateforme technologique du HDH :
    a. mise en place d'un espace de préparation et d'un espace d'analyse
    b. transfert d'une copie des données nécessaires au projet dans l'espace de préparation
    c. masquage des identifiants
    d. préparation des données (le cas échéant)
    e. transfert d'une copie des données vers l'espace d'analyse
4. Réalisation d'une étude par un porteur de projet sur la plateforme technologique du HDH

En outre, afin de sécuriser les relations avec les différentes parties prenantes dans le cadre de la réutilisation des données, les conditions générales d'utilisation (CGU) de la plateforme technologique du HDH sont signées par les porteurs de projet avant tout accès à la plateforme.

Traitement 1- [redacted]

a) <u>S'agissant des établissements de santé</u> :

Les établissements de santé ont un rôle de fournisseurs de données : ils réalisent, en tant que sous-traitants, les traitements jusqu'à l'alimentation de l'EDS. Les établissements de santé gèrent la préparation des données en vue de leur réplication dans l'EDS, sur la plateforme technologique du HDH. Ils s'engagent donc à collecter et mettre en qualité les données en amont de leur transfert.
Dans la mesure où les établissements de santé communiquent directement les données au HDH, ils devront s'assurer [redacted] que la base de données ne contient aucune donnée directement identifiante comme des noms, prénoms

ou numéros de sécurité sociale. Dans la mesure où le circuit de pseudonymisation de la Cnam devra être mobilisé, l'établissement de santé devra en outre attribuer à chaque personne un identifiant technique temporaire ou « numéro d'accrochage » qui permettra d'entrer dans le circuit de pseudonymisation de la Cnam et de faire le lien avec les données de santé transmises au HDH.
Les établissements de santé sont chargés de la dé-identification des données et, le cas échéant, de l'attribution d'un identifiant temporaire avant envoi de ces informations au HDH et à la Cnam respectivement. Le HDH les accompagne et les conseille pendant ces opérations préliminaires.

b) S'agissant du HDH :

Au stade de la collecte et de la conservation des données, le HDH est chargé de la mise en place d'un flux d'ingestion sécurisé pour le transfert des données vers la plateforme technologique et du masquage des pseudonymes à leur arrivée.

Traitement 2 - [redacted]

a) S'agissant du HDH :

Le HDH gère l'EDS et met à disposition des porteurs de projet les données contenues dans l'EDS. Il peut également exploiter les données pour en améliorer la qualité, faciliter leur réutilisation ou produire des requêtes simples.
La convention d'utilisation de la plateforme technologique qui sera conclue entre le HDH et le porteur de projet décrira les rôles et responsabilités de chacun dans le sens d'une responsabilité distincte de traitement.

b) S'agissant des porteurs de projet :

Le porteur de projet va réaliser son étude au sein d'un espace d'analyse qui lui est dédié sur la plateforme technologique. Le porteur de projet est responsable de la réalisation de cette étude au sein de son espace d'analyse.
Le porteur de projet n'a aucune responsabilité sur l'hébergement des données et la sécurité de la plateforme technologique mais :

- il garantit tout au long de l'étude que seuls les utilisateurs habilités accèdent à l'espace d'analyse ;
- il vérifie que les utilisateurs respectent au quotidien les bonnes pratiques en matière de sécurité de leur compte et de leur espace de travail (non partage des identifiants de connexion, préservation de la sécurité logicielle du poste de travail, verrouillage systématique de la session, etc.) ;
- il s'assure de la bonne mise en œuvre des mesures de sécurité, de gouvernance et d'organisation décrites, notamment grâce aux indicateurs de sécurité mis à disposition par le HDH.

Le porteur de projet met à jour trimestriellement la liste des utilisateurs et garantit que les utilisateurs participent aux actions de sensibilisation menées par le porteur de projet et le HDH en matière de sécurité et de protection des données.
Le porteur de projet veille à ce que les utilisateurs utilisent les services de la plateforme technologique de manière proportionnée aux besoins des tâches qui leur incombent afin de ne pas saturer les ressources de ces environnements de travail, notamment en tirant le meilleur parti des indicateurs d'utilisation des ressources mises à disposition par le HDH.
Le porteur de projet est responsable des imports de données anonymes dans son espace

d'analyse. A cet égard, il s'engage à veiller à l'intégrité des données importées et à la qualité de leur anonymisation.
Enfin, le porteur de projet est aussi responsable des exports de résultats générés par l'étude. En particulier, il doit veiller à ce que les exports soient strictement anonymes. La responsabilité en incombe au porteur de projet mais le HDH l'accompagne par des conseils et contrôle les exports.

c) S'agissant des établissements de santé :

Sous réserve que leurs traitements soient encadrés par une autorisation de la CNIL ou une méthodologie de référence, les établissements de santé sont susceptibles d'accéder eux-mêmes à leurs données sur la plateforme technologique aux fins de mener leurs propres projets ou à des fins de mise en qualité et préparation des données. Le cas échéant, ils devront respecter les CGU et ils disposeront de leur propre espace d'analyse (s'agissant de la réalisation d'un projet) ou de leur propre espace de préparation (s'agissant de la mise en qualité des données). Pour chacun de ces espaces de travail, l'établissement de santé concerné sera seul responsable des traitements qu'il réalise.

Traitement 3 - [redacted]

a) Sous-finalité 1 : [redacted]

[redacted] afin de contribuer aux missions du HDH de réunir, organiser et mettre à disposition les données du SNDS (par la facilitation de la compréhension et de l'usage des données présentes au catalogue, en fournissant le contexte et la documentation nécessaires), de contribuer à diffuser les normes de standardisation pour l'échange et l'exploitation des données de santé, en tenant compte des standards européens et internationaux, et de promouvoir l'innovation dans l'utilisation des données de santé. Ces missions restent à construire mais les pistes d'ores et déjà envisagées sont les suivantes :

- mettre à disposition un catalogue de métadonnées (informations sur la structure, statistiques descriptives, etc) interne et/ou externe à la plateforme [redacted] à des utilisateurs externes ayant les autorisations nécessaires pour leur permettre de contribuer à la production de métadonnées et ouvrir la version externe pour améliorer la connaissance générale des données de santé ;
- formater et standardiser les données [redacted] ;
- transformer les données pour améliorer la qualité métier (redressement, normalisation, création de nouvelles variables, etc.) en partenariat avec les responsables des données, développer et tester les logiques d'appariement direct ou indirect entre les bases de données à disposition et développer des clés de jointure [redacted] permettant d'augmenter la qualité et la simplicité d'usage des données.

b) Sous-finalité 2 : [redacted]

[redacted] sont destinataires des données de l'EDS afin de contribuer aux missions du HDH de promouvoir l'innovation dans l'utilisation des données de santé et de faciliter la mise à disposition de jeux de données. Ces missions restent à construire mais les pistes d'ores et déjà envisagées sont les suivantes :

- organiser des activités de découvertes et de formation en permettant aux participants

d'envoyer leurs algorithmes, de les lancer [redacted] sur la plateforme technologiques du Health Data Hub et de partager les performances obtenues (i.e., le "score" que l'algorithme a obtenu). Cela permettrait de confronter la pertinence des algorithmes développés dans le cadre de ces activités ludiques (data challenge, hackathon, etc.) à des données réelles tout en préservant la confidentialité et la sécurité de ces dernières.

- réaliser des analyses basiques pour répondre à des questions simples d'intérêt général pour lesquelles une démarche d'autorisation serait disproportionnée [redacted] ;
- créer des indicateurs en réalisant des analyses basiques pour répondre à des questions simples d'intérêt général pour lesquelles une démarche d'autorisation serait disproportionnée [redacted]
- ouvrir des données en générant et mettant à disposition des données synthétiques sur la base des données ou de leurs métadonnées, des jeux de données anonymisées [redacted], et en mettant à disposition au sein de la plateforme technologique des données à des utilisateurs externes ayant les autorisations nécessaires pour leur permettre de contribuer à la production de données synthétiques ou anonymisées.

c) Sous-finalité 3 : [redacted]

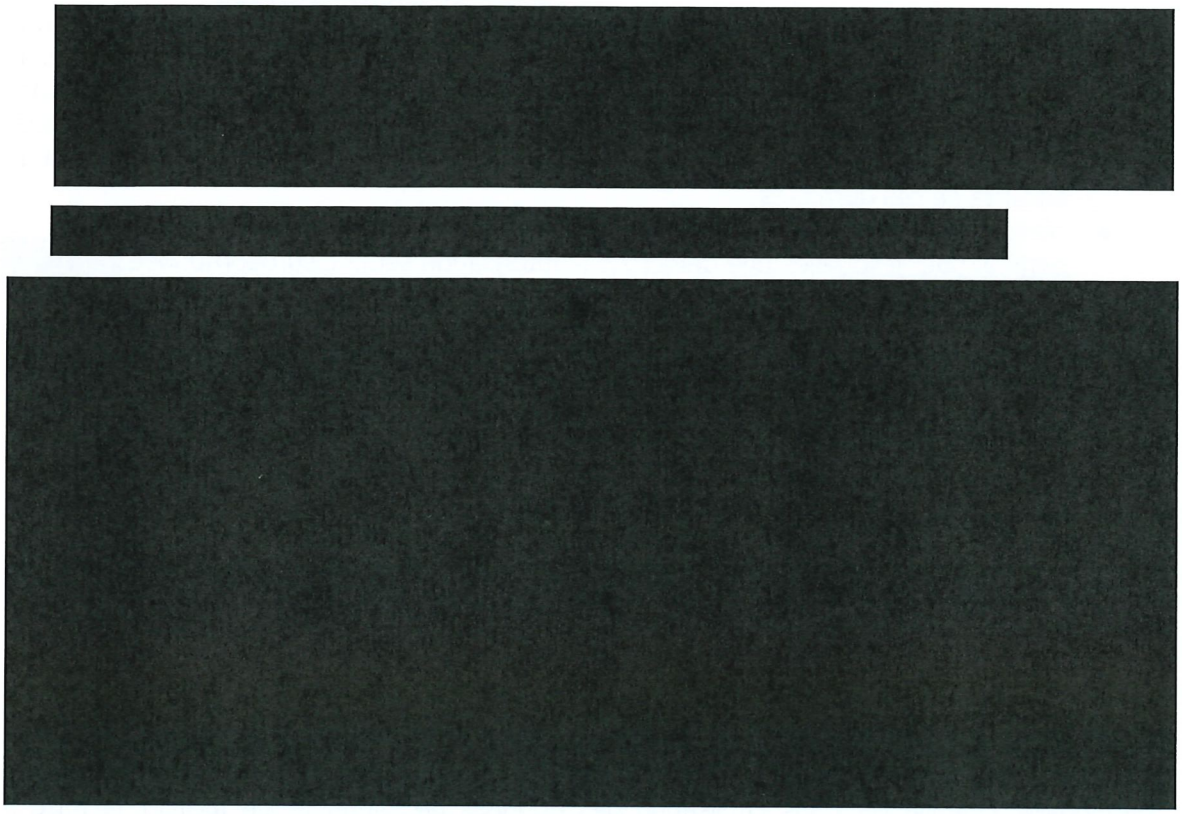
[redacted] sont destinataires des données de l'EDS afin de contribuer à la mission aux missions du HDH de promouvoir l'innovation dans l'utilisation des données de santé, et de faciliter la mise à disposition de jeux de données de santé présentant un faible risque d'impact sur la vie privée l'exploitation des données. Ces missions restent à construire mais les pistes d'ores et déjà envisagées sont les suivantes :

- mettre à disposition des données en transférant des données dans un autre environnement sécurisé en partenariat avec l'établissement de santé ;
- développer des outils de traitement des données en mettant en place et en partageant une bibliothèque de programmes et d'algorithmes permettant de traiter et de faciliter les interactions avec les données et/ou leurs métadonnées (génération, description, manipulation, standardisation, vérification, exploitation, etc.). Ces outils permettront de développer et de partager des applications de requêtes/extraction de données via des interfaces-outils simplifiés.
- définir des méthodes de référence de valorisation des données (e.g., utilisation de données de vie réelle dans le cadre d'études cliniques).
- standardiser les données, par exemple en travailler sur leur format et en promouvant des [redacted]
- mettre en place un catalogue de métadonnées ;
- ouvrir des jeux de données (échantillons, anonymes et synthétiques).

Traitement 4 - [redacted]

a) Sous-finalité 1 : [redacted]

Le HDH est responsable de la gestion des accès à la plateforme technologique (authentification, gestion des droits, etc.). Des données à caractère personnel sont nécessaires pour la création du compte et la gestion des accès et des identités sur la plateforme technologique. Elles ne sont accessibles et utilisées que par les collaborateurs du HDH chargés de cette tâche :



1.1.3. Quels sont les référentiels applicables ?

Outre le RGPD et la loi Informatique et Libertés qui sont naturellement applicables, les référentiels encadrant les traitements de données par le HDH sont les suivants :

- Politique de sécurité des systèmes d'information de l'Etat du 17 juillet 2014 ;
- Arrêté du 1er octobre 2015 portant approbation de la politique de sécurité des systèmes d'information pour les ministères chargés des affaires sociales (PSSI MCAS) ;
- Référentiel général de sécurité V2 du 13 juin 2014 ;
- Arrêté du 22 mars 2017 relatif au référentiel de sécurité du SNDS (qui renvoie au RGS, à la PGSSI-Santé et à la PSSI-MCAS). A la date de rédaction de la présente AIPD, des travaux ont été lancés pour mettre à jour le référentiel de sécurité du SNDS ;
- Loi n°2019-774 du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé ;
- Décret n°2021-848 du 29 juin 2021 relatif au traitement de données à caractère personnel dénommé « système national des données de santé » ;
- Délibération CNIL 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.

1.2. Données, processus et supports

1.2.1. Quelles sont les données traitées ? Description des données, destinataires et durées de conservation

Traitement 1 - [redacted]

a) Données :

➢ **Sources des données :**

Les données [redacted] issues d'extractions :

- des données des quatre centres participants
- de la base principale du SNDS.

Une mise-à-jour annuelle des données des établissements de santé et de la base principale du SNDS est réalisée au sein de l'EDS EMC2. [redacted]

Les données des établissements de santé et de la base principale du SNDS [redacted].

Les tableaux suivants détaillent chacune des sources et la typologie des données qui sont utilisées dans le cadre du projet.

Source de données #1

Base de données	Dossiers patients des Hospices Civils de Lyon / Entrepôt de données de santé[1]
Producteur	Hospices Civils de Lyon (HCL)
Encadrement réglementaire	Pour les dossiers patients : Code de la santé publique, Sixième partie : Etablissements et services de santé (Articles L6111-1 à L6441-1) Pour l'EDS : engagement de conformité n°2227726
Collecte des données	Les données sont issues d'une collecte préalable dans le cadre du soin à l'hôpital.
[redacted]	[redacted]

Source de données #2

Base de données	Dossiers patients du centre Léon Bérard
Producteur	Centre Léon Bérard (CLB)
Encadrement réglementaire	Code de la santé publique, Sixième partie : Etablissements et services de santé (Articles L6111-1 à L6441-1)

[1] L'EDS des HCL est en cours de mise en œuvre. Si celui-ci n'est pas terminé lorsque les données devront être versées dans l'EDS EMC2, alors elles seront extraites directement du système d'information hospitalier.

Collecte des données	Les données sont issues d'une collecte préalable dans le cadre du soin à l'hôpital.
[redacted]	[redacted]

Source de données #3

Base de données	Dossiers patients de la Fondation Hôpital Saint Joseph
Producteur	Fondation Hôpital Saint Joseph (FHSJ)
Encadrement réglementaire	Code de la santé publique, Sixième partie : Etablissements et services de santé (Articles L6111-1 à L6441-1)
Collecte des données	Les données sont issues d'une collecte préalable dans le cadre du soin à l'hôpital.
[redacted]	[redacted]

Source de données #4

Base de données	Dossiers patients du Centre Hospitalier Régional Universitaire de Nancy
Producteur	Centre Hospitalier Régional Universitaire de Nancy (CHRU de Nancy)
Encadrement réglementaire	Code de la santé publique, Articles L6111-1 à L6441-1 (Sixième partie - Etablissements et services de santé - de la Partie législative)
Collecte des données	Les données sont issues d'une collecte préalable dans le cadre du soin à l'hôpital.
[redacted]	[redacted]

Source de données #5

Base de données	Système National des Données de Santé (SNDS) - Base principale
Producteur	Caisse nationale de l'Assurance Maladie (CNAM)
Encadrement réglementaire	Code de la santé publique, Articles L1461-1 à L1461-7 (Chapitre Ier - Système national des données de santé - du Titre IV du Livre IV de la Première partie de la Partie législative)

Typologie de données disponibles	La base principale du SNDS contient des données relatives aux : • Informations sur le bénéficiaire (sexe, mois et année de naissance, rang de naissance, lieu de résidence, date de décès, régime, couverture maladie universelle complémentaire, aide à la complémentaire santé, indice de défavorisation) ; • Informations sur les professionnels de santé (spécialité, mode d'exercice, sexe, âge, département d'implantation) • Pathologies, notamment les affections de longue durée et les diagnostics des séjours hospitaliers ; • Dépenses et remboursements des prestations en soins de ville, en établissements de santé (consommations de soins hospitaliers ou en ville, prescriptions, dispositifs médicaux, montants et indemnités,etc.) • Causes médicales de décès • Informations sur de potentiels handicaps
Collecte des données	La CNAM consolide et pseudonymise les données : - de l'Assurance Maladie (consommations de soins en ville et en établissement remontées dans le SNIIRAM -Système national d'information inter-régimes de l'Assurance maladie) depuis 2006 - hospitalières du PMSI (Programme de Médicalisation des Systèmes d'Information) issues de l'ATIH, depuis 2006 - sur les causes médicales de décès du CépiDC, depuis 2006 (aujourd'hui 2006-2020) - sur les données relatives au handicap (MDPH) issues des maisons départementales des personnes handicapées (pas encore intégrée au SNDS à date)
[redacted]	[redacted]

➢ **Données de l'EDS :**

Les données extraites et transmises sur la plateforme du HDH ne comportent pas de données directement identifiantes, et contiennent uniquement des données structurées pseudonymisées. Dans le cadre de l'appariement direct à la base principale du SNDS, les centres participants auront néanmoins vocation à extraire le NIR (Numéro d'Inscription au Répertoire), la date de naissance et le sexe des patients. Les NIRs seront transmis à la CNAM pour l'appariement à la base principale du SNDS et exclusivement utilisés à cet effet ; ils n'intègrent pas l'EDS EMC2.

Les données recueillies et transférées par les centres participants incluront :

- Les données transférées au HDH :

[redacted]

[redacted]

- Les données transférées à la CNAM pour l'appariement à la base principale du SNDS (flux non intégré dans l'EDS EMC2) :

[redacted]

Les données des centres participants permettent notamment d'obtenir des informations sur les résultats d'examens biologiques et cliniques, les facteurs de risque du patient, les indications de prescription qui ne sont pas présentes dans les données de la base principale du SNDS.

Les données extraites de la base principale du SNDS relatives à la population incluse contiennent :

[redacted]

Ces données couvrent la période [redacted] à partir de la date de mise à disposition des données pour l'ensemble de la population présente dans l'EDS afin d'avoir un suivi rétrospectif et prospectif au long cours des patients dans l'EDS.

Les données de la base principale du SNDS sont nécessaires pour répondre aux objectifs de l'EDS EMC2 sur l'étude du parcours de soins, notamment en apportant des données longitudinales exhaustives sur les soins de ville, les hospitalisations dans d'autres établissements de santé ou encore d'autres informations d'intérêt (causes médicales de décès, vaccination Covid, etc.).

Le niveau de détail sur les variables sensibles (au sens du référentiel SNDS) dans l'EDS EMC2 sera :

[redacted]

Les dates exactes de soins sont indispensables pour reconstruire et évaluer de manière précise le parcours de soins des patients. La date de décès permet de réaliser des études de

mortalité et des analyses de survie. Il sera également intéressant de disposer du mois de naissance (en plus de l'année) afin de calculer l'âge précis, qui peut avoir son importance, notamment pour des études incluant des enfants en bas-âge. Enfin, il sera nécessaire de disposer de la commune de résidence en vue de l'utilisation l'indice de défavorisation communal pour prendre en compte le niveau socio-économique au niveau de certaines études.

La convention bipartite signée entre le HDH et chaque établissement de santé doit notamment préciser le périmètre, la fréquence de mise à jour et la sécurisation du transfert des données. La qualité des données relève de la responsabilité des producteurs de données tels que la CNAM et les établissements de santé.

Concernant les données provenant des établissements de santé, il s'agit de données médicales et médico-administratives déjà collectées. Elles peuvent comporter des erreurs dues à la nature du processus de remontées d'informations qui seront toujours présentes dans la base de données issue de l'appariement. Toutefois, les erreurs présentes dans cette base sont sans impact sur les droits et libertés des personnes.

Vous trouverez ci-dessous les grandes catégories données visant à intégrer l'EDS EMC2 :

Catégorie de données	Description
Données relatives aux séjours hospitaliers	[redacted]
Informations relatives aux diagnostics	[redacted]
Données relatives aux traitements et actes médicaux	[redacted]
Examens biologiques et leurs résultats	[redacted]
Examens cliniques et leurs résultats	[redacted]
Dispositifs médicaux	[redacted]
Facteurs de risques	[redacted]
Évaluation des patients	[redacted]
Caractéristiques sociodémographiques du patient	[redacted]
Données de remboursement de soins de ville	[redacted]

Causes médicales de décès	[redacted]
Données relatives au handicap (MDPH)	[redacted]

Les variables sensibles (identifiants potentiels au sens du référentiel SNDS) intégrées dans l'EDS EMC 2 sont :

Variables sensibles (identifiants potentiels)
Année + mois de naissance
Date de soins (jj/mm/aaaa)
Date de décès (jj/mm/aaaa)
Commune de résidence

b) Destinataires :

Les destinataires de ces données sont les personnes intervenant dans la constitution et la mise en œuvre de l'entrepôt, à savoir :

- d'une part, au sein du HDH, les équipes techniques et scientifiques qui définissent le périmètre des données, leur format et qui, de manière générale, sont en charge de l'organisation de l'entrepôt, [redacted]
- d'autre part, les sous-traitants du HDH, soit :
  - Les 4 établissements de santé chargés de fournir les données dans l'EDS EMC2 (HCL, CLB, CHRU de Nancy, FHSJ) ;
  - Microsoft [redacted]

c) Durée de conservation :

Chaque centre hospitalier contribue à l'EDS EMC2 en fournissant des données structurées pour des patients ayant été hospitalisés dans le centre participant à partir de 2022. Les données hospitalières intégrées dans l'EDS ont été collectées [redacted], elles seront conservées dans l'EDS pour [redacted] pour chaque patient inclus, en fonction de la disponibilité des données dans les établissements partenaires. La temporalité des données et la période d'inclusion des patients est variable en fonction des centres, la date de début de la reprise d'historique des données cliniques [redacted] a été déterminée au regard de la disponibilité des données dans les systèmes sources hospitaliers.

