

Chapitre XI

Le système national des données de santé : un vaste gisement d'informations à mieux exploiter

PRÉSENTATION

Créé par la loi en 2016, le système national des données de santé (SNDS) est un ensemble de bases de données issues des activités de remboursement et de tarification de l'assurance maladie, de données médico-économiques relatives aux séjours hospitaliers et d'informations sur les causes médicales de décès. Son périmètre a été étendu en 2019 pour intégrer davantage de données de nature clinique. Un organisme public, la Plateforme des données de santé, également dénommé « Health data hub » dans ses communications à rayonnement international, a été créé afin de mettre ces données à disposition des acteurs économiques et de la recherche.

Le SNDS inclut notamment le système national d'information interrégimes de l'assurance maladie (Sniiram), que la Cour avait étudié en 2016. Elle avait alors relevé que les données personnelles de santé, utiles pour mieux évaluer l'efficacité des soins, pour contrôler certaines activités et pour améliorer l'efficacité du système de santé, étaient insuffisamment utilisées. Elle avait critiqué les freins mis à l'accès à ces informations, au détriment de la recherche et de l'innovation.

La Cour a examiné, huit ans après ce rapport, dans quelle mesure les données du SNDS sont mieux exploitées. Elle constate qu'en dépit de fortes ambitions, les réalisations demeurent incomplètes : l'élargissement du contenu du SNDS est en pratique limité et les délais de mise à disposition des données restent anormalement longs, ce qui freine l'essor des usages (I). Un nouvel élan est aujourd'hui nécessaire pour lever un blocage institutionnel sur la nationalité de l'entreprise qui doit héberger ces données, qui perdure depuis plusieurs années. Pour résoudre ce blocage, et en attendant une solution de long terme, compatible avec le droit européen, des mesures palliatives doivent être décidées à court et à moyen terme (II).

Chiffres-clés

11 sources de données dans la base principale du SNDS et 10 dans un ensemble de bases complémentaires, dénommé « catalogue ».

192 extractions de données accompagnées et 266 projets déposés en 2022 au guichet dont la Plateforme des données de santé assure le secrétariat.

54 % des sollicitations de la caisse nationale de l'assurance maladie (Cnam) pour une mise à disposition de données du SNDS entre 2019 et septembre 2023 étaient issues de l'industrie pharmaceutique et de sociétés commerciales.

Fin 2023, 110 demandes pendantes de mise à disposition de données, autorisées par la Commission nationale de l'informatique et des libertés (Cnil) avant le 31 décembre 2022, avec un délai moyen d'attente de 29 mois depuis l'autorisation.

Comptes de la Plateforme des données de santé en 2023 : 9,65 M€ de subventions et des disponibilités bancaires (29 M€) très supérieures aux charges d'exploitation de l'année (19,9 M€).

I - Des réalisations toujours modestes au regard des ambitions

La création du SNDS par la [loi n° 2016-41 du 26 janvier 2016](#) a été principalement motivée par la recherche d'une ouverture de l'accès aux données personnelles de santé. La stratégie adoptée par les pouvoirs publics a consisté à élargir les finalités susceptibles d'être poursuivies par les porteurs de projet et à enrichir les bases de données concernées. Parallèlement, les modalités d'accès aux bases de données ont été aménagées.

Ces ambitions ont toutefois été contrariées par l'insuffisance de l'enrichissement du contenu disponible du SNDS. Les usages ont globalement augmenté mais le potentiel d'utilisation des données est resté sous-exploité, en particulier par les acteurs privés.

A - Une stratégie d'ouverture portée par les potentialités novatrices de l'intelligence artificielle

Les pouvoirs publics ont étendu les cas d'usages autorisés des données de santé, notamment au regard des nouvelles perspectives offertes par l'essor des techniques d'intelligence artificielle. Le nouveau cadre juridique posé par les lois de 2016 et de 2019 vise à favoriser ces usages au service de finalités enrichies.

1 - Une ambition forte au service de finalités enrichies

Le système national d'information interrégimes de l'assurance maladie (Sniiram), géré par la Cnam enregistre les données de liquidation des bénéficiaires de l'assurance maladie obligatoire. Il était initialement conçu pour répondre aux besoins des différents régimes et caisses de sécurité sociale.

Le SNDS, qui englobe désormais le Sniiram, a considérablement enrichi ces finalités. Il vise à contribuer à la recherche et à l'innovation, à mieux informer sur la santé, sur l'offre de soins et sur la prise en charge médico-sociale et par là-même à servir à la définition, à la mise en œuvre et à l'évaluation des politiques de santé et de protection sociale. Par exemple, l'Agence nationale de santé publique utilise le SNDS à des fins de surveillance et de veille sanitaires, pour le suivi des maladies à déclaration obligatoire.

