HEALTH
DATA HUB

Health Data Hub

Séminaire BNDMR

09/02/2023

Le Health Data Hub est un service à destination des acteurs de l’ écosystème des données de santé

Le Health Data Hub a été impulsé et soutenu dans le cadre de politiques publiques d'ampleur

Les usages des données de santé se multiplient et accéder aux sources de données dans des délais les plus courts possibles est essentiel. Le Health Data Hub est une structure publique créée fin 2019 visant à faciliter l'accès aux données de santé aux projets d'intérêt public, dans la continuité de l'ouverture du SNDS en 2016.

Un **guichet unique** sur les données de Santé en France

Une **plateforme sécurisée** et **à l'état de l'art**

Un **catalogue de données**, incluant une des plus larges bases de données médico-administratives au monde

Un projet impulsé et soutenu par des politiques publiques d'ampleur

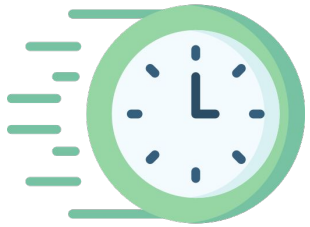
Créé à la suite du **rapport Villani** promouvant la mise en place de grandes plateformes de partage de données pour favoriser l'IA

Projet phare et un contributeur actif de la **Stratégie Nationale pour l'intelligence artificielle**

L'un des grands projets du **FTAP** et l'un des rares auditionnés en mars 2022 par la Ministre Amélie de Montchalin

L'une des actions de la **feuille de route du Numérique en Santé** 2019-2022 du Ministère de la Santé et de la Prévention

En près de 3 ans, le Health Data Hub est devenu une structure innovante et fonctionnelle



Créé seulement 4 mois après la promulgation de sa loi constitutive, le HDH a franchi d'importants jalons dès ses premières années d'activité

Mise en production d'une plateforme technologique en seulement un an

Recrutement rapide de collaborateurs pour atteindre 85 personnes au 30 novembre 2022

Accompagnement de 80 projets dont une demi douzaine sont déjà terminés ou ont déjà publié des résultats

Développement de liens avec tout l'écosystème, y compris la société civile, concrétisés par **plus d'une centaine de partenariats actifs**

Près de 440 interventions publiques depuis 2021 et **des participations dans de grands projets comme PariSanté Campus, FIAC, DARWIN**

Le Health Data Hub est co-responsable de traitement de la base principale du SNDS et responsable de traitement du catalogue

Quelles données sont disponibles dans la base principale du SNDS ?

- ✓ Le SNDS est une **compilation de bases de données préexistantes alimentées par des flux de données**
- ✓ **Ces bases de données sont déjà reliées entre elles** afin d'augmenter les opportunités de recherche et d'innovation
- ✓ **Les bases de données externes peuvent être chaînées au SNDS**

Le SNDS est rendu disponible sur le portail de l'Assurance maladie, après avis favorable du comité éthique et scientifique (CESREES) et d'une autorisation de la CNIL.

3 bases de données principales

SNIIRAM
DCIR

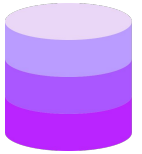
Données de remboursements de l'Assurance maladie

depuis 2006

PMSI

Programme de médicalisation des systèmes d'information

permet de décrire l'activité médicale des établissements de santé publics et privés, d'abord à des fins d'allocation budgétaire des hôpitaux



CépiDc

Causes médicales de décès

à être associé au SNIIRAM. *La base de données est anonymisée*

Première liste des bases de la base principale et du catalogue du SNDS portée par l'arrêté du 12 mai 2022

DGS

SIDEP : Système d'information national de dépistage populationnel de la Covid-19

Système d'information Vaccin Covid

Banque Nationale pour les Maladies Rares (**BNDMR**)