Des données de la base principale du SNDS à partir de 2015 sont également extraites pour complémenter l'EDS EMC2, [redacted]. La date de début de la reprise d'historique des données SNDS a été déterminée, tout d'abord, au regard de l'exhaustivité des remontées des régimes dans le SNDS à partir de cette date ; avant, certains régimes - et donc populations - ne sont pas présents dans les données. [redacted] Les données de la base principale du SNDS permettront notamment de disposer de données sur les consommations de soins de ville, le parcours de soins dans d'autres établissements de santé et les causes médicales de décès (le cas échéant) pour les patients inclus dans l'EDS EMC2. [redacted]

La base principale du SNDS permet également de constituer une population témoin (et de disposer des données associées) ainsi que d'identifier et de récupérer des données relatives aux enfants des femmes sélectionnées dans les centres ayant accouché depuis 2015. Les données de la base principale du SNDS seront également récupérées pour ces deux populations à partir de 2015 [redacted].

Concernant les pseudonymes fournis par les établissements de santé, ils ne sont pas conservés sur la plateforme technologique. Les identifiants de stockage générés par le HDH sont quant à eux conservés pendant la durée de conservation de la base de données correspondante.

Traitement 2- [redacted]

a) Données :

Les données concernées sont les données de santé pseudonymisées contenues dans l'EDS. Il s'agit de données issues des bases de données des établissements de santé.
Lors de la préparation des données [redacted] au sein de l'espace de préparation, des identifiants spécifiques au projet sont générés en remplacement des identifiants de stockage. Ensuite, [redacted] déclenche un automate qui transfère les données voulues vers l'espace d'analyse cible. De nouveaux identifiants spécifiques à chaque espace d'analyse sont générés à chaque fois, si bien que les données relatives à une même personne seront rattachées à autant d'identifiants projet différents que le nombre d'espaces d'analyse où elles sont mises à disposition.

b) Destinataires :

Les destinataires de ces données sont les personnes intervenant dans la constitution et la mise en œuvre de l'entrepôt, à savoir :

- d'une part, au sein du HDH, les équipes techniques et scientifiques qui définissent le périmètre des données, leur format et qui, de manière générale, sont en charge de l'organisation de l'entrepôt, [redacted] Ensuite, les « utilisateurs projet » traitent les données pour mener leurs projets dans l'espace qui leur est dédié ;

- d'autre part, les sous-traitants du HDH, soit :
  - Les 4 établissements de santé chargés de fournir les données dans l'EDS EMC2 (HCL, CLB, CHRU de Nancy, FHSJ) ;
  - Microsoft [redacted]

Dans le cadre de la réutilisation des données de l'EDS, les utilisateurs porteurs du projet n'ont accès qu'à l'espace qui leur est dédié et n'ont accès à aucun autre espace sur la plateforme technologique. Cet espace permet ainsi d'assurer un cloisonnement strict des projets en s'appuyant sur les droits des utilisateurs qui y sont rattachés. Il n'est possible de mettre à disposition dans cet espace d'analyse que les données nécessaires à la tenue de ce projet.

c) Durée de conservation :

Pendant la phase de préparation et d'exposition, les données sont conservées sur la plateforme technologique en deux endroits distincts : d'abord au sein de l'espace de préparation [redacted], ensuite au sein de l'espace d'analyse pour les utilisateurs projet.

- Au sein de l'espace de préparation, la durée de conservation des données sera conditionnée par le besoin du projet en termes de mise à jour des données dans son espace d'analyse au cours de son cycle de vie :
  - Pour les projets ne nécessitant pas de mise à jour des données pendant leur déroulé (une livraison initiale et unique dans l'espace d'analyse suffit pour mener à bien l'étude) : [redacted]
    période pendant laquelle le porteur de projet peut solliciter des données corrigées et/ou complémentaires si cette demande est conforme aux besoins définis.
  - [redacted]

[redacted]

- Au sein de l'espace d'analyse, la durée de conservation est fondée sur le temps nécessaire à la conduite du projet. Cette durée de conservation diffère donc d'un projet à un autre et est déterminée par le responsable de traitement du projet sous le contrôle de la CNIL. Les identifiants projet générés par le HDH sont également conservés pour la durée du projet. [redacted]

Traitement 3 - [redacted]

a) Sous-finalité 1 : [redacted]

i) **Données** : données de l'EDS
ii) **Destinataires :** [redacted] au sein du HDH sont destinataires des données notamment pour mettre en place un catalogue de métadonnées, travailler sur le formatage des données et l'amélioration des données.
iii) **Durée de conservation :** déterminée projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

b) Sous-finalité 2 : [redacted]

i) **Données :** données de l'EDS.
ii) **Destinataires :** [redacted] au sein du HDH sont destinataires des données notamment pour réaliser des analyses basiques afin de répondre à des questions simples d'intérêt général ou pour créer des indicateurs pour lesquelles une démarche d'autorisation serait disproportionnée [redacted], organiser des activités de découverte et de formation en permettant aux participants d'envoyer leurs algorithmes, et mettre à disposition des données synthétiques ou anonymes notamment à des utilisateurs externes pour leur permettre de contribuer à la protection de ces données.
iii) **Durée de conservation :** déterminée projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

c) Sous-finalité 3 : [redacted]

i) **Données** : données de l'EDS
ii) **Destinataires :** [redacted] au sein du HDH sont destinataires des données notamment pour mettre à disposition des données dans un espace sécurisé en partenariat avec le responsable de données, de développer des outils de traitement des données et de définir les méthodes de référence de valorisation des données. travailler sur la standardisation des données, mettre en place un catalogue de métadonnées, ouvrir des jeux de données (échantillons, anonymes et synthétiques) ou pour mettre en place et

développer des outils d'exploitation des données par exemple.

iii) **Durée de conservation :** déterminée projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

Traitement 4 - [redacted]

a) Sous-finalité 1 : [redacted]

i) **Données** : les données concernées sont les données à caractère personnel des utilisateurs de la plateforme technologique, à savoir leurs nom, prénom, données de contact professionnelles (mail, adresse postale, numéro de téléphone), fonction, organisme de rattachement et identifiant du générateur de jeton logiciel.

ii) **Destinataires** : les destinataires de ces données sont [redacted]

iii) **Durée de conservation** : ces données seront conservées pendant toute la durée d'existence du compte utilisateur [redacted]

[redacted]

[redacted]

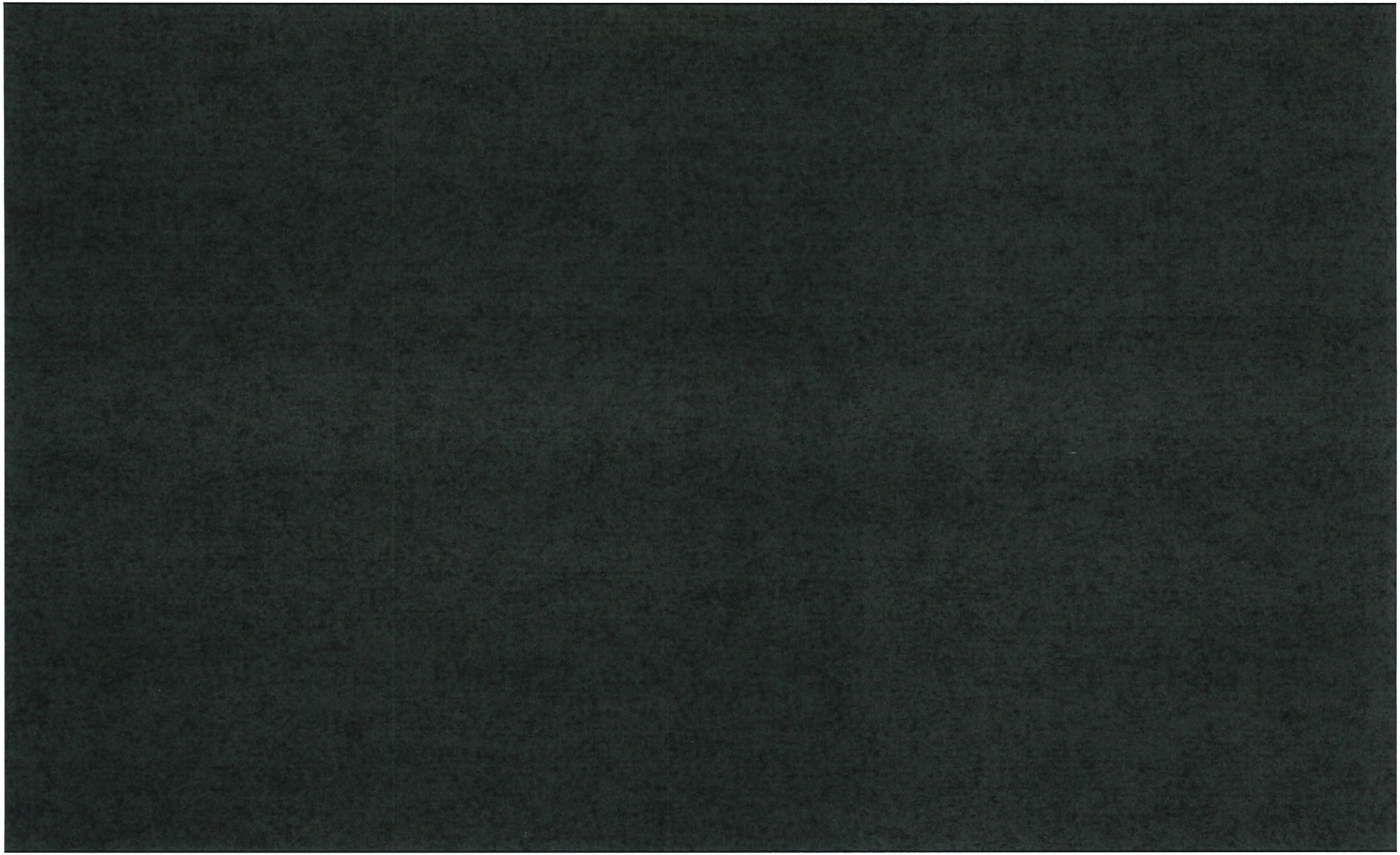
[redacted]

1.2.2. Quel est le cycle de vie des données ?

Les différentes étapes portant sur les données dans le cadre de la constitution de l'EDS EMC2 sont les suivantes :

- **Ciblage** - Au niveau de chaque centre, la population incluse est ciblée dans ses données hospitalières
- **Extraction** - Une fois la population ciblée, chaque centre extrait les données nécessaires (pour l'appariement et la constitution de l'EDS EMC2)
- **Préparation** - Au niveau de chaque centre, les données extraites sont mises en qualité.
- **Transfert & appariement** - L'ensemble des établissements partenaires transfèrent les données mentionnées au paragraphe 3.3 à la CNAM pour l'appariement direct.
- [redacted]

[redacted]



Traitement 1- [redacted]

- **Extraction des données :**

[redacted]

- **Préparation et mise en qualité :**

[redacted]

- **Mise en place d'un flux d'ingestion sécurisé pour le transfert de la copie des données de l'établissement de santé :**

[redacted]

- **Extraction SNDS et transfert de la CNAM vers le HDH :**

[redacted]

2 https://www.cnil.fr/sites/default/files/atoms/files/guide_pratique_circuits_nir_recherche_en_sante.pdf

[redacted]

- **Ingestion des données de santé :** [redacted]
- **Génération des identifiants propres à la plateforme technologique du HDH :** [redacted]
- **Validation des données :** [redacted]
- **Stockage des données de santé dans la plateforme technologique :** [redacted]
- **Persistance des données dans la plateforme technologique :** [redacted]

Traitement 2- [redacted]

Sur la plateforme technologique du HDH :

- **Préparation de la base finale :** [redacted]
- **Création d'un espace d'analyse :** [redacted]

- [redacted]
- **Copie des données dans l'espace de préparation :** [redacted]
- **Gestion des identifiants :** [redacted]
- **Préparation des données dans l'espace de préparation :** [redacted]
- **Exposition des données dans l'espace d'analyse :** [redacted]
- **Gestion des demandes d'imports et exports de données formulées par les utilisateurs :** [redacted]

[redacted]

Via un système-fils :

[redacted]

Traitement 3 - [redacted]

a) Sous-finalité 1 : [redacted]

- **Mise en place d'un catalogue de métadonnées** : production, mise à disposition et partage de métadonnées (informations sur la structure, statistiques descriptives, etc.), à travers un catalogue interne et/ou externe à la plateforme notamment pour permettre à des utilisateurs externes ayant les autorisations nécessaires de contribuer à la production de métadonnées.
- **Formatage et standardisation des données.**
- **Amélioration des données** : transformation des données pour améliorer la qualité métier (redressement, normalisation, création de nouvelles variables, etc.).

b) Sous-finalité 2 : [redacted]

- **Organisation d'hackathons, studyathon et data challenges** : organisation d'activités de découverte et de formations en permettant aux participants d'envoyer leurs algorithmes, de les lancer [redacted] sur la plateforme technologique et de partager les performances obtenues.
- **Réponses à des questions ponctuelles et création d'indicateurs** : réalisation d'analyses basiques pour répondre à des questions simples d'intérêt général ou pour créer des indicateurs simples pour lesquelles une démarche d'autorisation serait disproportionnée [redacted]. Création d'indicateurs : Sans qu'il n'y ait de sollicitation extérieure, les experts du SNDS au sein du HDH pourraient publier des statistiques régulièrement mises à jour à partir des données du SNDS.
- **Définition de méthodes de référence de valorisation des données** : Définition de méthodes de référence permettant de définir de nouvelles façons de valoriser les données de santé (e.g., utilisation de données de vie réelle dans le cadre d'études cliniques)

Traitement 4 : [redacted]

a) Sous-finalité 1 - [redacted]

- **Recueil des données des utilisateurs de la plateforme technologique** : [redacted]

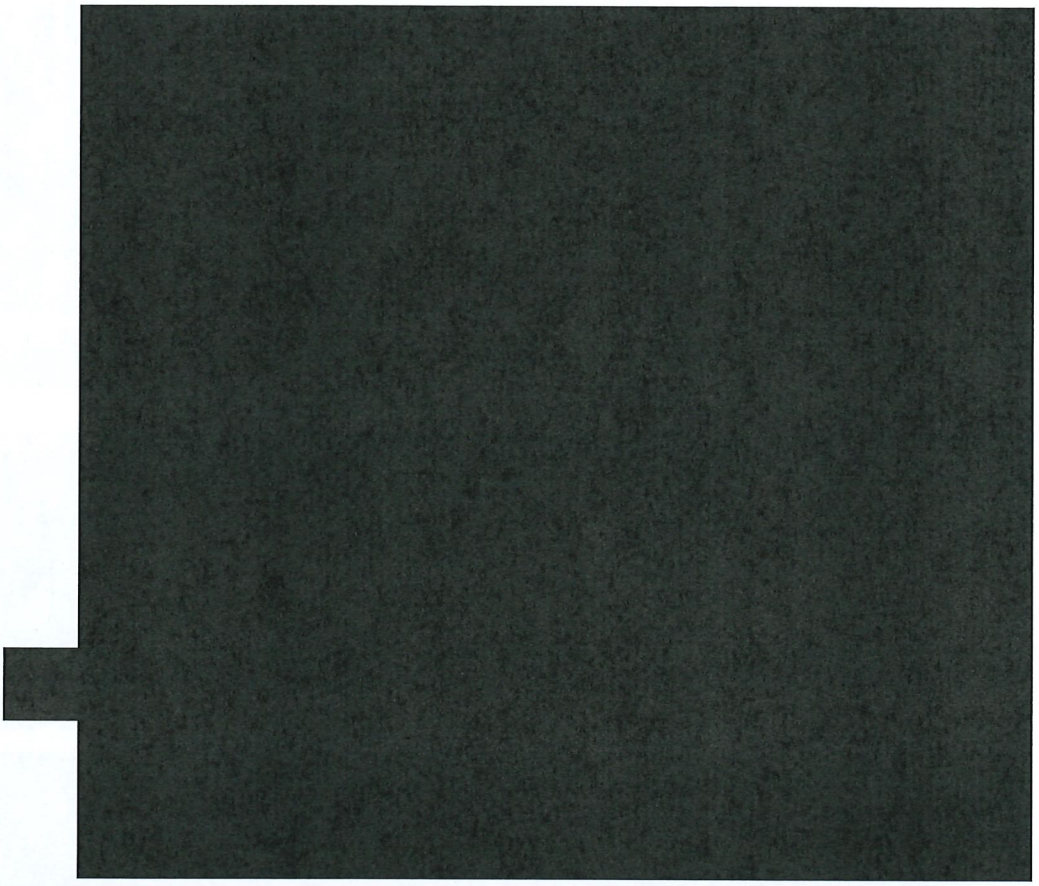
[redacted]

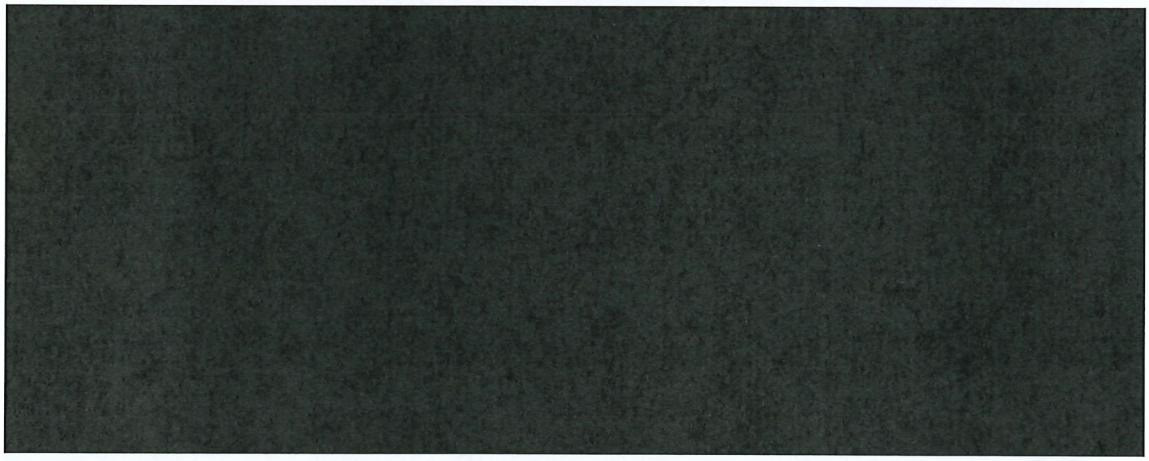
- **Modification des données des utilisateurs de la plateforme technologique :** [redacted]
- **Suppression des données des utilisateurs de la plateforme technologique :** [redacted]

[redacted]

1.2.3. Quels sont les supports de données ?

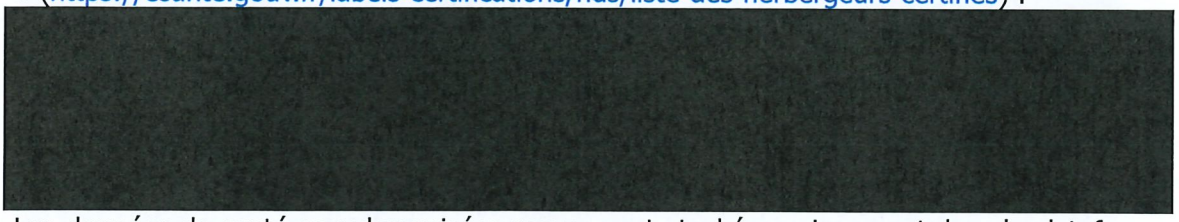
[redacted]



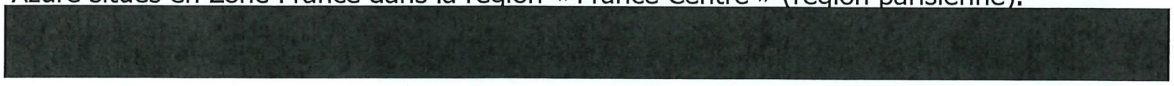


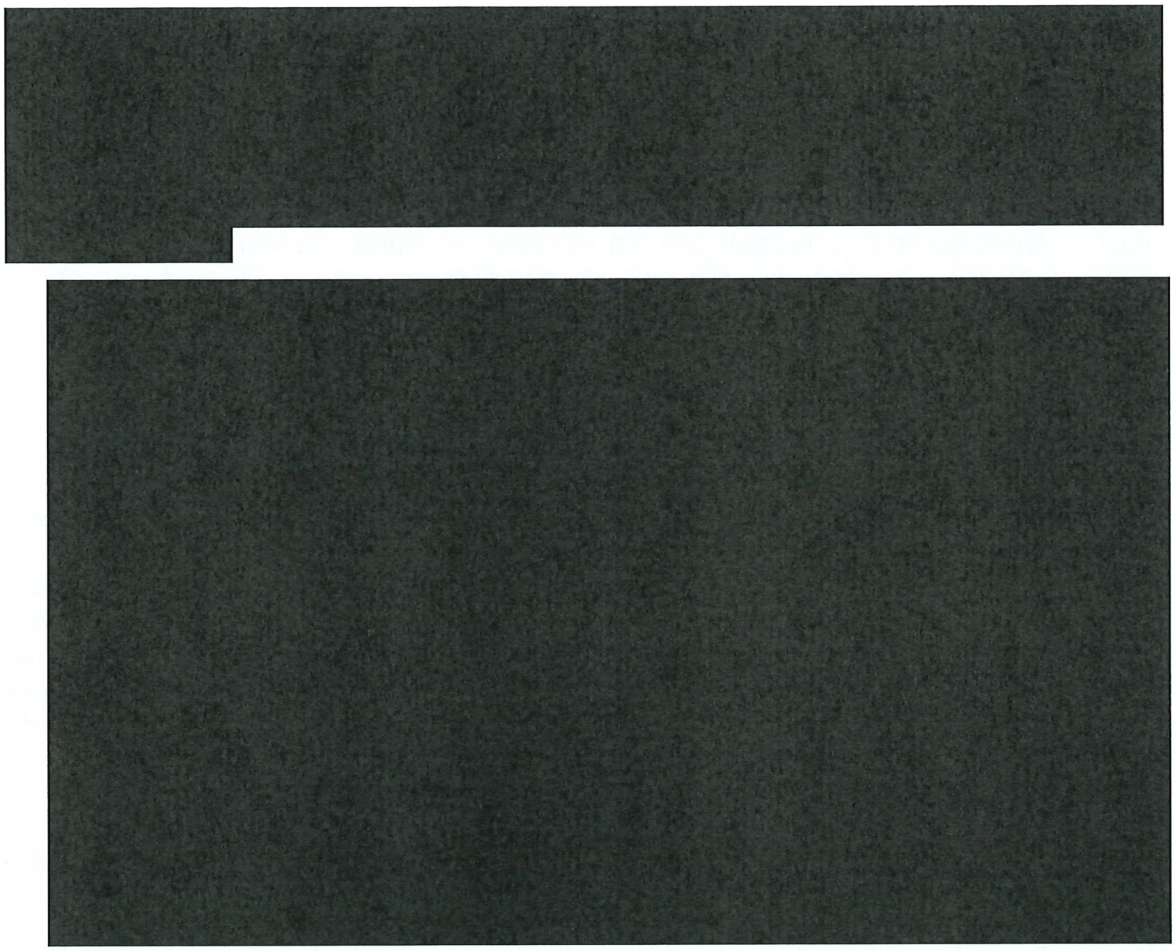
- **Plateforme technologique du HDH** :

La plateforme technologique repose sur les offres de Microsoft Azure, une solution d'hébergement dans le Cloud disposant de la certification « Hébergeur de données de santé » (https://esante.gouv.fr/labels-certifications/hds/liste-des-herbergeurs-certifies) :



Les données de santé pseudonymisées reçues sont stockées uniquement dans la plateforme technologique. Cette plateforme est hébergée dans les centres de données de Microsoft Azure situés en Zone France dans la région « France Centre » (région parisienne).





Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2. Principes fondamentaux

2.1. Mesures garantissant la proportionnalité et la nécessité du traitement

2.1.1. Finalités et fondements

2.1.1.1. Finalités poursuivies par l'EDS

Le traitement a pour objectif la mise en œuvre d'un entrepôt de données de santé en vue de la réutilisation des données qu'il contient. La finalité est identique à celle prévue par le référentiel sur les entrepôts de données de santé.

En aucun cas les données ne seront traitées à des fins de promotion des produits mentionnés au II de l'article L. 5311-1 du code de la santé publique en direction de professionnels de santé ou d'établissements de santé, à des fins d'exclusion de garanties des contrats d'assurance ou de modification de cotisations ou de primes d'assurance d'un individu ou d'un groupe d'individus présentant un même risque.

Le HDH a été créé par la loi n° 2019-774 du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé et ses missions sont définies à l'article L.1462-1 du code de la santé publique.

La constitution d'un entrepôt de données hospitalières multicentriques est un enjeu majeur pour le développement de la recherche "en vie réelle" en France. Au sein du HDH, l'entrepôt EMC2 regroupe des données cliniques et paracliniques liées à la prise en charge intra-hospitalière des patients, enrichies des données de la base principale du SNDS, ouvrant de nouvelles perspectives pour la recherche en santé.

La base principale du SNDS est un système de bases de données médico-administratives parmi les plus volumineuses et exhaustives du monde, liées aux remboursements des actes et des soins des bénéficiaires de l'ensemble des régimes d'assurance maladie obligatoire. Il présente un intérêt majeur pour la recherche et l'innovation en santé en France notamment dans les domaines épidémiologique, pharmaco-épidémiologique et médico-économique.
Néanmoins, l'absence de données cliniques ou de données sur les résultats des examens complémentaires au sein de la base principale du SNDS limite la portée de nombreuses études. L'entrepôt EMC2 vise à combler - au moins en partie - ce manque en appariant les données cliniques et paracliniques fournies par quatre établissements de santé français avec les données individuelles de la base principale du SNDS pour les patients concernés, complétées d'un échantillon de la base principale du SNDS afin de pouvoir mener des analyses comparatives en population générale.

La mise en œuvre de l'entrepôt EMC2 a pour finalité de permettre la réutilisation des données qu'il contient à des fins de recherche, d'étude et d'évaluation dans le domaine de la santé. Tout responsable de traitement réalisant les formalités adéquates et en premier lieu les établissements hospitaliers partenaires, les équipes de recherche internes à l'EMA pourront réaliser ces études. Il s'agira en particulier de la conduite d'études d'intérêt public de différentes natures telles que :

- l'observation et l'évaluation de la prise en charge des patients,
- la caractérisation des populations de patients,
- l'évaluation de l'utilisation / des pratiques, de l'efficacité et de la sécurité en vie réelle (hôpital et ville) des produits de santé (en particulier médicaments et dispositifs médicaux (DM) inscrits au remboursement ou en accès précoce),
- l'évaluation médico-économique (efficience et 'impact budgétaire) des produits de santé administrés (médicaments) ou utilisés (DM) en vie réelle,
- le développement et l'évaluation d'actions de prévention pour les pathologies d'intérêt,
- la conception et la validation d'outils d'aide à l'interprétation des signaux, au diagnostic ou à la prise en charge préventive ou curative.

La liste précédente des études ne se veut pas exhaustive : elle a été déterminée au regard des problématiques d'importance identifiées par les établissements hospitaliers partenaires ainsi que par l'EMA. L'entrepôt EMC2 sera ouvert à toute réutilisation des données dans une finalité d'intérêt public et pour lequel son périmètre est pertinent : contribuer à la recherche, aux études, à l'évaluation et à l'innovation dans les domaines de la santé et de la prise en charge médico-sociale.

Par rapport à des études équivalentes conduites sur la base principale du SNDS seule, l'entrepôt EMC2 permettra des analyses sur des sous-groupes de patients non identifiables à partir des seules données du PMSI mais repérables à partir des critères cliniques ou paracliniques transmis par les 4 établissements de santé. Par exemple, pour de nombreuses pathologies, le PMSI (via la caractérisation CIM-10 des diagnostics lors des séjours

hospitaliers) ne permet pas de différencier les patients selon la sévérité ou le stade de la maladie (insuffisance cardiaque, cancer, etc.). Les données cliniques ou les résultats d'examens fournis par les 4 établissements de santé de l'entrepôt EMC2 permettront d'identifier et de caractériser des sous-groupes de patients jusque là non identifiables et donc de réaliser des analyses stratifiées selon différents phénotypes ou stades de sévérité, et ainsi d'affiner les études.

2.1.1.2. Justification de l'intérêt public

La mise en œuvre de l'entrepôt EMC2 présente un caractère d'intérêt public pour plusieurs raisons.

En rassemblant des données cliniques chaînées avec celles de la base principale du SNDS, l'entrepôt EMC2 réunit des données de santé d'une grande richesse permettant d'appréhender les variétés des parcours de santé en disposant de données cliniques complétées par des données issues des remboursements de soins. Cette complémentarité des sources de données est propice au développement de recherches au bénéfice des patients, de la communauté scientifique et plus largement de la société en permettant l'amélioration de la qualité des prises en charge préventives et curatives sur le territoire. Le regroupement de ces données favorise en effet la mise en œuvre d'études permettant par exemple de mieux appréhender les parcours de soins, d'identifier des actions de prévention existantes et à mettre en place, d'évaluer la sécurité et l'efficacité des médicaments au sein de sous-groupes de patients particuliers, de favoriser le développement et la validation d'algorithmes afin d'améliorer l'établissement de diagnostics et la caractérisation d'une population d'intérêt. [redacted] des données présentes dans l'entrepôt facilitera en outre la conduite d'études fédérées à l'échelle européenne voire internationale et la réutilisation d'algorithmes développés sur d'autres sources de données au sein de l'entrepôt.

- **L'entrepôt EMC2 apporte un bénéfice important pour la société et la communauté scientifique**

Les autorités administratives (françaises ou européennes) régulant l'accès au marché et le remboursement des produits de santé, appuyées par des équipes de recherche spécialisées en pharmaco-épidémiologie, ont pour mission d'assurer le suivi de l'utilisation, la veille sur les effets indésirables, la mesure de l'efficacité pour les patients et pour le système de santé des produits de santé. La réalisation de ces études doit mobiliser des "données de vie réelle" afin de compléter les informations fournies par les industriels en phase de pré-commercialisation et issues des essais cliniques. L'entrepôt EMC2 permettra notamment de documenter ces sujets pour un panel important de médicaments et de DM commercialisés en France et utilisés auprès de différentes populations de patients (grossesse et périnatalité, pédiatrie, patients adultes tous âges et toutes pathologies, avec une forte représentation de l'onco-hématologie).

- **Les données contenues dans l'entrepôt EMC2 ne seront pas exploitées à des fins interdites**

Les données de l'entrepôt EMC2 ne seront exploitées ni à des fins de promotion des produits mentionnés au II de l'article L. 5311-1 du code de la santé publique en direction de professionnels de santé ou d'établissements de santé, ni à des fins d'exclusion de garanties des contrats d'assurance, ni de modification de cotisations ou de primes d'assurance d'un individu ou d'un groupe d'individus présentant un même risque.

En premier lieu, le HDH ne constitue pas l'entrepôt dans la perspective de répondre à des finalités interdites. En effet, en tant qu'acteur visant au titre de l'article L1461-1 du code de la santé publique à garantir un accès aisé et unifié, transparent et sécurisé aux données de santé pour améliorer la qualité des soins et l'accompagnement des patients, le HDH est investi d'une mission d'intérêt public et n'a pas vocation dans ce cadre à permettre la poursuite de finalités interdites.

En second lieu, dans le cadre de la réutilisation des données à des fins de recherche par des responsables de traitements tiers, le comité scientifique de l'entrepôt prendra acte de l'engagement de ces derniers, inscrit dans le dossier de recherche, à ne pas poursuivre ces finalités interdites, conformément au point 5.1 de l'arrêté du 22 mars 2017 relatif au référentiel de sécurité applicable au Système national des données de santé. Le principe de limitation des finalités avec une mise en garde contre le détournement des finalités sera abordé lors des sessions de sensibilisation des utilisateurs conformément au point 3.5 Sensibilisation de l'arrêté précité.

- **Des dispositions sont prises afin de garantir l'intégrité scientifique et la qualité des études et de prévenir le risque de produire des résultats biaisés**

Afin de garantir la qualité des études menées à partir des données de l'entrepôt, chacune d'entre elles fait l'objet d'un avis du comité scientifique et éthique de l'EDS. Ce dernier est chargé de se prononcer sur la pertinence et la qualité scientifique de chacun des projets qui lui sera soumis. Pour cela, il dispose d'expertises croisées à travers la représentation des fournisseurs de données, d'experts scientifiques et de représentants d'associations de patients. Par ailleurs, les données collectées dans l'EDS passent par un processus de standardisation et de mise en qualité permettant aux utilisateurs finaux de l'EDS de mener leurs études sur des bases de données fiables et standardisées, ce qui réduit le délai de préparation des données pour les utilisateurs finaux, et réduit d'autant le délai d'obtention des résultats tout en améliorant leur pertinence.

- **Des mesures sont prises afin d'assurer la transparence des études menées sur les données de l'entrepôt EMC2**

Afin que les citoyens puissent prendre connaissance des études réalisées sur les données de l'entrepôt, le HDH va déployer sur son site internet un portail de transparence relatif aux EDS, à l'image de celui existant pour les projets de recherche. L'entrepôt EMC2 y sera référencé et renverra vers les fiches du répertoire public du HDH contenant les informations de l'article 14 du RGPD propres à chaque projet de recherche réalisé à partir des données de cet entrepôt. Le répertoire public contiendra un filtre permettant d'identifier facilement les études réalisées sur les données de l'entrepôt EMC2.
Enfin, une information générale sur la mise en œuvre de l'EDS sera réalisée par le HDH via une campagne d'information publique.

En tout état de cause, le HDH propose de communiquer tous les trois ans à la CNIL un rapport sur le fonctionnement de l'entrepôt et sur les recherches réalisées à partir des données qu'il contient.

- **L'entrepôt EMC2 s'inscrit dans une démarche de science ouverte sur l'interopérabilité d'EDS**

Les scripts permettant de convertir les données du SNDS [redacted] sont ouverts à tous, et la constitution de l'entrepôt permettra d'améliorer la première version mise à disposition par le HDH, en prenant en compte les spécificités de chaînage entre les données hospitalières et la base principale du SNDS. D'autres projets pourront donc s'appuyer

librement sur ces premiers travaux. De plus, la constitution d'un entrepôt répondant aux spécifications du CDM permettra également l'exécution locale de requêtes standardisées développées dans le cadre de collaboration de recherche. Cette possibilité d'utiliser des scripts standardisés pour la conduite d'analyses selon un modèle fédéré permettra à l'entrepôt EMC2 de contribuer aux efforts de recherche multi-bases de données pour la production de preuves robustes au niveau national et international.

2.1.1.3. Base juridique et exception à l'interdiction de traiter des données de santé

Dans la mesure où les traitements de données de santé à caractère personnel sont nécessaires à l'exécution des missions d'intérêt public du HDH, le fondement juridique du traitement des données au sens du RGPD est donc l'article 6-1-e (« *le traitement est nécessaire à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement* ») et, s'agissant particulièrement des données de santé, l'exception de l'article 9-2-i est mobilisée (« le traitement est nécessaire pour des motifs d'intérêt public dans le domaine de la santé publique, tels que la protection contre les menaces transfrontalières graves pesant sur la santé, ou aux fins de garantir des normes élevées de qualité et de sécurité des soins de santé et des médicaments ou des dispositifs médicaux, sur la base du droit de l'Union ou du droit de l'État membre qui prévoit des mesures appropriées et spécifiques pour la sauvegarde des droits et libertés de la personne concernée, notamment le secret professionnel »).

Concernant les traitements de données relatives aux utilisateurs de la plateforme technologique, le fondement juridique est l'article 6-1-c du RGPD (« *le traitement est nécessaire au respect d'une obligation légale à laquelle le responsable du traitement est soumis* »).

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.1.2. Les données collectées sont-elles adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées (minimisation des données) ?

Traitement 1- [redacted]

La cohorte intégrée dans l'EDS EMC2 est constituée de patients ayant été hospitalisés en MCO (Médecine-Chirurgie-Obstétrique) dans l'un des quatre centres hospitaliers participants depuis 2022. Chaque centre sélectionne un sous-groupe de patients parmi les patients hospitalisés dans son établissement sur la période d'inclusion, en se basant :

- sur sa capacité à informer le patient du traitement de ses données
- Sur sa capacité à récupérer les différentes données patients
- sur les domaines médicaux d'intérêt identifiés pour l'EDS EMC2 repérés, par exemple, grâce aux codes médicaux associés aux séjours ou autres codes traceurs. Parmi les domaines médicaux d'intérêt figurent notamment les spécialités nutrition/métabolisme/cardiologie, neurologie, oncologie, gériatrie, pédiatrie, infectiologie, hépatologie, appareil locomoteur, immunologie et allergologie. Les critères d'extraction utilisé par chaque établissement hospitalier pour alimenter l'EDS EMC2 seront clairement documentés et mis à disposition des utilisateurs

Chaque centre fournit à l'EDS EMC2 de nouveaux patients chaque année ; il s'agit d'une cohorte ouverte. Les contributions en nombre de patients apportés par chaque centre sera variable en fonction des années et dépendra notamment de la maturité des établissements hospitaliers et des pathologies d'intérêt ciblées dans chaque centre. Le nombre de patients intégrés dans l'EDS est estimé entre 300 000 et 500 000 par année à partir des chiffres fournis par les centres participants sur le nombre de patients par spécialité médicale sur une période de 12 mois calendaires.

Afin de pouvoir réaliser des études épidémiologiques sur des effets embryo/foetotoxiques liés à certains médicaments, il est également envisagé d'identifier les enfants nés depuis 2015 des femmes patientes des centres incluses dans la cohorte de l'EDS EMC2. Ces enfants seront identifiés dans la base principale du SNDS [redacted] présent dans le PMSI et leurs données SNDS seront intégrées dans l'EDS EMC2. L'année 2015 a été choisie comme date de début de prise en compte des naissances pour coller à la date de disponibilité des premières données SNDS dans l'EDS EMC2.

Dans le but de faciliter la réalisation d'études épidémiologiques (type cas-témoin), il est également envisagé d'intégrer dans l'EDS EMC2 des données de la base principale du SNDS pour une population témoin. Cette dernière sera tirée au sort dans ces mêmes données sur la base des nouveaux patients inclus suivant un ratio 1:3 (1 patient pour 3 témoins) apparié sur le sexe, l'année de naissance et le département de résidence. Afin de ne pas trop agrandir la population de l'EDS EMC2, un ratio de trois témoins a été retenu car généralement la puissance statistique des analyses cas-témoin n'est pas améliorée au-delà[3].

Les données extraites et transmises sur la plateforme du HDH ne comportent pas de données directement identifiantes, et contiennent uniquement des données structurées pseudonymisées. [redacted] à partir de la date de collecte dans le cadre du soin.

Les données des centres participants permettent notamment d'obtenir des informations sur les résultats d'examens biologiques et cliniques, les facteurs de risque du patient, les indications de prescription qui ne sont pas présentes dans les données de la base principale du SNDS.

Les données extraites de la base principale du SNDS relatives à la population incluse couvrent la période [redacted] la date de mise à disposition des données pour l'ensemble de la population présente dans l'EDS afin d'avoir un suivi rétrospectif et prospectif au long cours des patients dans l'EDS.

Les données de la base principale du SNDS sont nécessaires pour répondre aux objectifs de l'EDS EMC2 sur l'étude du parcours de soins, notamment en apportant des données longitudinales exhaustives sur les soins de ville, les hospitalisations dans d'autres

[3] Kang et al. (2009). The Effect of Increasing Control-to-case Ratio on Statistical Power in a Simulated Case-control SNP Association Study. Genomics & Informatics Vol. 7(3) 148-151, September 2009

établissements de santé ou encore d'autres informations d'intérêt (causes médicales de décès, vaccination Covid, etc.).

Les dates exactes de soins sont indispensables pour reconstruire et évaluer de manière précise le parcours de soins des patients. La date de décès permet de réaliser des études de mortalité et des analyses de survie. Il sera également intéressant de disposer du mois de naissance (en plus de l'année) afin de calculer l'âge précis, qui peut avoir son importance, notamment pour des études incluant des enfants en bas-âge. Enfin, il sera nécessaire de disposer de la commune de résidence en vue de l'utilisation l'indice de défavorisation communal pour prendre en compte le niveau socio-économique au niveau de certaines études.

La convention bipartite signée entre le HDH et chaque établissement de santé doit notamment préciser le périmètre, la fréquence de mise à jour et la sécurisation du transfert des données. La qualité des données relève de la responsabilité des producteurs de données tels que la CNAM et les établissements de santé.

Concernant les données provenant des établissements de santé, il s'agit de données médicales et médico-administratives déjà collectées. Elles peuvent comporter des erreurs dues à la nature du processus de remontées d'informations qui seront toujours présentes dans la base de données issue de l'appariement. Toutefois, les erreurs présentes dans cette base sont sans impact sur les droits et libertés des personnes.

Traitement 2 - [redacted]

Quel que soit le projet, les formalités applicables doivent être accomplies et la pertinence des données est vérifiée à cette occasion. La plateforme technologique n'héberge que les données ainsi validées lors d'un processus d'accès réglementaire.
De surcroît, une équipe projet n'a accès qu'aux données strictement nécessaires à son projet et elle n'a en aucun cas accès aux données stockées dans la plateforme technologique pour d'autres projets. Ceci est assuré techniquement par l'architecture de la plateforme technologique [redacted]

En outre, s'agissant des appariements de données réalisés au sein de la plateforme technologique, l'architecture de cette dernière garantit que les utilisateurs projet n'ont accès qu'aux données appariées strictement nécessaires à la réalisation de l'étude. Pendant la phase d'appariement préalable, la minimisation des données est assurée :

- en cas d'appariement direct (ou déterministe) : le lien est établi entre plusieurs jeux de données concernant des personnes grâce à un identifiant commun. Seules les données rattachées à cet identifiant sont donc transmises sur la plateforme technologique.
- en cas d'appariement indirect (ou probabiliste) : les personnes sont sélectionnées en fonction d'informations définies sans avoir d'identifiant en clair. Ces personnes devront correspondre de manière probabiliste aux informations souhaitées pour que la correspondance soit la plus optimale possible. Cette correspondance est établie généralement en plusieurs étapes, par tâtonnement, en jouant sur les différentes variables. La sélection est rendue plus facile lorsque le nombre de variables transmises est important (avec des variables le plus discriminantes possibles sur la personne). Le périmètre des variables utilisables pour la phase d'appariement est déterminé projet par projet dans l'autorisation délivrée par la CNIL.

Enfin, il est à noter que, sous réserve que leurs traitements soient encadrés par une autorisation de la CNIL ou une méthodologie de référence, les établissements de santé sont susceptibles d'accéder eux-mêmes aux données de l'EDS sur la plateforme technologique à

des fins de mise en qualité et préparation des données. Le cas échéant, ils devront respecter les CGU et ils disposeront de leur propre espace de préparation réservé aux opérations de mise en qualité des données.

Traitement 3 - [redacted]

a) Sous-finalité 1 : [redacted]

Les données sont issues de l'EDS. Le périmètre des données nécessaires est déterminé projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

b) Sous-finalité 2 : [redacted]

Les données sont issues de l'EDS. Le périmètre des données nécessaires est déterminé projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

c) Sous-finalité 3 : [redacted]

Les données sont issues de l'EDS. Le périmètre des données nécessaires est déterminé projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

Traitement 4 - [redacted]

a) Sous-finalité 1 : [redacted]

En vertu de sa mission de mise à disposition des données prévue par l'article L. 1462-1 du code de la santé publique et afin de respecter l'obligation légale posée par l'article L. 1461-1 du même code selon laquelle l'accès aux données du SNDS s'effectue dans des conditions assurant « *la confidentialité et l'intégrité des données et la traçabilité des accès et des autres traitements* », le HDH doit traiter les données à caractère personnel relatives aux utilisateurs. Les données collectées pour la gestion des comptes sont circonscrites au strict nécessaire tel que décrit dans la réponse à la question 1.2.1 *Quelles sont les données traitées?*.

[redacted]

Acceptable / Améliorable / À corriger	Commentaires d'évaluation	Mesures correctives

Acceptable	- La pertinence des données est déterminée en amont par le responsable de traitement du projet et vérifiée lors de la procédure d'accès aux données sous le contrôle de la CNIL. - La grande diversité et l'important volume de traces collectées sont justifiés par les finalités poursuivies par l'EDS, telles que définies en partie 2.1.1. Finalités et fondements.	