Inserm

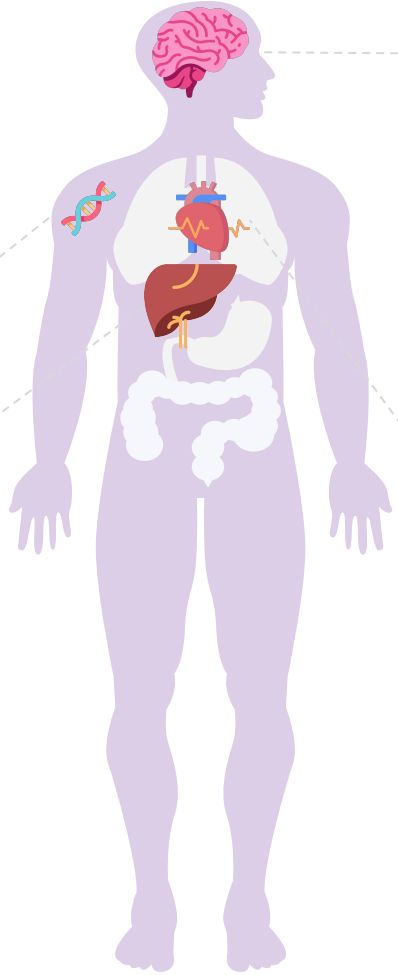
HEPATHER : Cohorte de patients infectés par le virus de l'hépatite B ou C

DREES

EPICOV : Enquête nationale sur l'épidémie du Covid-19

DÉPISTAGE DES CANCERS Centre de coordination Occitanie

Base e-sis : SI sur le dépistage du cancer du sein



Memento : cohorte de patients ayant des signes pouvant évoquer une forme débutante de la maladie d'Alzheimer ou une maladie apparentée

CHU BDX

OSCOUR : Organisation de la surveillance coordonnée des urgences

Santé publique France

MDO : Liste des maladies obligatoires

ars

E-must : Recueil prospectif d'informations concernant les Infarctus Du Myocarde

ESME : Épidémio-Stratégie Médico-Economique du cancer du sein

unicancer

Base ATU : données collectées dans le cadre des accès précoces

Roche

Base liée à l'épidémie de Covid-19

Base principale — l'Assurance Maladie — HEALTH DATA HUB

Catalogue — HEALTH DATA HUB

L’exemple du projet Parcours IDM en IDF : pour mieux comprendre le parcours de la maladie ischémique chronique post-IDM, deux sources de données sont nécessaires

Porteur du projet

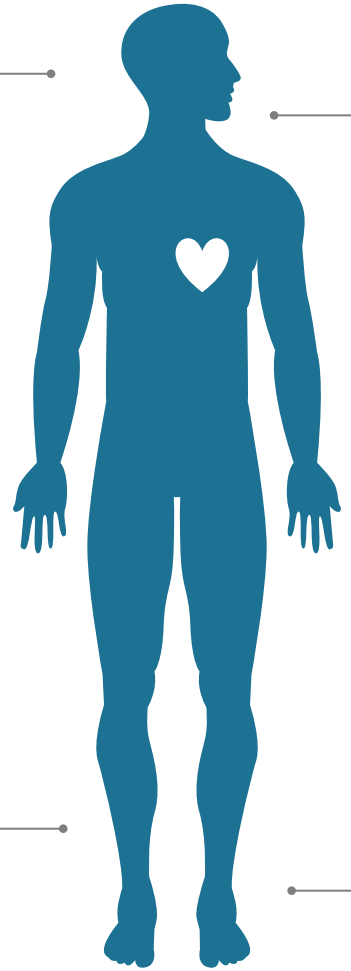
ARS Ile de France

Chiffres clés

- 115 000 IDM par an en France
- 15% : taux de mortalité à 1 an post-IDM aigu

Partenaires

- CH de Versailles
- GIP SESAN
- INSERM
- Asso de patients “Alliance du Coeur”



Mieux connaître le devenir des patients en post-IDM

Décrire le nombre, la fréquence et le délai de survenue des troubles du rythme en post-IDM

Impact indirect : améliorer la prise en charge des patients

Réduire les complications et les ré-hospitalisations, améliorer la qualité de vie et réduire la mortalité à long terme

Amélioration du dépistage précoce des FDR

Réduction des inégalités de soin

Bases mobilisées et intérêt du croisement

Registre e-must : données cliniques pré-hospitalières et hospitalières

SNDS : données post-hospitalières et pré-hospitalières

Pour passer à l'échelle et massifier les données, il peut être utile de capitaliser sur ces étapes

1

Capitaliser sur la réalisation de l'appariement

La réalisation d'un appariement entre deux sources de données requiert **un travail scientifique de plusieurs mois**

(lorsque l'appariement est direct, il demande l'articulation entre plusieurs acteurs, la préparation des données identifiantes, la formation à l'utilisation d'un soft ...)