2.1.3. Les données sont-elles exactes et tenues à jour ?

Traitement 1 - [redacted]

La convention bipartite signée entre le HDH et chaque établissement de santé doit notamment préciser le périmètre, la fréquence de mise à jour et la sécurisation du transfert des données. La qualité des données relève de la responsabilité des producteurs de données tels que la Cnam et les établissements de santé.

Concernant les données provenant des établissements de santé, il s'agit de données médicales et médico-administratives déjà collectées. Elles peuvent comporter des erreurs dues à la nature du processus de remontées d'informations qui seront toujours présentes dans la base de données issue de l'appariement. Toutefois, les erreurs présentes dans cette base sont sans impact sur les droits et libertés des personnes.

Traitement 2 - [redacted]

S'agissant des mêmes données de santé que pour le traitement 1, leur qualité et leur fréquence de mise à jour relèvent du même mécanisme qui s'appuie sur la responsabilité des établissements de santé.

Traitement 3 - [redacted]

a) Sous-finalité 1 : [redacted]

Il s'agit des données de l'EDS visées au traitement 1 qui sont mises à jour selon les modalités définies avec les établissements de santé.

b) Sous-finalité 2 : [redacted]

Il s'agit des données de l'EDS visées au traitement 1 qui sont mises à jour selon les modalités définies avec les établissements de santé.

c) Sous-finalité 3: [redacted]

Il s'agit des données de l'EDS visées au traitement 1 qui sont mises à jour selon les modalités définies avec les établissements de santé.

Traitement 4 - [redacted]

a) Sous-finalité 1 : [redacted]

[redacted]

Acceptable / Améliorable / À corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.1.4. Quelle est la durée de conservation des données ?

Traitement 1 - [redacted]

Les données cliniques contenues dans l'EDS seront conservées pour une durée de 20 ans à compter de leur collecte dans le cadre du soin et les données de la base principale du SNDS seront conservées pour une durée de 20 ans à compter de leur versement dans l'EDS, ce qui permettra de répondre à la nécessité d'avoir un suivi longitudinal important pour la réalisation d'études pharmaco-épidémiologiques et de pharmacovigilance d'intérêt. Une fois ces délais expirés, les données seront supprimées de l'EDS pour laisser place à la réception d'un nouveau jeu de données, à l'occasion de la mise à jour annuelle de l'EDS.

Concernant les projets mobilisant les données de l'EDS, les données sont conservées au sein de l'espace de stockage pendant la durée du projet qui est fondée sur le temps nécessaire à la conduite de celui-ci. Cette durée de conservation diffère donc d'un projet à un autre et est déterminée par le responsable de traitement du projet. [redacted]

[redacted]

Traitement 2 - [redacted]

Pendant la phase de préparation et d'exposition, les données sont conservées sur la plateforme technologique en deux endroits distincts : d'abord au sein de l'espace de préparation [redacted], ensuite au sein de l'espace d'analyse pour les utilisateurs projet.

- Au sein de l'espace de préparation, la durée de conservation des données sera conditionnée par le besoin du projet en termes de mise à jour des données dans son espace d'analyse au cours de son cycle de vie :
  - Pour les projets ne nécessitant pas de mise à jour des données pendant leur déroulé (une livraison initiale et unique dans l'espace d'analyse suffit pour mener à bien l'étude) : [redacted], période pendant laquelle le porteur de projet peut solliciter des données corrigées et/ou complémentaires si cette demande est conforme aux besoins définis.

[redacted]

- Au sein de l'espace d'analyse, la durée de conservation est fondée sur le temps nécessaire à la conduite du projet. Cette durée de conservation diffère donc d'un projet à un autre et est déterminée par le responsable de traitement du projet sous le contrôle de la CNIL. Les identifiants projet générés par le HDH sont également conservés pour la durée du projet. [redacted]

Traitement 3 - [redacted]

a) Sous-finalité 1 : [redacted]

La durée de conservation des données nécessaires est déterminée projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

b) Sous-finalité 2 : [redacted]

La durée de conservation des données nécessaires est déterminée projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

c) Sous-finalité 3 : [redacted]

La durée de conservation des données nécessaires est déterminée projet par projet en fonction des axes de travail du HDH (Direction des données et Guichet).

Traitement 4 - [redacted]

a) Sous-finalité 1 : [redacted]

[redacted]

Acceptable / Améliorable / À	Commentaires d'évaluation	Mesures correctives

corriger		
Acceptable		

2.2. Mesures protectrices des droits des personnes concernées

2.2.1. Comment les personnes concernées sont-elles informées à propos du traitement ?

2.2.1.1. Modalités générales d'information des personnes concernées par les données cliniques appariées à la base principale du SNDS

- **Information des patients adultes et mineurs et titulaires de l'autorité parentale des patients mineurs :**

Les patients adultes et mineurs concernés par les données de l'EDS seront informés de manière individuelle et collective que les données collectées lors de leur prise en charge sont versées dans l'EDS et appariées avec les données de la base principale du SNDS. L'information s'opérera préalablement au versement des données dans l'EDS, a minima un mois avant. Dans l'hypothèse où l'information individuelle d'une personne dont les données sont pertinentes pour l'EDS n'est pas possible (par exemple, l'établissement de santé n'a plus ses coordonnées), alors l'établissement ne les transmettra pas au HDH.

L'information est écrite et fournie aux établissements de santé par le HDH. Elle comporte toutes les mentions d'information de l'article 14 du RGPD. La note d'information délivrée aux patients adultes, testée auprès d'une association de patients, est disponible en annexe 4.1 de l'argumentaire.

Pour le cas particulier des mineurs dont des données pourraient être versées dans l'EDS, ces derniers recevront une note d'information adaptée et testée auprès d'un groupe d'enfants, également fournie par le HDH aux établissements de santé, disponible en annexe 4.2 de l'argumentaire. Les titulaires de l'autorité parentale des patients mineurs seront également informés selon la note d'information disponible en annexe 4.3 de l'argumentaire. Seules les données des patients mineurs dont les titulaires de l'autorité parentale pourront être informés seront versées dans l'EDS.
Pour le cas particulier des majeurs faisant l'objet d'une mesure de protection, l'information individuelle de ces derniers ainsi que de leurs tuteurs impliquant une complexité supplémentaire, leurs données ne seront pas incluses dans l'EDS.

Chaque établissement de santé informera individuellement l'ensemble des personnes (patients adultes, patients mineurs et titulaires de l'autorité parentale des patients mineurs) selon des modalités propres à chaque établissement qui sont décrites au point 2.2.1.2. et validées par le HDH.

Une information collective relative à la mise en œuvre de l'EDS EMC2 sera assurée par ailleurs par le HDH via le portail de transparence mentionné au point 2.2.1.3. et via une campagne d'information générale sur la mise en œuvre de l'EDS.

L'information pour chaque projet de recherche sera enfin assurée via le portail de transparence précité auquel renvoie la note d'information individuelle relative à l'EDS. Ce

dernier contiendra les liens vers les fiches projets du répertoire public du HDH renseignées par les responsables de traitement des projets.

- **Information des professionnels de santé :**

Chaque établissement de santé informera individuellement les professionnels de santé exerçant au sein de l'établissement pendant ou postérieurement à la mise en œuvre de l'EDS.

Cette information s'opérera préalablement au versement des données dans l'EDS (a minima un mois avant), au moyen d'une note d'information comportant tous les éléments de l'article 14 du RGPD fournie par le HDH et remise selon des modalités propres à chaque établissement qui sont décrites au point 2.2.1.2. et validées par le HDH. La note d'information délivrée aux professionnels est disponible en annexe 4.4 de l'argumentaire.

> 2.2.1.2. Modalités de délivrance de l'information des personnes concernés par les données cliniques appariées à la base principale du SNDS propres à chaque établissement de santé

- **Délivrance de l'information par les HCL :**
  - **Pour les patients mineurs et adultes**

Les nouveaux patients et les patients en cours de suivi seront informés individuellement par la remise en main propre de la note d'information au moment de leur visite dans l'établissement ou par courrier.
Les patients n'étant plus suivis ou suivis préalablement à la constitution de l'EDS et ayant une adresse électronique renseignée seront informés par courriel avec accusé de réception.

  - **Pour les titulaires de l'autorité parentale des patients mineurs**

Les titulaires de l'autorité parentale des nouveaux patients mineurs et patients mineurs en cours de suivi seront informés par la remise en main propre de la note d'information au moment de leur visite dans l'établissement. Si l'un ou les deux titulaires n'est pas présent, il sera informé par courrier ou par courriel avec accusé de réception.
Les titulaires de l'autorité parentale des patients mineurs n'étant plus suivis seront informés par courriel avec accusé de réception sous réserve que leurs coordonnées soient disponibles. A défaut, les données du patient mineur concerné ne seront pas versées dans l'EDS.

  - **Pour les professionnels de santé**

Le personnel salarié des HCL sera informé individuellement par un courrier postal joint au bulletin de paie ou au contrat d'embauche. Une information générale ainsi qu'une information en Commission Médicale d'Etablissement (CME) seront réalisées et diffusées à tous les professionnels via des canaux de communication internes : site Intranet, lettres bimensuelles d'information par courriel (Flash HCL) et Revue interne Tonic.

- **Délivrance de l'information par le CHRU de Nancy :**
  - **Pour les patients mineurs et adultes**

L'ensemble des patients sera informé individuellement par une note d'information intégrée dans le livret d'accueil remis en main propre au patient hospitalisé. Les patients hospitalisés

entre le 1er janvier 2023 et la date d'intégration de l'information dans le livret d'accueil recevront une note d'information individuelle par courrier.

- **Pour les titulaires de l'autorité parentale des patients mineurs**

Les titulaires de l'autorité parentale des nouveaux patients mineurs et patients mineurs en cours de suivi seront informés par une note d'information intégrée dans le livret d'accueil remis en main propre au patient mineur hospitalisé. Si l'un ou les deux titulaires n'est pas présent, il sera informé par courrier.
Les titulaires de l'autorité parentale des patients mineurs hospitalisés entre le 1er janvier 2023 et la date d'intégration de l'information dans le livret d'accueil seront informés selon les modalités les plus appropriées (courriel ou à défaut par courrier postal).

- **Pour les professionnels de santé**

L'information sera diffusée en commission médicale d'établissement, sur l'intranet de celui-ci et à l'aide d'affiches dans les lieux de repos des personnels. La note d'information individuelle sera par ailleurs jointe à la fiche de paie des professionnels.

- **Délivrance de l'information par le CLB :**

  - **Pour les patients mineurs et adultes**

L'ensemble des patients sera informé individuellement par une note d'information intégrée au portail de transparence Unicancer/Mesdonnées, à l'image du dispositif d'information prévu dans la MR004. En effet, les patients de CLB dont les données seront versées dans l'EDS ont reçu une note d'information individuelle au moment de la collecte des données lors de la prise en charge au CLB détaillant :

- la possible réutilisation de ces données à des fins de recherche
- et l'existence du portail de transparence Unicancer/Mesdonnées auxquels les patients peuvent se reporter pour consulter les recherches réalisées sur leurs données.

Ainsi, le portail Unicancer/Mes données contiendra un encart dédié à l'EDS EMC2 avec un renvoi vers le portail de transparence EMC2 pour connaître la liste des études.

  - **Pour les titulaires de l'autorité parentale des patients mineurs**

Les titulaires de l'autorité parentale des patients mineurs seront informés par une note d'information intégrée au portail de transparence Unicancer/Mesdonnées.

  - **Pour les professionnels**

L'information sera diffusée en commission ou en conférence médicale d'établissement, sur l'intranet de celui-ci et à l'aide d'affiches dans les lieux de repos des personnels. La note d'information individuelle sera par ailleurs jointe à la fiche de paie des professionnels.

- **Délivrance de l'information par la FHSJ :**

  - **Pour les patients mineurs et adultes**

L'ensemble des patients sera informé individuellement par une note d'information transmise par courriel. Les nouveaux patients seront également informés par la remise d'une note d'information en main propre à l'accueil par le clinicien.

  - **Pour les titulaires de l'autorité parentale des patients mineurs**

Les titulaires de l'autorité parentale des patients mineurs seront informés par une note d'information transmise par courriel. Les titulaires de l'autorité parentale des nouveaux patients mineurs seront également informés par la remise d'une note d'information en main propre à l'accueil par le clinicien ou par courriel.

- **Pour les professionnels de santé**

L'information des professionnels sera assurée en commission médicale d'établissement et par la newsletter de l' établissement. La note d'information sera également déposée sur le coffre-fort numérique des professionnels de santé, qui seront notifiés par mail lors de l'ajout du document.

2.2.1.3. Modalités d'information des personnes concernées par les données de la base principale du SNDS seule (population témoin)

Le HDH mettra en œuvre un portail de transparence sur les EDS dynamique et accessible sur son site web. Ce portail référencera l'EDS EMC2 et contiendra les notes d'information collective conformes à l'article 14 du RGPD à destination des patients mineurs et adultes, des titulaires de l'autorité parentale et des professionnels de santé concernés par les données de la base principale du SNDS seule (population témoin). Ces notes d'information sont disponibles en annexes 4.5, 4.6, 4.7 et 4.8 de l'argumentaire.
Par ailleurs, le HDH assurera une campagne d'information générale sur la mise en œuvre de l'EDS.
Enfin, le portail de transparence précité contiendra les liens vers les fiches du répertoire public du HDH contenant les informations de l'article 14 du RGPD propres à chaque projet de recherche portant sur les données de l'EDS. Ainsi, les personnes concernées par ces recherches en seront informées, préalablement à leur mise en œuvre quels qu'en soient les responsables de traitement. Par ailleurs, le répertoire public contiendra un filtre permettant aux personnes concernées de retrouver facilement les études qui les concernent menées sur l'entrepôt EMC2, lorsqu'elles ne consulteront pas directement le portail de transparence.
Enfin, le HDH assurera une campagne d'information générale sur la mise en œuvre de l'EDS.

2.2.1.4. Modalités d'information des utilisateurs de la plateforme

Enfin, concernant les fonctionnalités de gestion des comptes utilisateurs et de traçabilité de leurs activités, les utilisateurs sont informés du traitement de leurs données lors du circuit d'arrivée sur la plateforme technologique, notamment grâce à la signature des conditions générales d'utilisation. Ces CGU sont envoyées [redacted] aux utilisateurs et détaillent les modalités d'exercice des droits relatifs à leurs données à caractère personnel.

Acceptable / Améliorable / À corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.2.2. Si applicable, comment le consentement des personnes concernées est-il obtenu ?

Ce point n'est pas applicable à l'EDS :

- Le consentement n'étant pas la base juridique retenue pour l'EDS, ce point n'est pas

applicable. Concernant les projets réutilisant les données de l'EDS, les traitements à des fins de recherche dans le domaine de la santé ne sont pas soumis à consentement préalable.

- Concernant les données collectées et utilisées pour la gestion des comptes des utilisateurs : le traitement est fondé sur le respect d'une obligation légale et les données sont recueillies directement auprès des utilisateurs.
- Concernant les traces de l'activité des comptes utilisateurs : la collecte des traces pour répondre au besoin de la sécurité de la plateforme répond à une obligation légale du HDH.

Acceptable / Améliorable / À corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.2.3. Comment les personnes concernées peuvent-elles exercer leurs droits d'accès et droit à la portabilité ?

Droit à la portabilité :

Les traitements de données mis en œuvre dans le cadre de l'EDS n'étant fondés ni sur le consentement de la personne ni sur l'exécution d'un contrat, les conditions d'applicabilité du droit à la portabilité prévu par l'article 20 du RGPD ne sont pas réunies.

Droit d'accès :

- Arrivée de la demande

Le canal d'entrée principal des demandes d'exercice de droit est le point de contact du DPO du HDH : par voie postale à l'adresse "À l'attention du Délégué à la protection des données, 9 rue Georges Pitard, 75015 Paris", par voie électronique à l'adresse dpd@health-data-hub.fr. Le DPO et son équipe ont accès à cette boîte aux lettres électronique. La personne concernée aura également la possibilité d'exercer ses droits via le formulaire du concentrateur du HDH, qui fera le lien avec l'EDS.

La boîte aux lettres électronique du DPO du HDH envoie un accusé de réception automatique, précisant à la personne concernée que sa demande a bien été prise en compte et qu'elle recevra une réponse sous un mois maximum.

Le DPO du HDH inscrit la demande au Tableau de suivi des demandes de droit et informe de cette demande l'établissement de santé concerné.

La personne concernée a également la possibilité d'exercer ses droits auprès de l'établissement de santé qui collecte ses données. Dans ce cas, l'établissement de santé :

- Mène les actions techniques nécessaires sur les données sources ;
- Informe le HDH dès réception de la demande, afin qu'il puisse mener les actions techniques sur les données de l'EDS, le cas échéant.

Les coordonnées du DPO du HDH et du DPO de l'établissement de santé concerné seront fournies au sein de la note d'information relative à l'EDS.

En cas de demande introduite à la fois auprès du HDH et d'un établissement de santé, le HDH se coordonnera avec l'établissement concerné en vue de déterminer les modalités de traitement de la demande et de réponse à la personne.

- Vérification de l'identité de la personne

En cas de doute raisonnable sur l'identité du demandeur, le DPO du HDH répondra manuellement en demandant un titre d'identité. Cette action suspend les délais prévus à l'article 12 du RGPD.

Si le demandeur ne répond pas à la demande du DPO du HDH sous 3 mois, la demande sera clôturée. La personne concernée sera informée par courriel de la clôture de sa demande et des motifs de cette clôture.

- Actions techniques

L'EDS étant chaîné à la base principale du SNDS, le circuit de pseudonymisation de la base principale du SNDS permettra de retrouver les données associées aux personnes qui exercent leur droit d'accès. En effet, dans la mesure où les identifiants utilisés dans la base principale du SNDS sont obtenus par un procédé cryptographique irréversible appliqué au NIR, le lien entre les données de la base principale du SNDS et la personne qui exerce ses droits pourra être fait en appliquant ces opérations de pseudonymisation du NIR, selon le circuit de pseudonymisation habituel qui fait intervenir la CNAM.

Dès lors que la personne fournira les informations nécessaires pour entrer dans le circuit de pseudonymisation, la CNAM sera en mesure de communiquer au HDH l'identifiant associé aux données de la copie de la base principale transmises au HDH. Ainsi, le HDH sera capable de vérifier que des données relatives à cette personne sont présentes ou non sur la plateforme technologique.
Si la personne concernée n'est pas retrouvée, le DPO du HDH répond négativement à la demande, sous un mois à compter de sa réception.

Si la personne concernée est retrouvée dans l'EDS :

- Si la demande porte sur la confirmation que les données sont ou ne sont pas traitées, le DPO du HDH indiquera à la personne que ses données sont ou ne sont pas présentes dans l'EDS.
- Si la demande porte sur l'obtention d'informations sur le traitement de données, le DPO du HDH fournira à la personne les informations demandées, en s'appuyant sur la lettre d'information délivrée par les ES et en détaillant, dans la mesure du possible, davantage les informations.
- Si la demande vise à obtenir l'accès aux données ou une copie des données :

[redacted]

[redacted]Lorsque le droit d'accès s'exerce par voie postale, nous envisageons que les données soient envoyées par courrier avec accusé de

réception afin de garantir que seul le bon destinataire accède au pli. Pour limiter l'accès aux données au sein du HDH, les données seraient imprimées par l'équipe DPO du HDH et un courrier type accompagnerait l'extraction des données. Les risques sur la confidentialité des données présentées par la voie postale seront exposés aux personnes et un envoi par voie électronique privilégié dans la mesure du possible. A défaut, une remise en main propre pourrait être proposée afin de limiter la circulation des données, selon les modalités suivantes :

- Remise en main propre dans les locaux du HDH, si la demande a été initialement formulée auprès du HDH ;
- Remise en main propre dans les locaux de l'établissement de santé qui suit la personne concernée, si la demande a été initialement formulée auprès de cet établissement de santé.

Concernant le droit d'accès des utilisateurs de la plateforme technologique à leurs données personnelles, le Délégué à la Protection des données du HDH est le point de contact des utilisateurs qui souhaitent l'exercer. Les coordonnées du DPD sont communiquées aux utilisateurs avant leur arrivée sur la plateforme technologique et sont rappelées dans les CGU de la plateforme technologique, qu'ils signent en amont de leur accès à la plateforme.

- Réponse à la demande

Le DPO du HDH répond à la personne concernée sous un mois, par courrier électronique ou, le cas échéant, par courrier postal.

Si la demande n'est pas recevable, le DPO du HDH répond négativement à la personne concernée, en lui exposant les motifs du refus.

Si la demande est recevable, le DPO du HDH répond positivement à la personne concernée, en lui expliquant les actions prises pour faire droit à sa demande.

S'il juge la demande complexe, le DPO informe sous un mois la personne concernée de l'allongement du délai de réponse et des motifs associés. Le DPO du HDH répondra à la personne sous trois mois maximum à compter de la réception de la demande.

Le DPO du HDH actualise le Tableau de suivi des demandes de droit.

Si la demande est directement reçue par l'établissement de santé, celui-ci se charge de :

- Répondre à la personne dans le délai d'un mois ;
- Transmettre la réponse au HDH sans délai.

Acceptable / Améliorable / À corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.2.4. Comment les personnes concernées peuvent-elles exercer leurs droits de rectification et droit à l'effacement (droit à l'oubli) ?

- Arrivée de la demande

Le canal d'entrée principal des demandes d'exercice de droit est le point de contact du DPO du HDH : par voie postale à l'adresse "À l'attention du Délégué à la protection des données, 9 rue Georges Pitard, 75015 Paris", par voie électronique à l'adresse dpd@health-data-hub.fr. Le DPO et son équipe ont accès à cette boîte aux lettres électronique. La personne concernée aura également la possibilité d'exercer ses droits via le formulaire du concentrateur du HDH, qui fera le lien avec l'EDS.

La boîte aux lettres électronique du DPO du HDH envoie un accusé de réception automatique, précisant à la personne concernée que sa demande a bien été prise en compte et qu'elle recevra une réponse sous un mois maximum.

Le DPO du HDH inscrit la demande au Tableau de suivi des demandes de droit et informe de cette demande l'établissement de santé concerné.

La personne concernée a également la possibilité d'exercer ses droits auprès de l'établissement de santé qui collecte ses données. Dans ce cas, l'établissement de santé :

- Mène les actions techniques nécessaires sur les données sources ;
- Informe le HDH dès réception de la demande, afin qu'il puisse mener les actions techniques sur les données de l'EDS, le cas échéant.

Les coordonnées du DPO du HDH et du DPO de l'établissement de santé concerné seront fournies au sein de la note d'information relative à l'EDS.

En cas de demande introduite à la fois auprès du HDH et d'un établissement de santé, le HDH se coordonnera avec l'établissement concerné en vue de déterminer les modalités de traitement de la demande et de réponse à la personne.

- Vérification de l'identité de la personne

En cas de doute raisonnable sur l'identité du demandeur, le DPO du HDH répondra manuellement en demandant un titre d'identité. Cette action suspend les délais prévus à l'article 12 du RGPD.

Si le demandeur ne répond pas à la demande du DPO du HDH sous 3 mois, la demande sera clôturée. La personne concernée sera informée par courriel de la clôture de sa demande et des motifs de cette clôture.

- Actions techniques

Droit de rectification :

Le HDH fait le lien avec l'établissement de santé ayant collecté les données qui aurait commis une erreur justifiant d'être corrigée, afin de bénéficier de leur connaissance de la personne et de leur expertise médicale, lorsque la donnée à rectifier est médicale.

L'établissement de santé concerné mène ensuite les actions de rectification nécessaires sur les données sources du dossier patient.

Concernant les données des utilisateurs de la plateforme technologique, ces derniers peuvent contacter leur autorité d'enregistrement [redacted]

[redacted]

Droit à l'effacement :

En premier lieu, le DPO du HDH vérifie que les conditions pour appliquer le droit à l'effacement sont réunies au sens de l'article 17 1. du RGPD et qu'aucune exception de l'article 17 3. n'est applicable.

- *Avant ingestion des données sur la plateforme :*

Si la demande est recevable, le HDH fait ensuite le lien avec l'établissement de santé ayant collecté les données.

L'établissement de santé concerné mène ensuite les actions d'effacement nécessaires sur l'extraction des données sources qui a vocation à être ingérée dans l'EDS.

- *Après ingestion des données sur la plateforme :*

[redacted]

Le HDH pourrait donc accepter les demandes d'exercice du droit à l'effacement pour les données de l'EDS mais ne les exécuterait pas au fil de l'eau et supprimerait toutes les données en même temps, lors d'une période dédiée à cette opération. L'exécution des demandes d'effacement en même temps, pendant une période limitée dans le temps, semble le meilleur compromis pour limiter les risques de sécurité soulevés.

Le HDH pourra répondre aux citoyens qui exercent leur droit d'effacement dans un délai d'un mois conformément au délai imposé par l'article 12 du RGPD en leur expliquant les enjeux de sécurité et comment il sera donné suite à leur demande. Le délai imposé par l'article 12 du RGPD porte sur les mesures prises à la suite d'une demande et sera ainsi respecté mais il n'emporte pas obligation que l'effacement soit exécuté concrètement dans ce délai d'un mois. Naturellement, le délai d'exécution doit être raisonnable mais il est laissé à l'appréciation du responsable de traitement qui doit en particulier prendre en compte le niveau de risques associé à l'opération. En l'espèce, une exécution des opérations d'effacement une fois par an semble le meilleur compromis mais, sous réserve de lever certaines contraintes pesant aujourd'hui sur l'architecture de la plateforme technologique, une fréquence d'exécution plus grande serait envisageable.

Concernant les données des utilisateurs de la plateforme technologique, leur traitement étant fondé sur une obligation légale, les utilisateurs ne peuvent pas demander leur effacement avant la fin de la période de conservation [redacted]

[redacted]

- Réponse à la demande

Le DPO du HDH répond à la personne concernée sous un mois, par courrier électronique ou, le cas échéant, par courrier postal.

Si la demande n'est pas recevable, le DPO du HDH répond négativement à la personne concernée, en lui exposant les motifs du refus.

Si la demande est recevable, le DPO du HDH répond positivement à la personne concernée, en lui expliquant les actions prises pour faire droit à sa demande.

S'il juge la demande complexe, le DPO informe sous un mois la personne concernée de l'allongement du délai de réponse et des motifs associés. Le DPO du HDH répondra à la personne sous trois mois maximum à compter de la réception de la demande.

Le DPO du HDH actualise le Tableau de suivi des demandes de droit.

Si la demande est directement reçue par l'établissement de santé, celui-ci se charge de :
- Répondre à la personne dans le délai d'un mois ;
- Transmettre la réponse au HDH sans délai.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.2.5. Comment les personnes concernées peuvent-elles exercer leurs droits de limitation et droit d'opposition ?

- Arrivée de la demande

Le canal d'entrée principal des demandes d'exercice de droit est le point de contact du DPO du HDH : par voie postale à l'adresse "À l'attention du Délégué à la protection des données, 9 rue Georges Pitard, 75015 Paris", par voie électronique à l'adresse dpd@health-data-hub.fr. Le DPO et son équipe ont accès à cette boîte aux lettres électronique. La personne concernée aura également la possibilité d'exercer ses droits via le formulaire du concentrateur du HDH, qui fera le lien avec l'EDS.

La boîte aux lettres électronique du DPO du HDH envoie un accusé de réception automatique, précisant à la personne concernée que sa demande a bien été prise en compte et qu'elle recevra une réponse sous un mois maximum.

Le DPO du HDH inscrit la demande au Tableau de suivi des demandes de droit et informe de cette demande l'établissement de santé concerné.

La personne concernée a également la possibilité d'exercer ses droits auprès de l'établissement de santé qui collecte ses données. Dans ce cas, l'établissement de santé :

- Mène les actions techniques nécessaires sur les données sources ;
- Informe le HDH dès réception de la demande, afin qu'il puisse mener les actions techniques sur les données de l'EDS, le cas échéant.

Les coordonnées du DPO du HDH et du DPO de l'établissement de santé concerné seront fournies au sein de la note d'information relative à l'EDS.

En cas de demande introduite à la fois auprès du HDH et d'un établissement de santé, le HDH se coordonnera avec l'établissement concerné en vue de déterminer les modalités de traitement de la demande et de réponse à la personne.

- Vérification de l'identité de la personne

En cas de doute raisonnable sur l'identité du demandeur, le DPO du HDH répondra manuellement en demandant un titre d'identité. Cette action suspend les délais prévus à l'article 12 du RGPD.

Si le demandeur ne répond pas à la demande du DPO du HDH sous 3 mois, la demande sera clôturée. La personne concernée sera informée par courriel de la clôture de sa demande et des motifs de cette clôture.

- Actions techniques

Droit à la limitation :

Le DPO du HDH vérifie que les conditions pour appliquer le droit à la limitation sont réunies au sens de l'article 18 1. du RGPD.

De même que pour le droit de rectification, une analyse au cas par cas s'imposera pour déterminer la meilleure réponse à apporter à une demande d'exercice du droit à la limitation mais les modalités de mise en œuvre du droit d'opposition devraient satisfaire également à l'exercice de ce droit.

Si la demande est recevable, le DPO du HDH demande aux établissements de santé de ne plus verser dans l'EDS les données concernées.

Droit d'opposition :

Le droit d'opposition sera pris en compte à différents niveaux :

- En amont du transfert des données vers le HDH, la personne peut exprimer son opposition auprès du HDH mais il n'est pas exclu qu'elle se manifeste aussi directement auprès de l'établissement de santé qui l'aura informée de son intention de transmettre les données au HDH. Dans les deux hypothèses, une interaction forte entre le HDH et les établissements de santé sera nécessaire pour que ces derniers puissent retenir les données à la source et ne pas les envoyer au HDH.
- En aval du transfert des données vers le HDH, il est donc trop tard pour empêcher leur circulation mais le HDH pourra donner suite au droit d'opposition en mettant en

œuvre une procédure d'exclusion, qui permettra de prendre en compte le droit d'opposition au niveau de l'EDS et pour les recherches ultérieures. Cela consiste, par un jeu d'automates, à ajouter les identifiants de stockage correspondant aux personnes qui ont exprimé leur opposition à une matrice d'exclusion. La matrice permet ensuite, de manière automatisée, de ne plus partager les données attenantes issues de l'EDS lors des futures mises à disposition de données aux porteurs de projets.
Le HDH informera la personne que son droit d'opposition a bien été pris en compte et lui garantira l'utilisation de la matrice d'exclusion pour les futures mises à disposition de données.

Par ailleurs, concernant les utilisateurs de la plateforme technologique, ces derniers ne peuvent pas s'opposer aux traitements de leurs données qui sont nécessaires à la gestion de leur compte utilisateur ni aux traitements de leurs traces d'activité du fait de l'obligation légale de maintenir la plateforme technologique en condition de sécurité. Ils peuvent néanmoins exercer leur droit à la limitation auprès du DPO du HDH. Les coordonnées du DPO seront communiquées aux utilisateurs avant leur arrivée sur la plateforme technologique et sont rappelées dans les CGU de la plateforme technologique, qu'ils signent en amont de leur accès à la plateforme.

[redacted]

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.2.6. Les obligations des sous-traitants sont-elles clairement définies et contractualisées ?

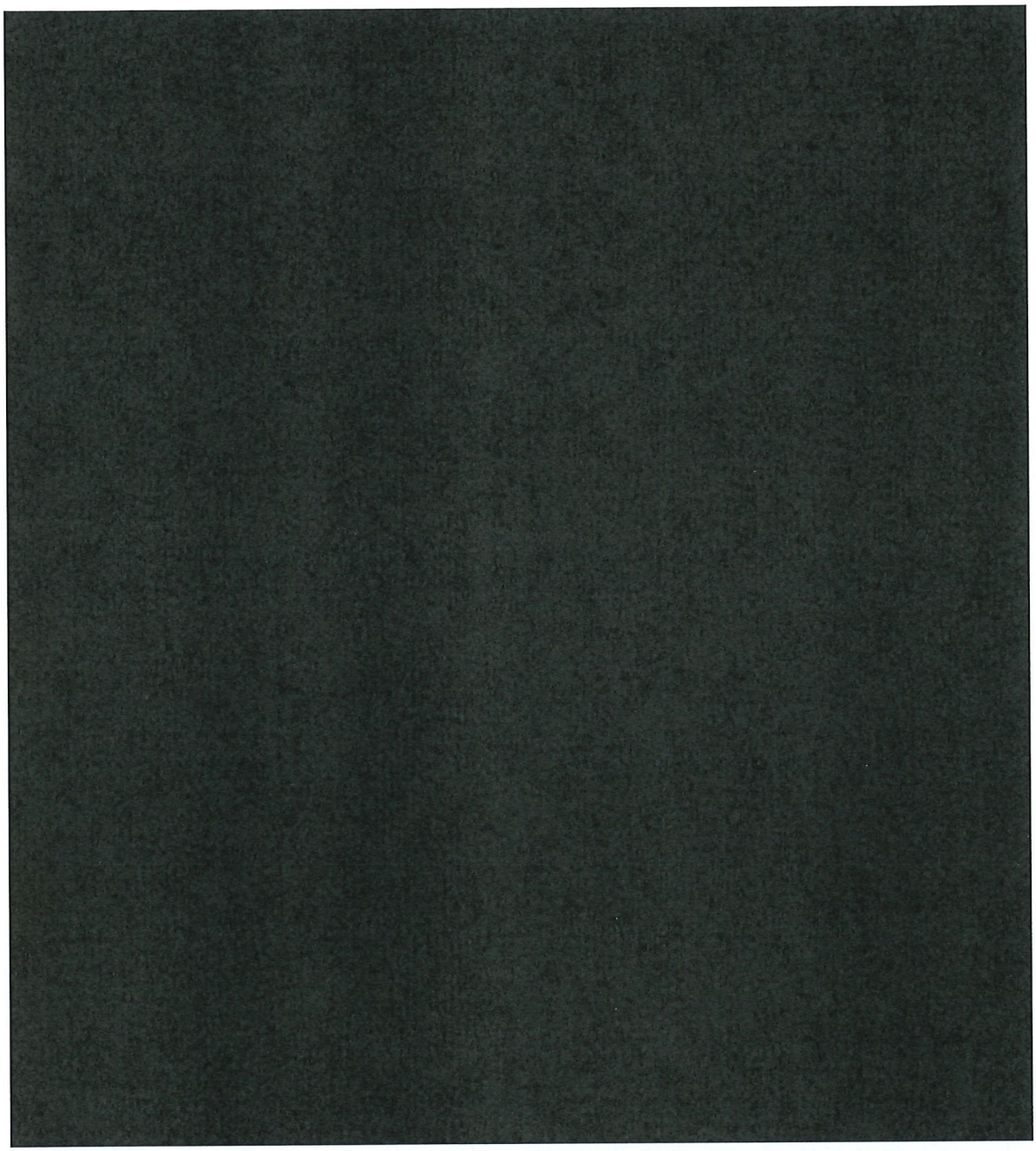
Six sous-traitants interviennent dans la réalisation des traitements de données liés à la mise en oeuvre de l'EDS :

a) Les établissements de santé :

Dans le cadre de la mise en œuvre de l'entrepôt, les HCL, le CLB, le CHRU de Nancy et la FHSJ, en leur qualité de fournisseurs de données, traiteront des données à caractère personnel pour le compte, sur instruction et sous l'autorité du responsable de traitement. Ces sous-traitants inscriront les traitements réalisés au sein de leur registre des activités de traitement conformément à l'article 30.2 du RGPD.
Conformément à l'article 28 du RGPD, la description des traitements, les obligations de ces sous-traitants y compris en matière de sécurité et de gestion des violations de données sont formalisées au sein de contrats entre le HDH et chaque établissement de santé.

b) Microsoft :

[redacted]



c) CDC Arkhinéo :

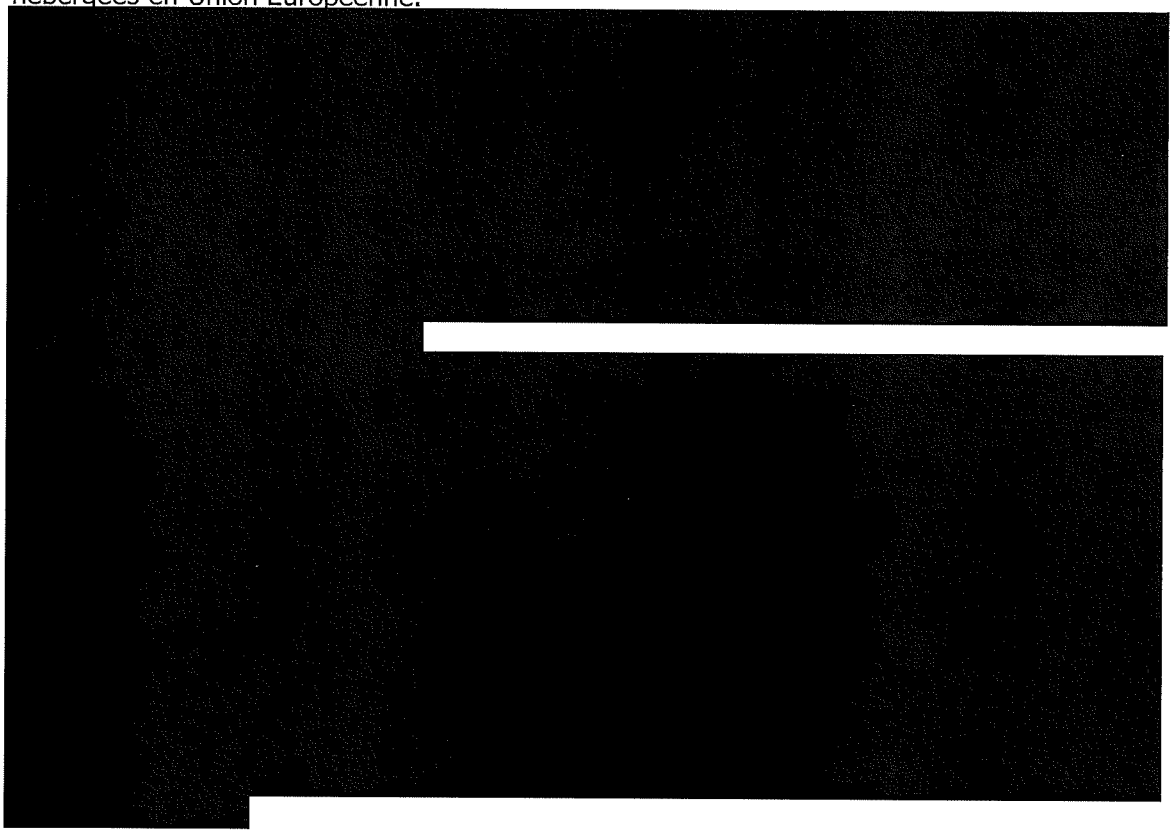
Des garanties sont apportées par le tiers archiveur CDC Arkhinéo sur la sécurité physique de ses centres de données.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

2.2.7. En cas de transfert de données en dehors de l'Union européenne, les données sont-elles protégées de manière équivalente ?

a) Sur la localisation des données au sein de l'Union Européenne :

Conformément au référentiel de sécurité du SNDS, les données de santé stockées sur la plateforme technologique sont hébergées en Union Européenne. Plus précisément, les données sont stockées dans les centres de données Microsoft situés en Zone France dans la région « France Centre » (région parisienne), certifiés « Hébergeur de données de santé ». Il est de la responsabilité du HDH de spécifier la zone géographique des ressources déployées et Microsoft s'engage contractuellement à ne pas stocker ni traiter les données en dehors de cette zone et de cette région géographiques. Le HDH met en œuvre des restrictions techniques et des contrôles pour s'assurer que les ressources déployées sont effectivement hébergées en Union Européenne.



b) Sur le risque d'application extraterritoriale des lois américaines :

Un arrêt de la CJUE du 16 juillet 2020, dit « Schrems II » , a invalidé le Privacy Shield au motif que « *la primauté des exigences relatives à la sécurité nationale, à l'intérêt public et au respect de la législation américaine, rendant ainsi possibles des ingérences dans les droits fondamentaux des personnes dont les données sont transférées vers ce pays* ». Le contrat entre le HDH et Microsoft s'appuie sur les clauses contractuelles types dont la validité a été confirmée par la Cour dans ce même arrêt. Néanmoins, dans la mesure où la Cour a tout de même pris soin de rappeler que les personnes concernées par des données transférées doivent bénéficier d'un niveau de protection équivalent à celui garanti au sein de l'UE et que cette protection doit être évaluée, au-delà des clauses contractuelles, en tenant compte du cadre juridique d'un éventuel accès par les autorités publiques du pays tiers, la vérification du niveau de protection des données des utilisateurs de la plateforme technologique concernés par un éventuel transfert de données hors UE a été menée par le HDH et Microsoft.
Le HDH s'est fait accompagner dans cette investigation par un cabinet d'avocats spécialisé en protection des données à caractère personnel, notamment dans les transferts de données internationaux. Cette analyse a abouti à la conclusion que l'argument selon lequel il existerait

un risque pour le traitement des données de santé du fait de l'usage de la solution Microsoft ne semble pas justifié après un examen approfondi (i) des lois de surveillance américaines, (ii) des mesures techniques, organisationnelles et contractuelles en place négociées par le HDH avec Microsoft, (iii) de la Décision Schrems II et notamment des dérogations de l'article 49 du RGPD qu'elle suggère d'appliquer et enfin (iv) des dispositions du chapitre V du RGPD régissant les conditions de transfert hors UE. L'analyse détaillée du cabinet d'avocats a été communiquée précédemment à la CNIL.

Cette analyse a par la suite été confirmée par la décision du Conseil d'Etat rendue le 23 novembre 2022 à propos du décret du 29 juin 2021 relatif au traitement de données à caractère personnel dénommé « Système national des données de santé », décision qui rejette tout manquement au RGPD.

Acceptable / Améliorable / A corriger	**Commentaires d'évaluation**	**Mesures correctives**
Acceptable		

[redacted]

3.1. Mesures existantes ou prévues

3.1.1. Mesures contribuant à traiter des risques liés à la sécurité des données

3.1.1.1. Chiffrement

[redacted]

[redacted]

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.1.2. Anonymisation

[redacted]

[redacted]

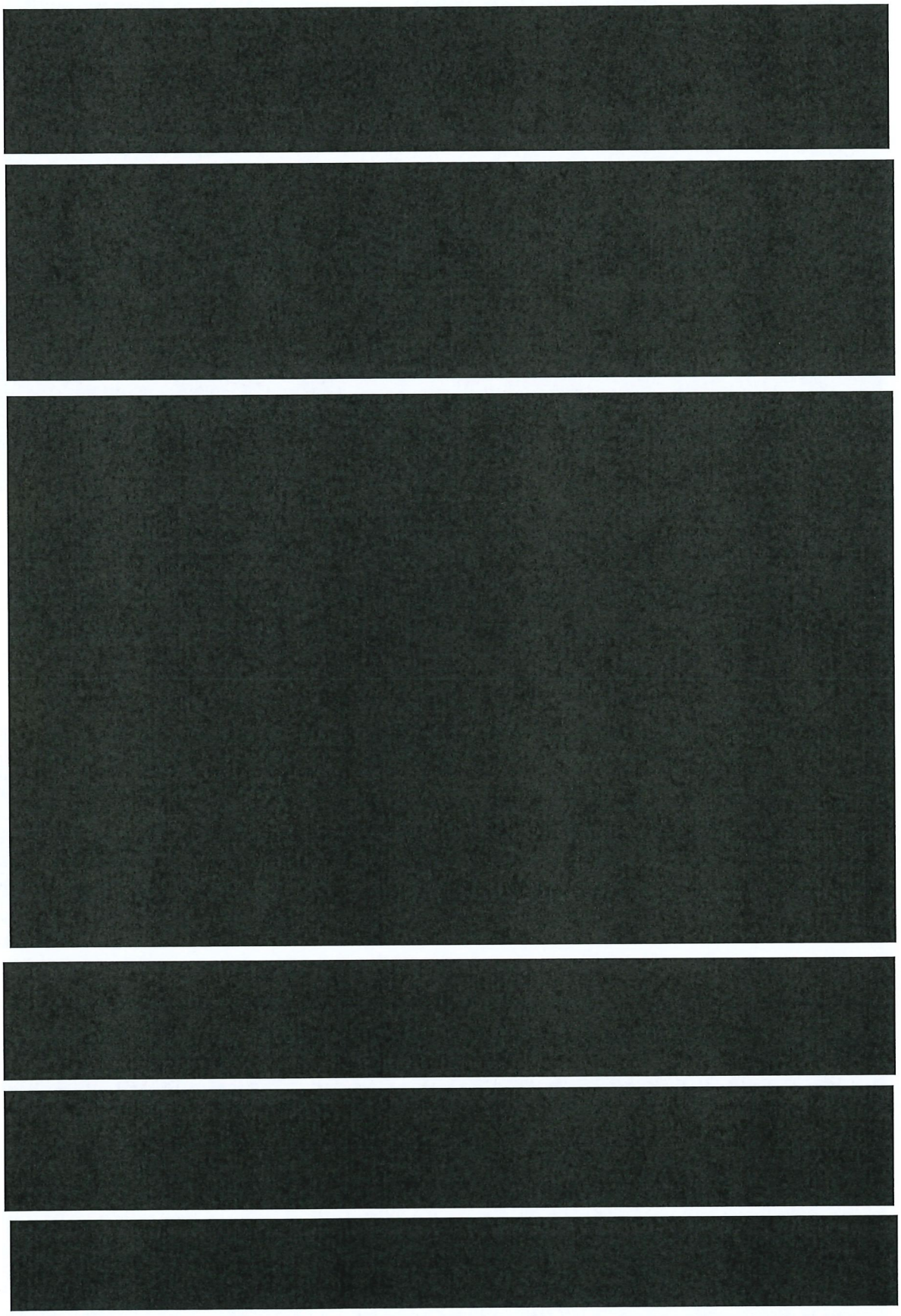
3.1.1.3. Cloisonnement des données (par rapport au reste du SI)

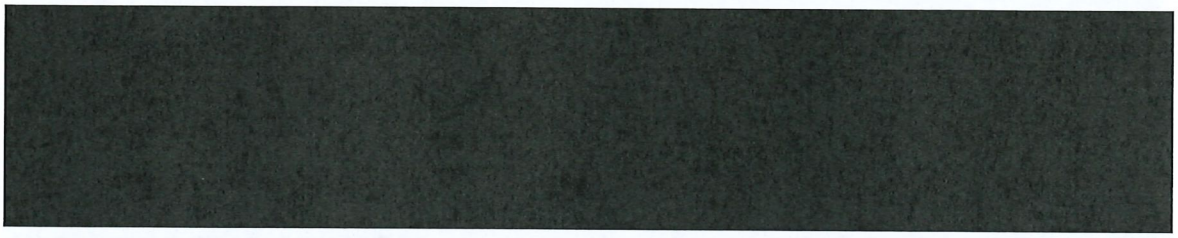
Cloisonnement des espaces :

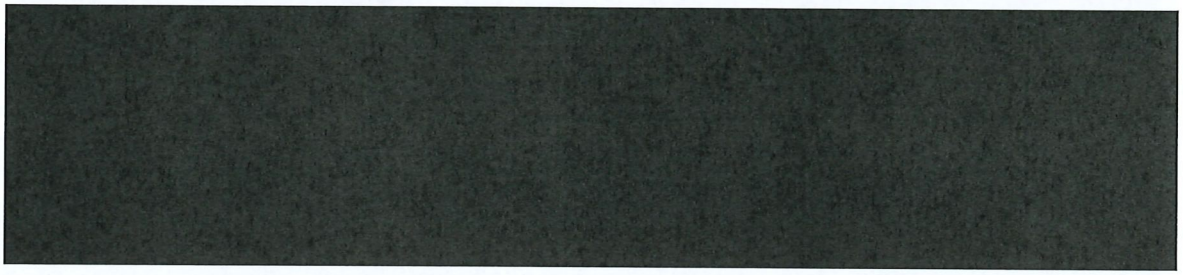
[redacted]

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable	Les données sont correctement cloisonnées.	