2

Capitaliser et réduire les temps de contractualisation

Le partage des données dans la perspective d'un projet donné requiert une **contractualisation qui fait intervenir plusieurs acteurs**, si la base est au catalogue, le **conventionnement se fait entre le dépositaire des données et le HDH** et cadre toutes les démarches futures

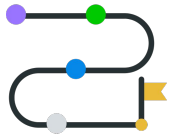
3

Capitaliser sur les travaux techniques et les actions de conformité au RGPD

La transmission des données requiert **leur documentation et la mise en place d'un canal de transfert** conforme aux exigences de sécurité. Ce travail requiert également des ressources
Il en est de même pour les actions relatives à **la mise en oeuvre du RGPD**

La mise à disposition d'une base au catalogue du HDH ne signifie pas que les données sont accessibles en open data

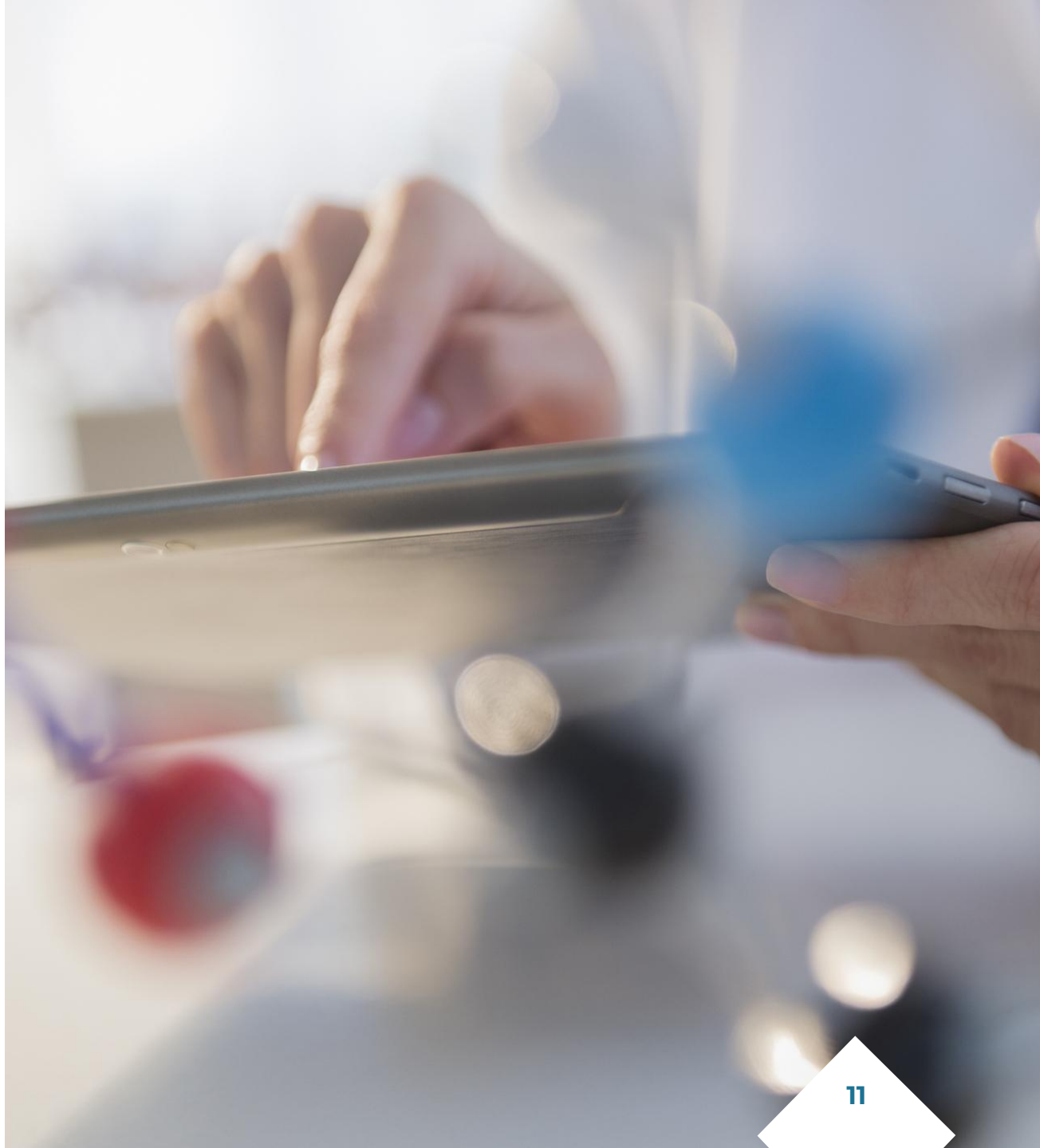
Intérêt pour le responsable de données : la mise au catalogue permet de mettre en visibilité sa base, enrichie par d'autres sources et notamment les données du SNDS historique, et documentée.