3.1.1.4. Contrôle des accès logiques

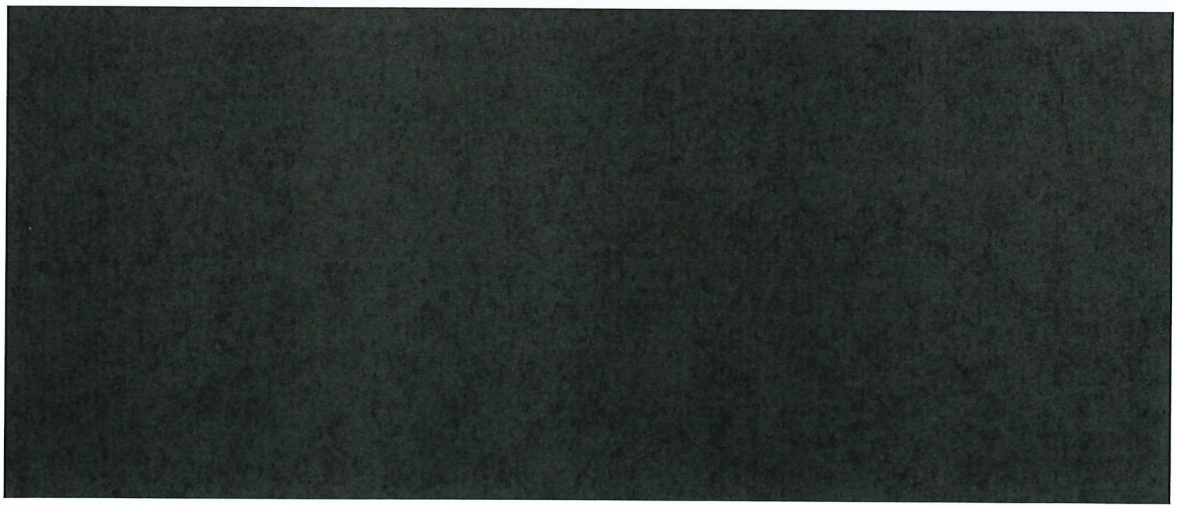


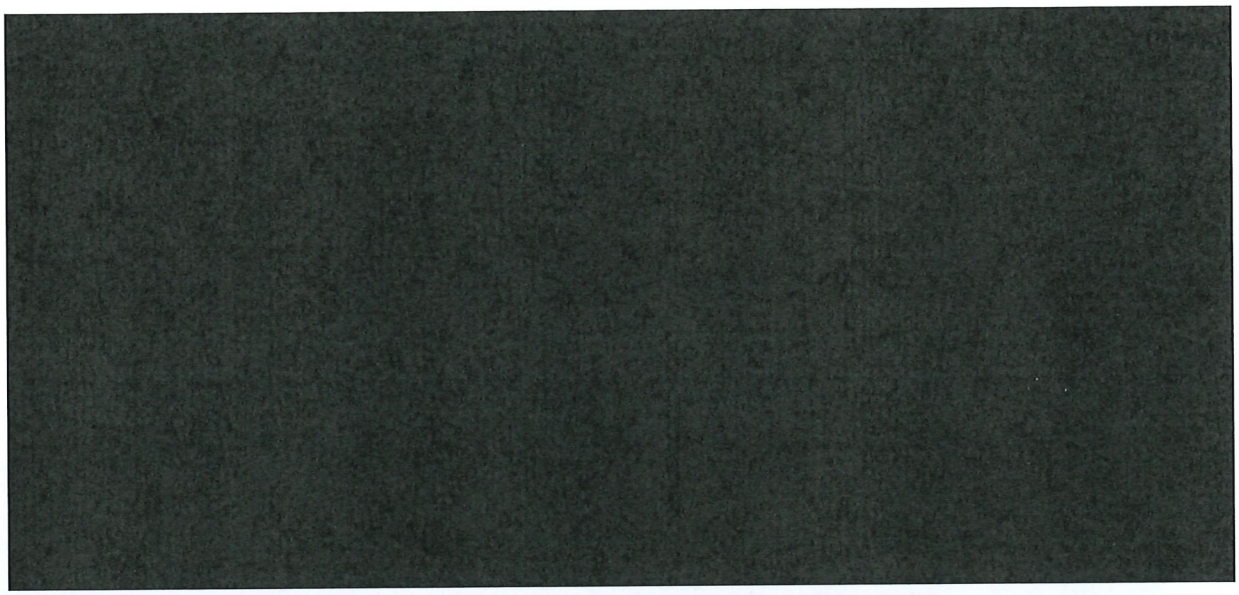




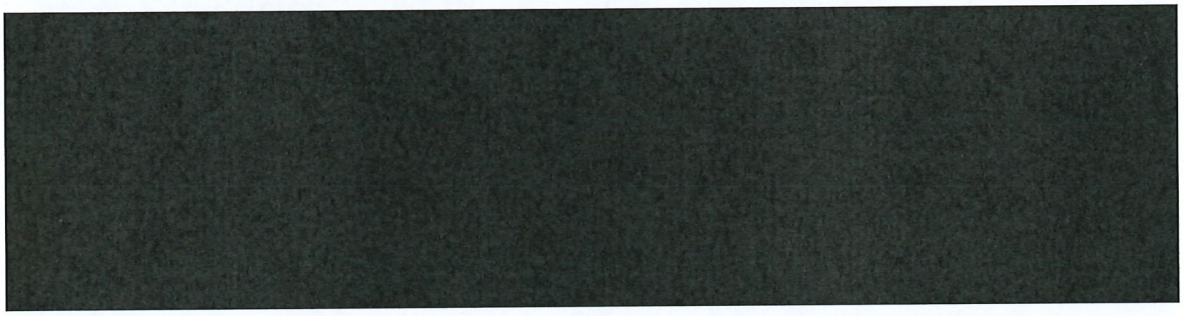
Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.1.5. Journalisation





3.1.1.6. Contrôle d'intégrité



Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.1.7. Archivage

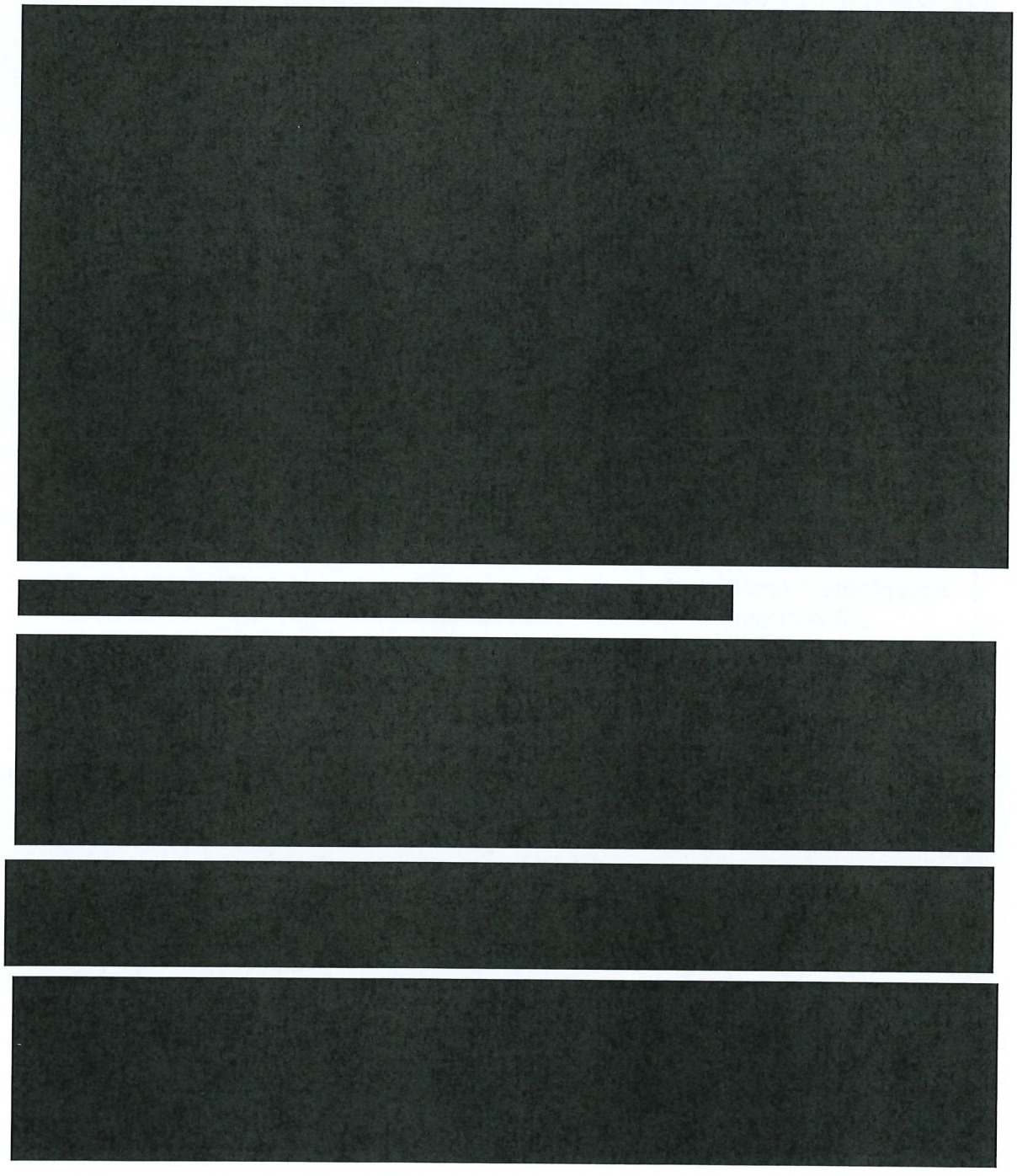
Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

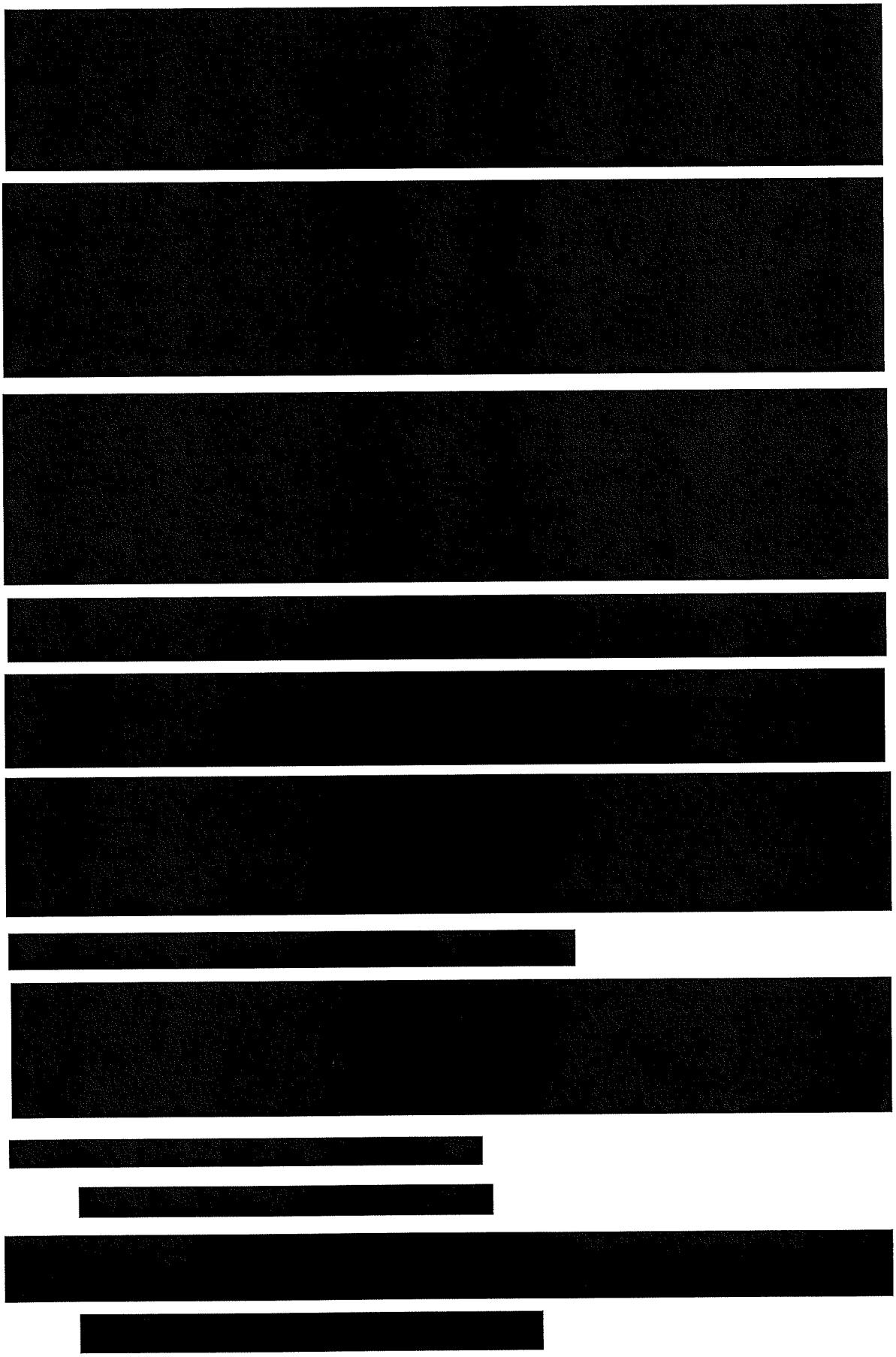
3.1.1.8. Sécurité des documents papiers

Les données de santé traitées sur la plateforme technologique ne sont pas au format papier.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.1.9. Pseudonymisation





[redacted]

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2. Mesures générales de sécurité

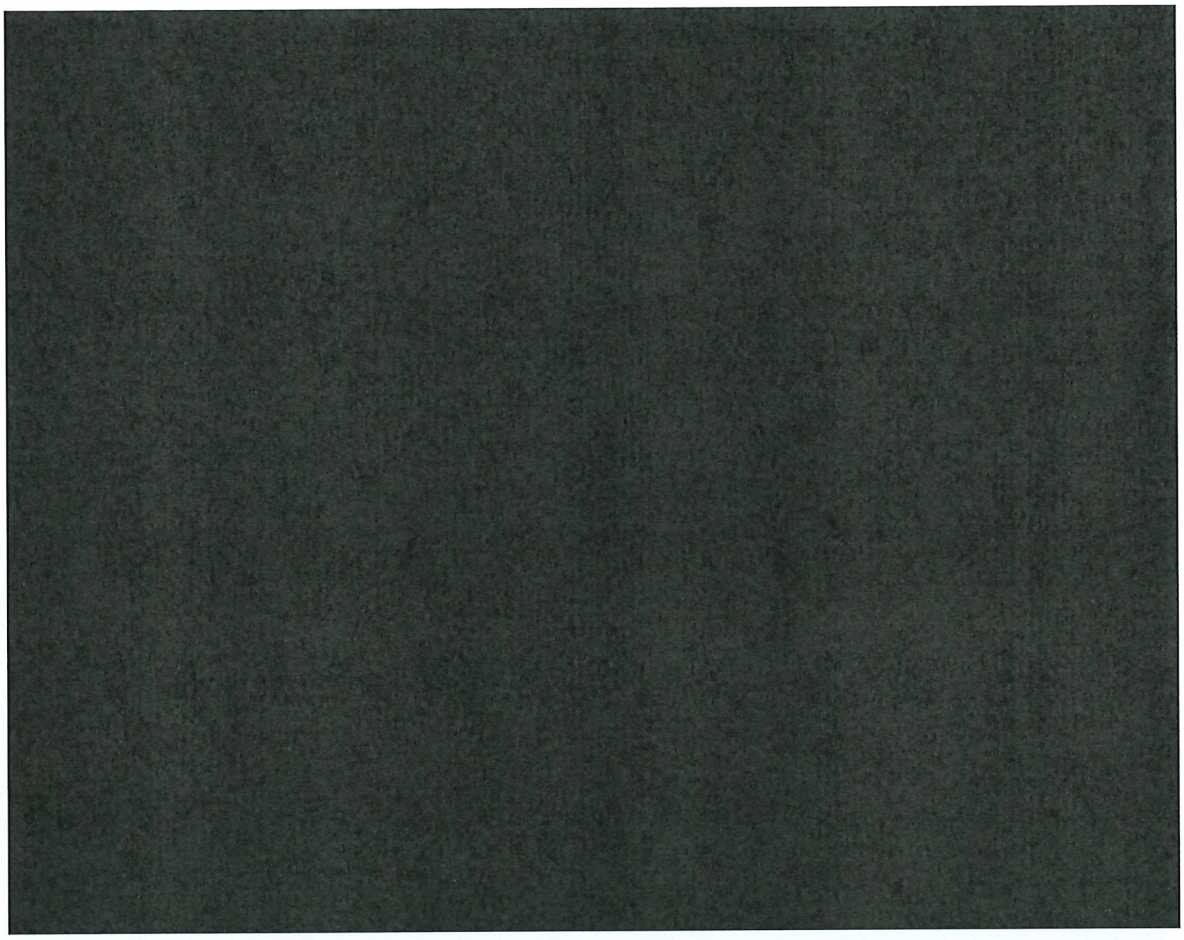
3.1.2.1. Sécurité de l'exploitation

[redacted]

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2.2. Lutte contre les logiciels malveillants

[redacted]



Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

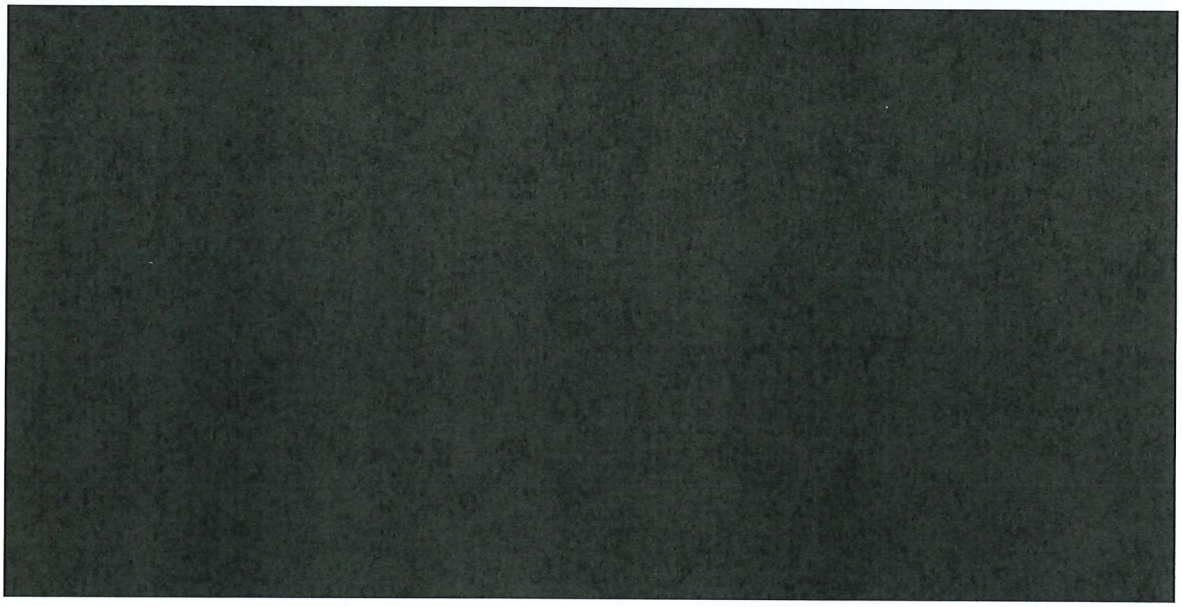
3.1.2.3. Sauvegardes

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2.4. Maintenance

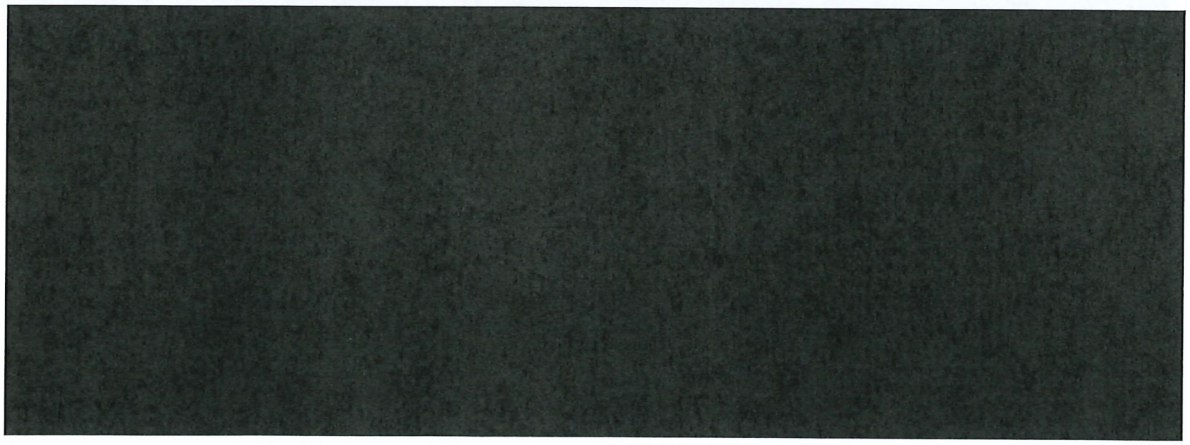
Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

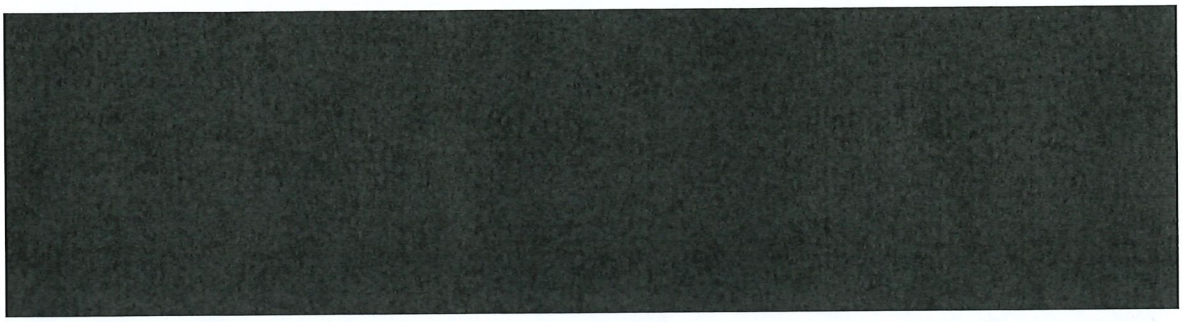
3.1.2.5. Sécurité des canaux informatiques (réseaux)



Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

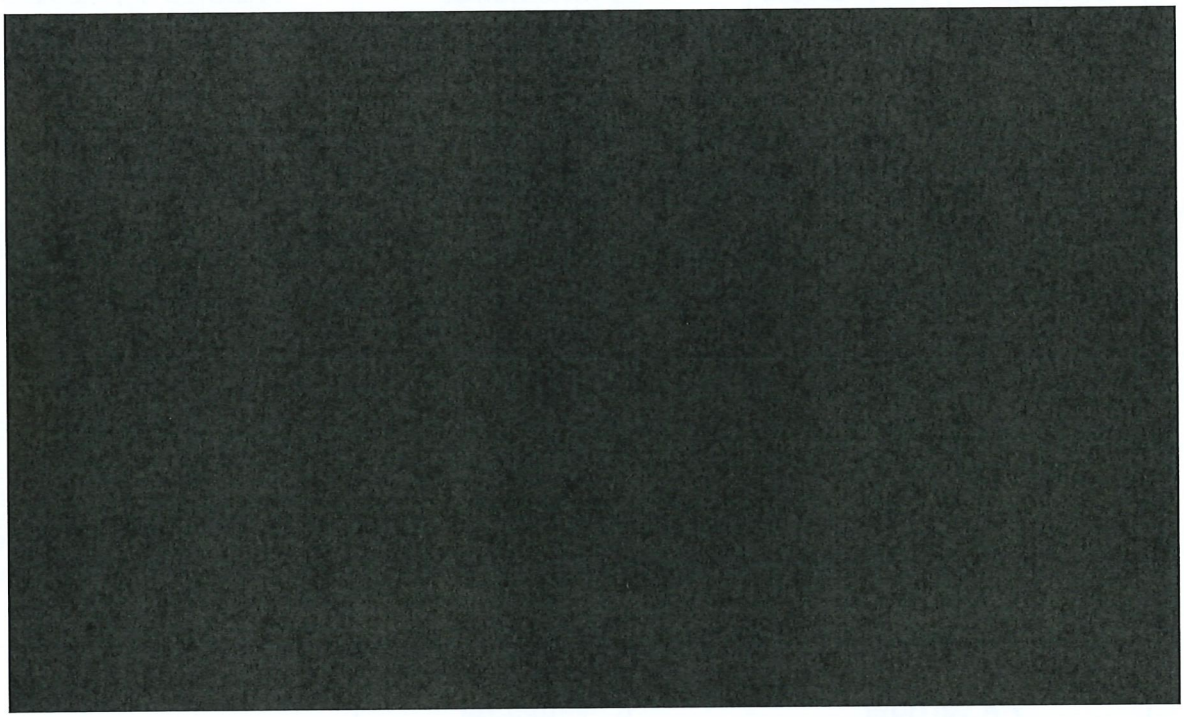
3.1.2.6. Sécurité physique





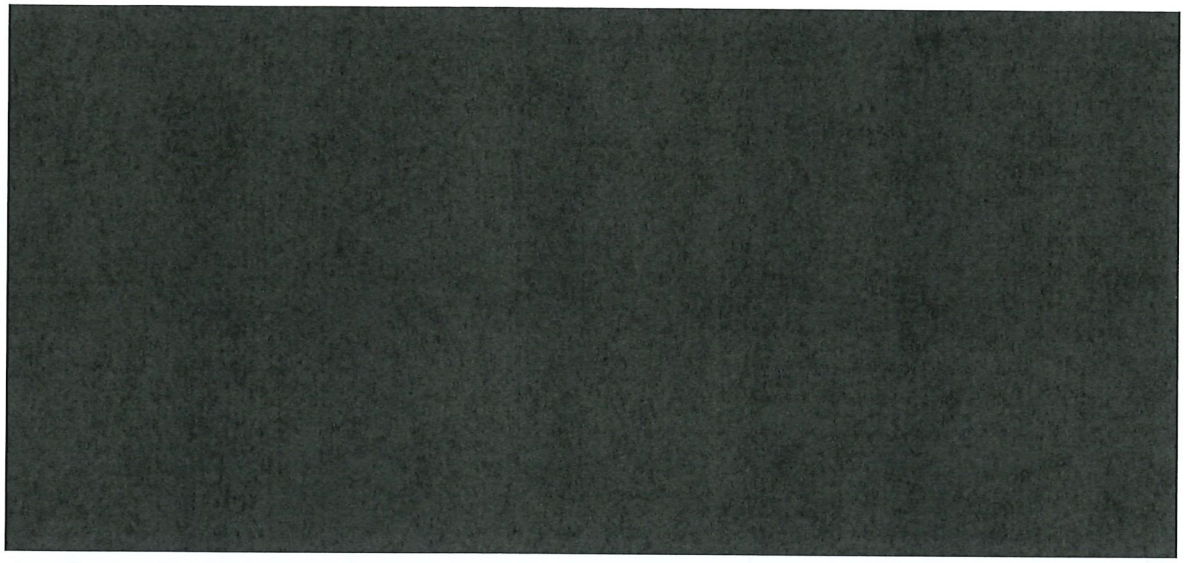
Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2.7. Traçabilité



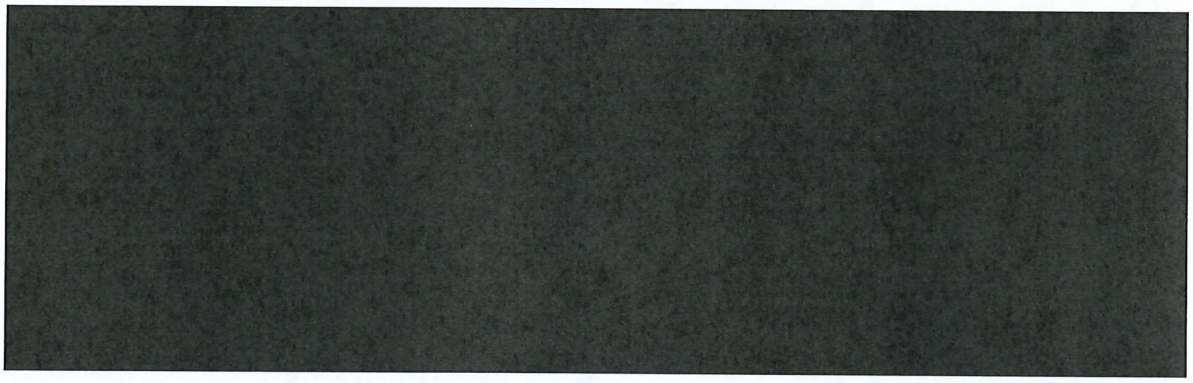
Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2.8. Sécurité du matériel



Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2.9. Eloignement des sources de risque



Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.2.10. Protection contre les sources de risque non humaines

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3. Mesures organisationnelles (gouvernance)

3.1.3.1. Organisation

Le HDH dispose d'une équipe dédiée à la protection des données personnelles composée d'un Délégué à la Protection des Données (DPO) et de deux juristes spécialisés en protection des données personnelles. Le DPO est le garant de la protection des données à caractère personnel au sein de l'organisme.
Le pôle DPO assure la sensibilisation des collaborateurs du HDH et des utilisateurs projet aux règles à respecter dans le cadre de traitements de données à caractère personnel. Il organise les procédures de réponses aux demandes d'exercice des droits et réclamations adressées par les personnes concernées par les traitements (y compris l'EDS) et, selon leur nature, les instruit ou les transmet aux services compétents.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3.2. Politique (gestion des règles)

Les principes de sécurité du HDH sont décrits dans la Politique de sécurité des systèmes d'information (PSSI). Les règles et procédures qui décrivent la mise en œuvre opérationnelle de la PSSI impliquent le DPO dès que la protection des données à caractère personnel est concernée.

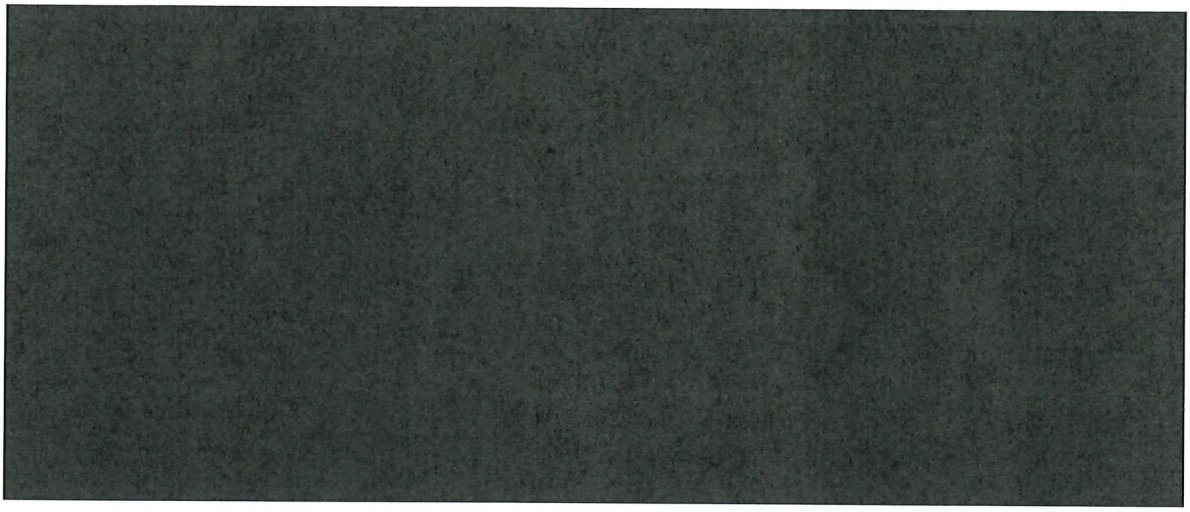
Le HDH dispose également d'une charte informatique décrivant notamment les règles de classification de l'information. Chaque agent du HDH s'engage par écrit à respecter cette charte lors de son circuit d'arrivée.

Les interactions avec la plateforme technologique du HDH sont aussi encadrées par les conditions générales d'utilisation. Chaque utilisateur de la plateforme s'engage à respecter ces conditions en les signant dans le cadre de la création de ses accès.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3.3. Gérer les risques

[redacted]



Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

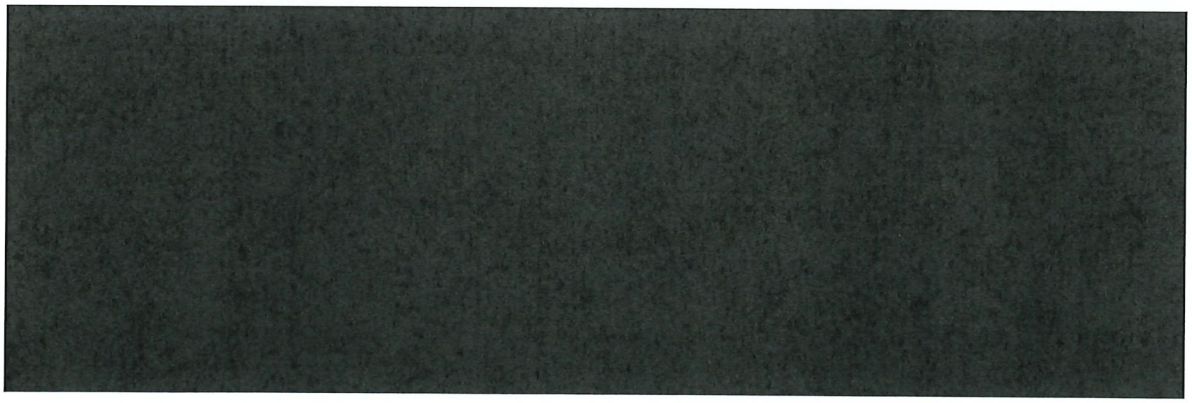
3.1.3.4. Intégrer la protection de la vie privée dans les projets

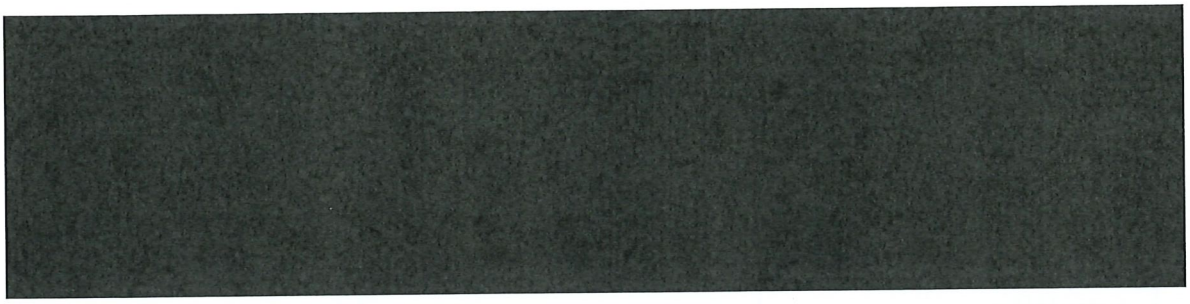
La collecte et la conservation des données n'ont pour seul objectif que de les mettre à disposition des utilisateurs habilités, soit parce qu'ils interviennent dans des projets présentant un caractère d'intérêt public, soit à des fins de recherche et développement quand il s'agit des data scientists du HDH.

Dans le cadre d'un projet, le porteur de projet est soumis à une démarche d'homologation pendant laquelle la protection de la vie privée est prise en compte, notamment par la réalisation d'une AIPD propre à ce projet.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3.5. Gestion des incidents de sécurité et les violations de données





Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3.6. Gestion des personnels

Un plan de sensibilisation sur la sécurité des systèmes d'information, sous la responsabilité du RSSI du HDH, et un plan de sensibilisation à la protection des données à caractère personnel, notamment des données de santé, sous la responsabilité du Délégué à la Protection des Données, sont mis en oeuvre tout le long de l'habilitation d'un utilisateur à accéder à la plateforme technologique, conformément au circuit d'arrivée et de départ.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3.7. Relation avec les tiers

Comme évoqué en 2.2.6, « Les obligations des sous-traitants sont-elles clairement définies et contractualisées ? », six sous-traitants interviennent dans le cadre de l'EDS. La relation du HDH avec ces tiers est décrite au paragraphe susmentionné.

Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

3.1.3.8. Superviser la protection de la vie privée

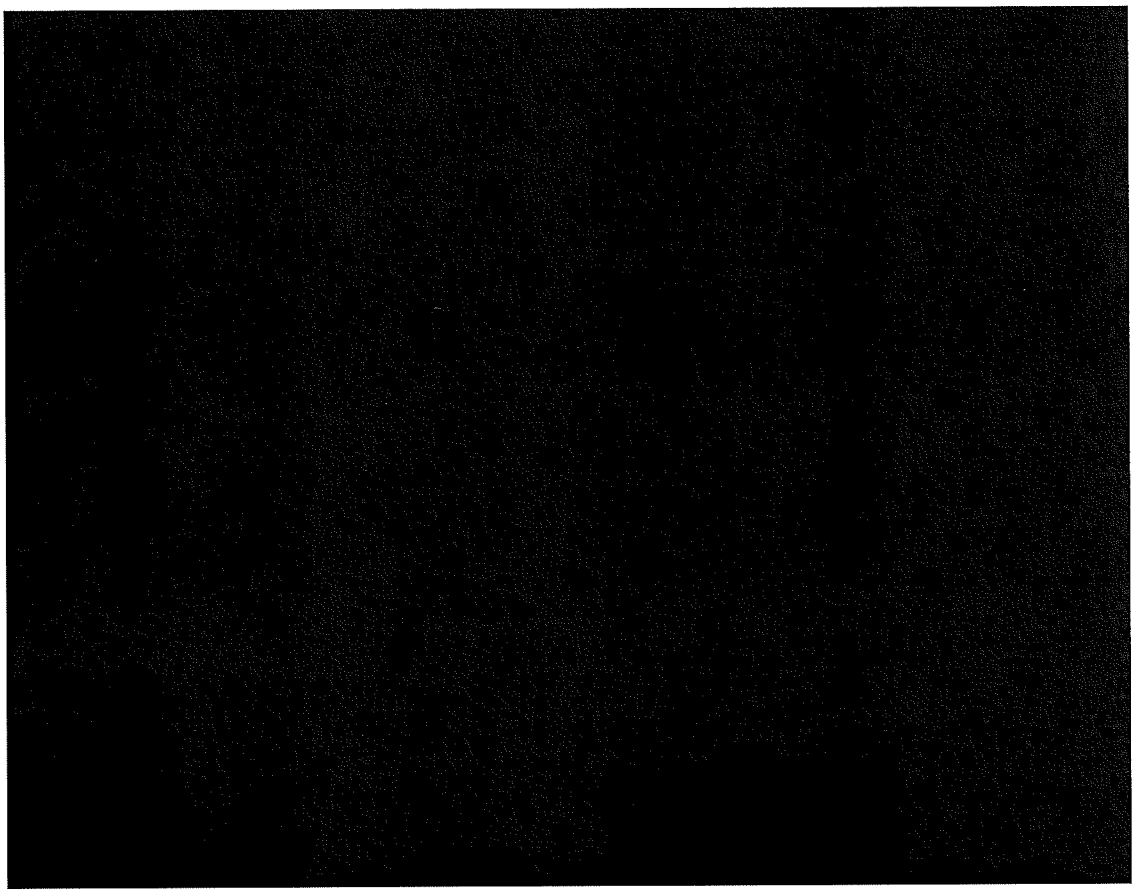
La supervision de la protection de la vie privée est assurée d'abord dans le cadre de la démarche d'homologation de la plateforme technologique qui s'appuie notamment sur une analyse de risques concernant la sécurité des systèmes d'information et un audit technique réalisé par un Prestataire d'audit de la sécurité des systèmes d'information qualifié par l'ANSSI.

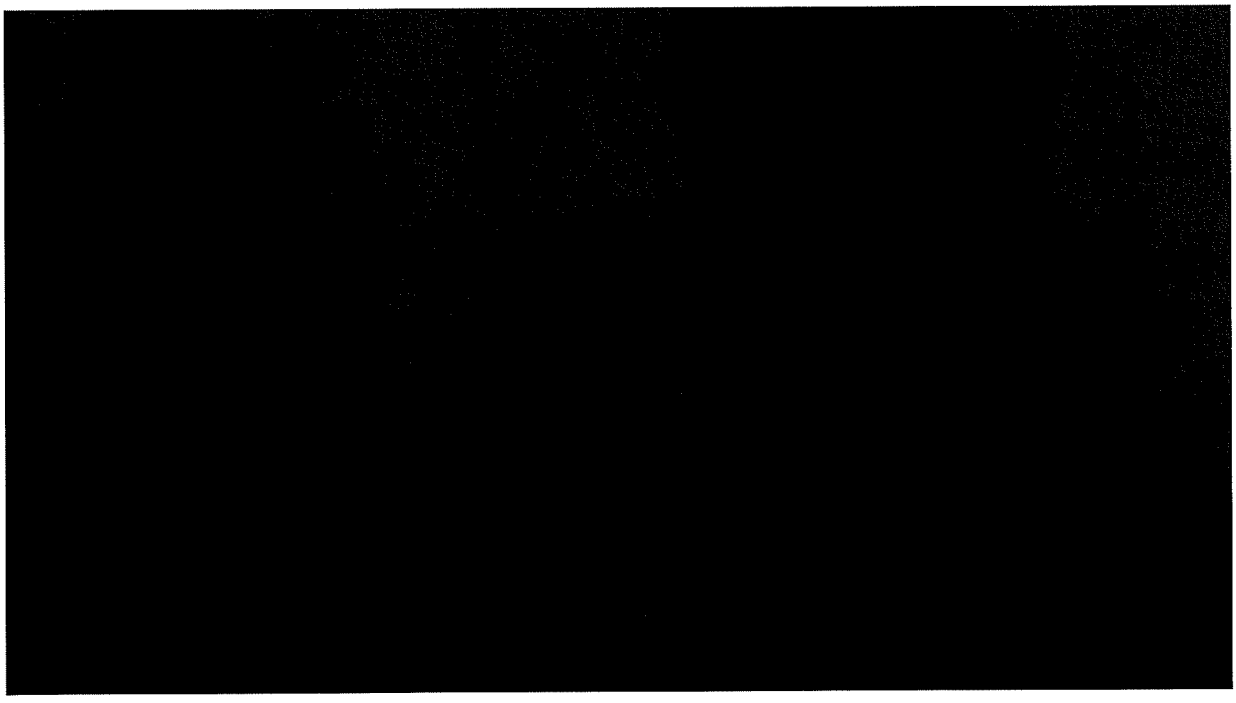
L'homologation de la plateforme technologique est mise à jour à chaque évolution majeure de la plateforme technologique.