Accompagnement du HDH à toutes les étapes : de la mise en oeuvre, de l'analyse de conformité de la base au sens du RGPD, à l'effort de documentation et de mise en place d'un canal de transmission (appui financier possible), jusqu'à la fourniture d'un DOI.

Modalités d'accès aux données dans le catalogue : qu'elles soient ou non au catalogue, les règles d'accès qui s'appliquent sont issues de la réglementation française, mais le HDH s'engage à assurer l'articulation entre le conseil scientifique de la base et le CESREES, à mettre en relation les utilisateurs intéressés et à informer le responsable de données des avancées des projets. Les modalités de valorisation scientifiques et économiques sont convenues dans la convention de partenariat nouée entre le HDH et le responsable de données.

La mise au catalogue de la BNDMR



La mise en oeuvre pratique de la mise au catalogue de la BNDMR

La BNDMR figure sur l'arrêté du 12 mai 2022 portant la liste des bases du catalogue

Signature de la convention de partenariat

- La convention précisant les engagements du HDH et de la BNDMR sera signée début février 2023.

Préparation des données

- Le HDH met à disposition de la BNDMR dès le mois de février une data engineer pour la mise en qualité des données, la pseudonymisation, la documentation, la production des métadonnées et la préparation des données d'appariement.

Exposition de la base

- Homogénéisation du catalogue de métadonnées sur l'ensemble du catalogue du HDH
- Visibilité et communication autour de la base pour encourager la réutilisation des données
- Point de contact avec le HDH pour la gouvernance d'accès aux données

Valorisation économique et scientifique

- Valorisation économique : pas de mise à disposition des données tant que le chantier tarification n'est pas finalisé.
- Application des règles internationales de signature de l'ICMJE. Citation et remerciements obligatoires. Mise en place d'un DOI.

Gouvernance et suivi du partenariat

- Comité de pilotage BNDMR/HDH : suivi du partenariat et de l'exécution de la convention (toutes les 6 semaines puis tous les 6 mois après duplication des données).
- Comité scientifique de la BNDMR : émet un avis consultatif sur les protocoles déposés au CESREES demandant un accès aux données de la BNDMR via le catalogue.

Pré-requis restants avant l'ingestion

- Mise à jour de référentiel de sécurité du SNDS par voie d'arrêté
- Autorisation CNIL pour stockage pérenne des bases de données

Un chaînage pérenne avec le SNDS pour de nouvelles opportunités de recherche dans le domaine des maladies rares

L'appariement des données de la **BNDMR** et du **SNDS** présente de nombreux avantages pour la recherche et le soin, dans un objectif d'**optimisation des schémas de prise en charge**.

Reconstruction des parcours de soins

Meilleure caractérisation des syndrômes polymalformatifs

Mise en évidence d'effets pharmacologiques fortuits

Avez-vous des questions ?

HEALTH
DATA HUB

Suivez-nous sur les réseaux sociaux !