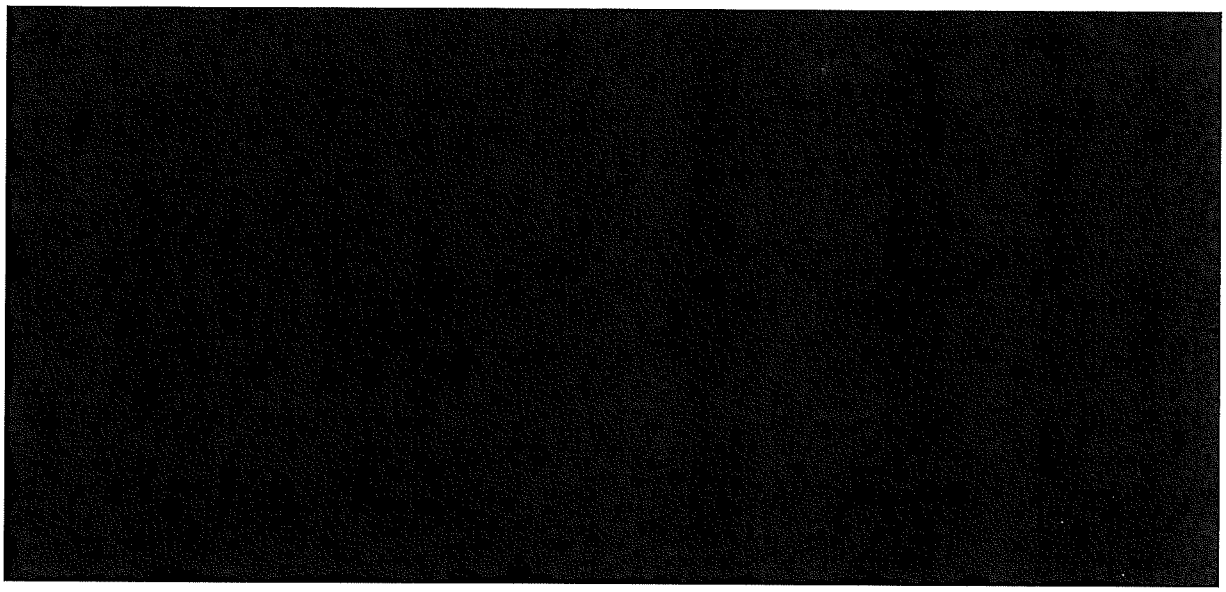
Ensuite, la présente AIPD permet de détailler les enjeux liés aux questions de protection des données à caractère personnel et les mesures techniques et organisationnelles prises pour protéger ces données dans le cadre de l'EDS.

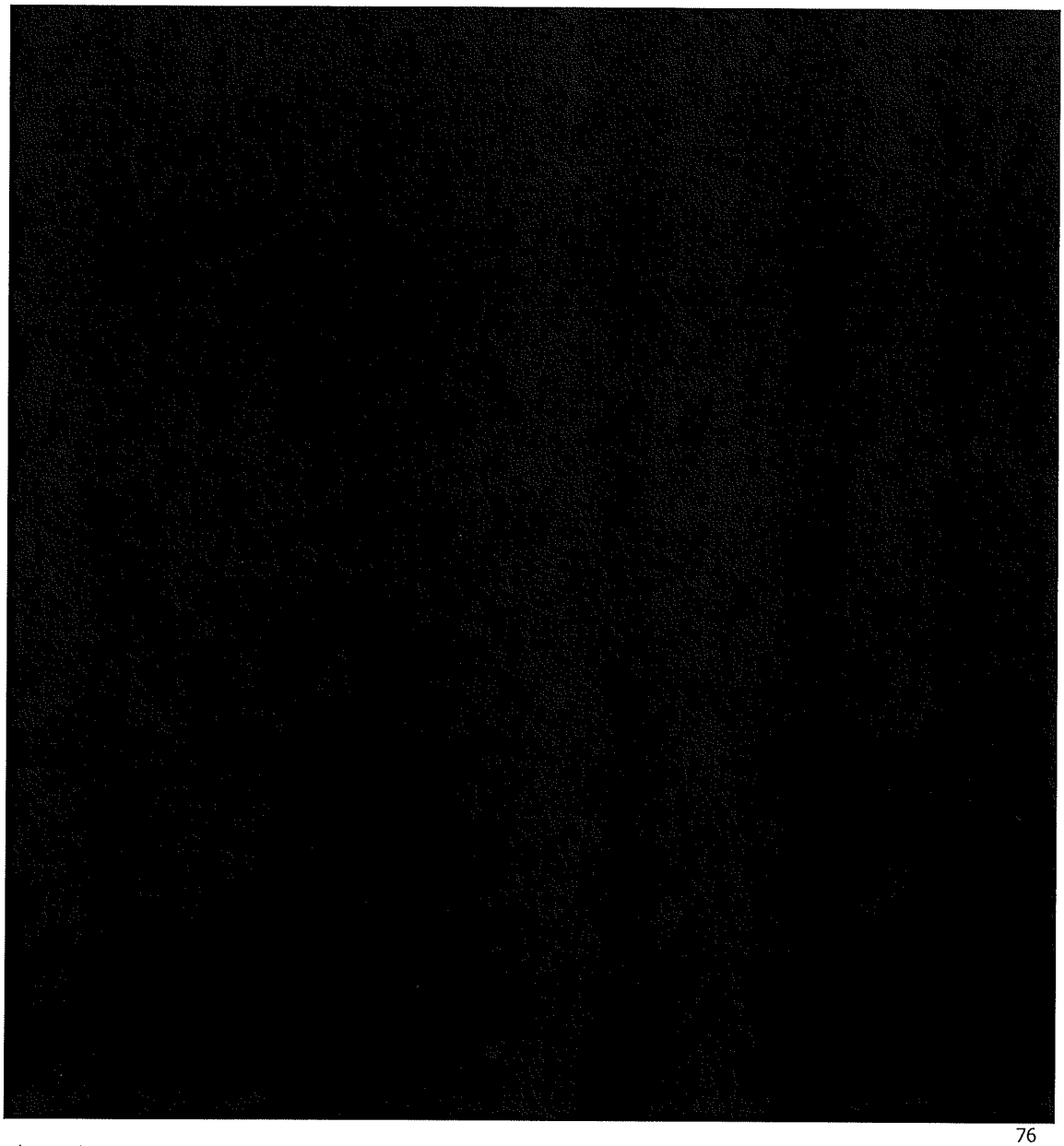
Acceptable / Améliorable / A corriger	Commentaires d'évaluation	Mesures correctives
Acceptable		

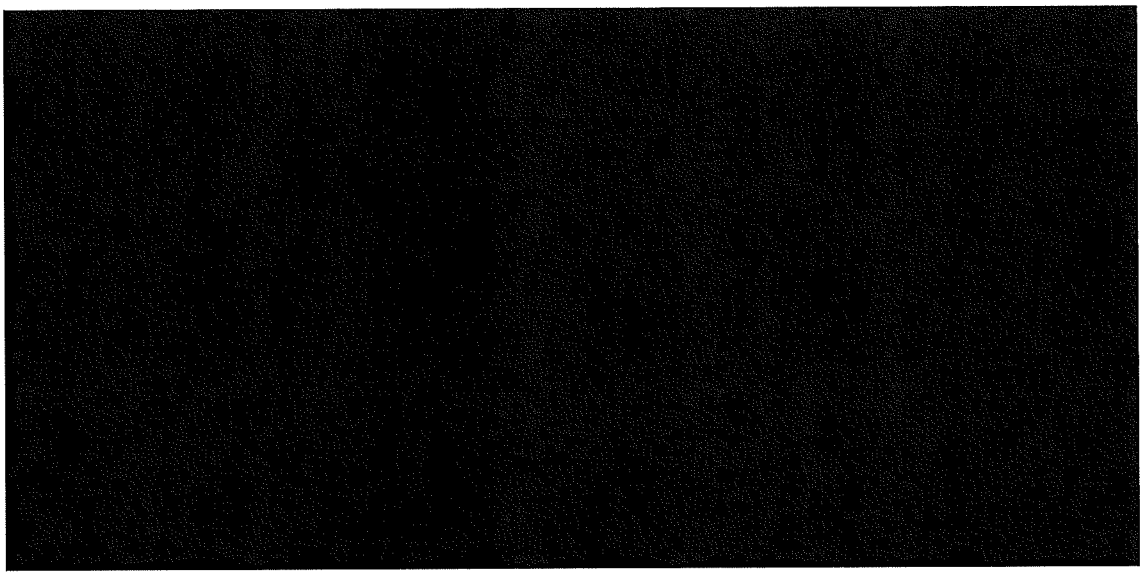
[redacted]

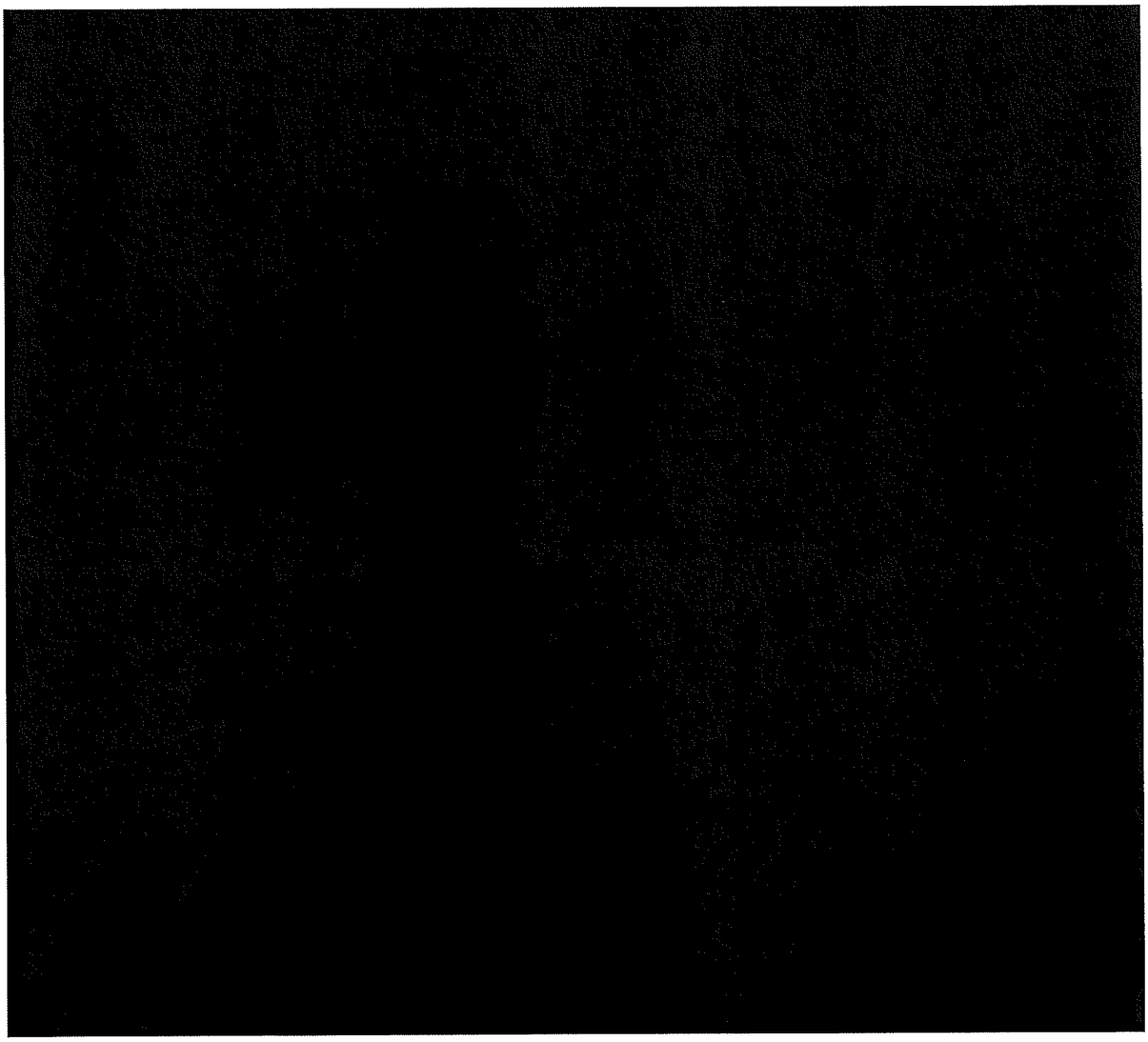


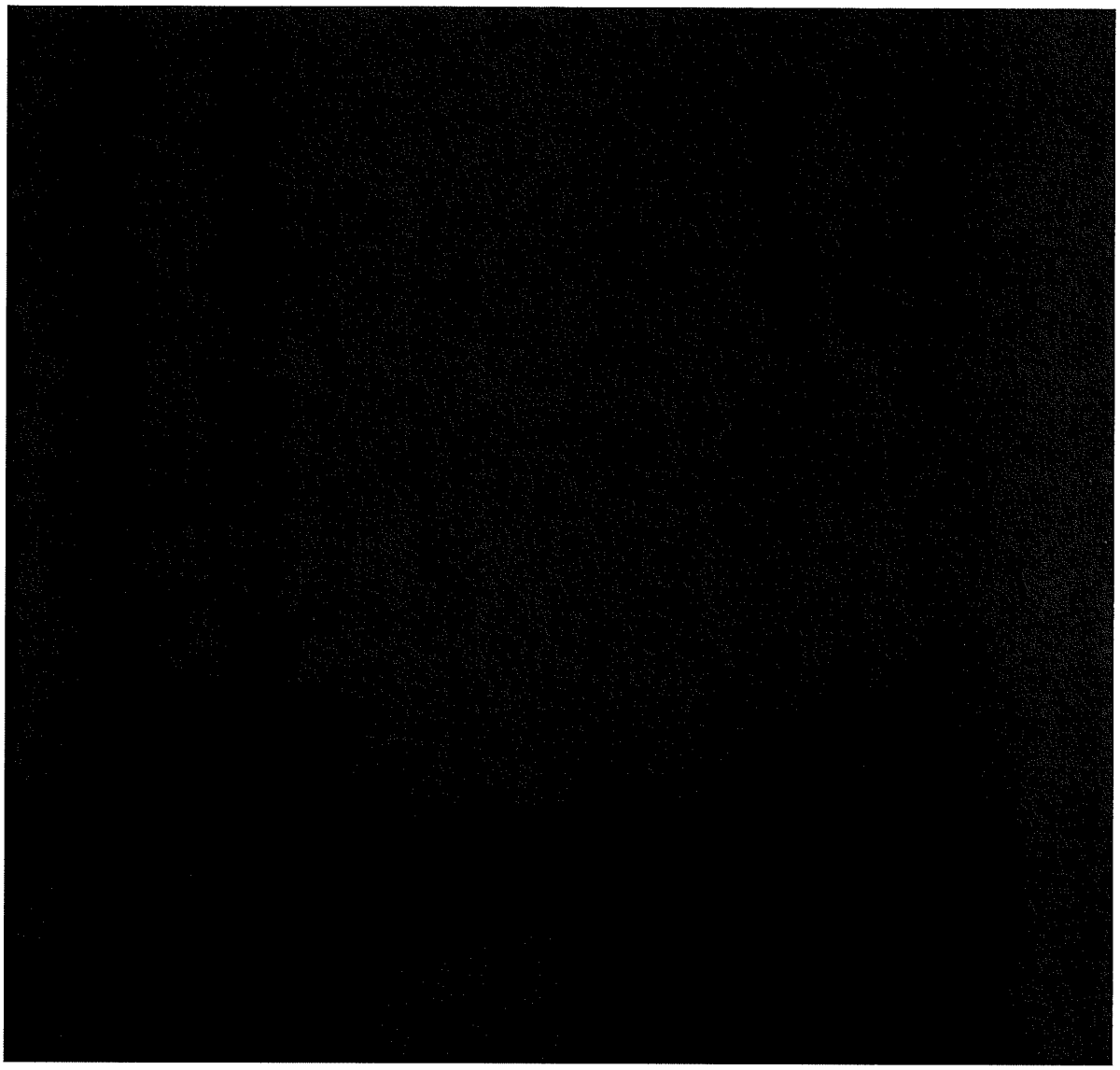


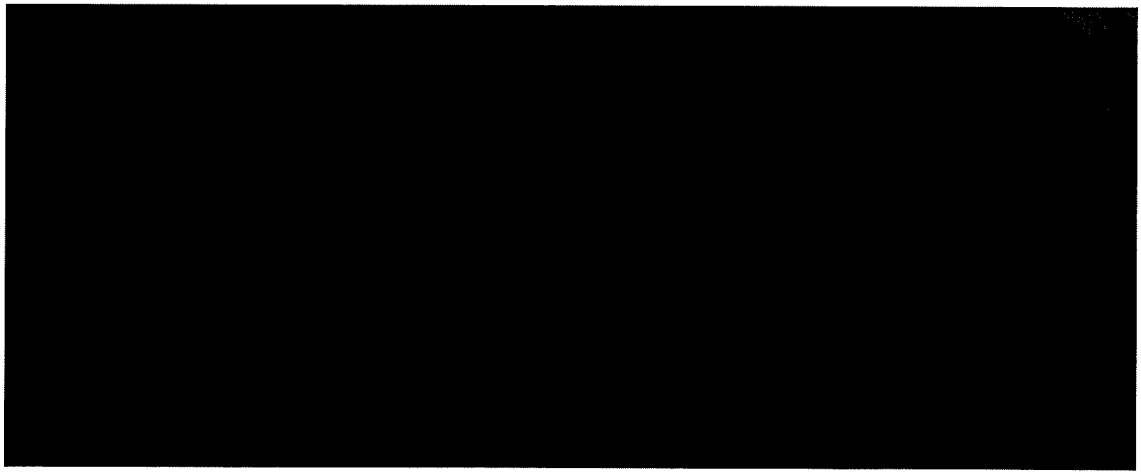


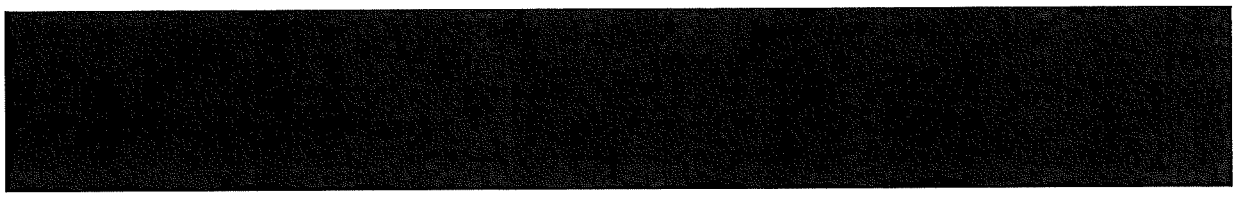












4. Formalisation de la validation

Avis du Délégué à la Protection des Données

L'EDS EMC2 répond à des enjeux majeurs pour la recherche en santé, dans la mesure où il a vocation à regrouper au sein d'une base multicentrique chaînée au SNDS des données qui permettront de traiter des questions relatives à la pharmacovigilance et à la pharmaco-épidémiologie.

En termes de conformité en matière de protection des données, les mesures détaillées dans l'AIPD conduiront au respect des principes du RGPD. Pour ce faire, l'EDS a pris en compte dès sa conception le Référentiel sur les entrepôts de données de santé de la CNIL. S'il n'y est pas totalement conforme en raison notamment du recours à des données issues de la base principale du SNDS, il s'en approche le plus possible.

Au regard de ses attributions prévues par les textes, le HDH semble pleinement légitime à porter ce projet d'EDS. Les catégories de données concernées et les durées de conservation associées ont été identifiées en considération des objectifs poursuivis par l'EDS, dont l'intérêt public a été démontré dans la présente AIPD. Également, les relations avec les différentes parties prenantes sont encadrées conformément au RGPD, au moyen d'instruments contractuels aboutis.

La transparence est une composante essentielle du projet, avec des modalités d'information individuelles propres à chaque établissement de santé qui octroient une transparence adaptée à leur environnement. À l'exception du droit à la portabilité, qui aurait assez peu de sens ici, les personnes concernées peuvent exercer la totalité de leurs droits, y compris le droit à l'effacement de leurs données. Le droit d'opposition peut quant à lui être activé de manière inconditionnelle, en ce qu'il n'est pas subordonné à la fourniture d'un motif particulier.

[redacted] À ce titre, le fait que l'EDS ait vocation la plateforme technologique du HDH telle qu'homologuée en novembre 2020 confère à ce projet un haut niveau de sécurité. De la même manière, les procédures qui encadreront l'accès aux données de l'EDS par les porteurs de projet ne sont pas nouvelles, elles sont les mêmes que celles qui sont déjà appliquées et pratiquées régulièrement par le HDH.

En conclusion, les conditions semblent réunies pour que l'EDS EMC2 soit mis en œuvre par le HDH de façon à garantir le respect des principes relatifs à la protection des données personnelles.

Date et signature :

Le 28 avril 2023
Léa Rizzuto

Décision du responsable de traitement

Le 28 avril 2023, Mme Stéphanie Combes, directrice de la Plateforme des données de santé et autorité qualifiée pour la sécurité des systèmes d'information, valide l'Analyse d'impact relative à la protection des données (AIPD) de l'entrepôt de données de santé EMC2 HDH.

La manière dont il est prévu de mettre en œuvre les mesures à la fois juridiques et techniques et de traiter les risques est en effet jugée acceptable au regard de ces enjeux. La mise en œuvre du plan d'actions devra être démontrée ainsi que l'amélioration continue de l'AIPD.

Date et signature :
Le 28 avril 2023
Stéphanie Combes