Ces objectifs pluriels reflètent les potentialités multiples des données personnelles de santé, dont l'utilisation peut concourir à améliorer la qualité des prises en charge et l'efficacité du système de soins en optimisant les parcours de soins.

Le SNDS est, en outre, porteur d'enjeux significatifs pour les acteurs économiques, ce qui justifie une régulation exigeante de l'accès, de manière à préserver la confidentialité des données.

Afin de concrétiser ces potentialités, la France s'est dotée d'une stratégie d'accélération en santé numérique dans le cadre du plan d'investissement « France 2030 », d'une feuille de route du numérique en santé 2023-2027 portée par la délégation au numérique en santé, et d'une stratégie d'accélération informatique en nuage dite « cloud⁴⁸⁶ au centre ».

⁴⁸⁶ L'informatique en nuage (*cloud computing* en anglais) fait référence à l'utilisation de la mémoire et des capacités de calcul des ordinateurs et des serveurs et liés par un réseau. Les données ne sont plus sur un ordinateur mais sur des serveurs distants interconnectés.

Bien que le programme d'ensemble ait ainsi été constitué, une stratégie d'usage secondaire de données originellement constituées pour la prise en charge des patients reste à définir et à mettre en œuvre par les pouvoirs publics. Des travaux de réflexion ont été conduits en 2023 à cet effet, dans le cadre plus large de la stratégie européenne pour les données, définie par la Commission européenne en 2020. Un règlement du 30 mai 2022 sur la gouvernance des données et un règlement sur l'espace européen des données de santé, qui devrait être adopté en 2024, visent à créer un marché unique des données et des services de santé numérique au sein de l'Union européenne, à des fins de recherche, d'innovation, d'élaboration de politiques publiques et de développement d'une médecine personnalisée.

2 - Des potentialités multipliées par l'intelligence artificielle

L'intelligence artificielle offre des perspectives nouvelles à l'utilisation du SNDS. Elle peut aider à concevoir des actions de prévention ou de soins plus personnalisées, et à ouvrir la possibilité d'analyses prédictives de la demande de soins à des fins d'amélioration du pilotage de l'offre.

Ces potentialités ont été mises en exergue dans un rapport de 2018 du député Cédric Villani⁴⁸⁷, sur la base duquel une stratégie pour l'intelligence artificielle 2018-2025 a été adoptée. Ce rapport a rappelé la rapidité des progrès en matière d'apprentissage automatique (*machine learning*)⁴⁸⁸ et profond (*deep learning*)⁴⁸⁹, tributaires, pour entraîner les algorithmes, de larges jeux de données structurées que le SNDS a vocation à fournir.

L'étude d'impact de la loi n° 2019-774 du 24 juillet 2019 a fait état de ces perspectives pour justifier une extension du périmètre du SNDS. Le développement des techniques d'intelligence artificielle appliquées aux données de santé devra toutefois tenir compte du cadre réglementaire européen en cours d'adoption⁴⁹⁰.

En pratique, les perspectives de l'intelligence artificielle sont prometteuses pour les hôpitaux compte tenu de la richesse des données de santé dont ils disposent. Ces dernières devraient être mises en commun,

⁴⁸⁷ Cédric Villani, *Donner du sens à l'intelligence artificielle : pour une stratégie nationale et européenne*, 2018.

⁴⁸⁸ Processus par lequel un algorithme évalue et améliore ses performances sans l'intervention d'un programmeur, en répétant son exécution sur des jeux de données jusqu'à obtenir, de manière régulière, des résultats pertinents.

⁴⁸⁹ Modalité d'apprentissage automatique qui utilise un réseau de neurones artificiels composé d'un grand nombre de couches dont chacune correspond à un niveau croissant de complexité dans le traitement et l'interprétation des données.

⁴⁹⁰ Ce cadre européen pourra inclure une régulation de l'intelligence artificielle, avec une possible entrée en vigueur en 2026.

sous un format homogène, dans des entrepôts de données de santé, à des fins de pilotage, de recherche et d'amélioration des soins.

Des projets ambitieux conciliant l'utilisation du SNDS et les techniques d'intelligence artificielle sont en cours, comme le projet *Deep.Piste* qui vise à améliorer la prise en charge du cancer du sein en développant une analyse automatique des mammographies et des facteurs de risque. L'analyse des examens d'imagerie médicale combinée aux données du SNDS permettra de distinguer les patientes nécessitant un suivi médical renforcé et celles pouvant bénéficier d'un parcours allégé. Le projet *Apsoren* a pour objectif la modélisation du parcours de soins des patients traumatisés crâniens afin de repérer les risques de mauvaise prise en charge. Enfin, le projet *Evired* cherche à préserver la vision des patients diabétiques grâce au caractère prédictif d'un système de diagnostic et d'aide à la décision.

B - Des mesures encourageantes pour développer les usages des données de santé

La mise en place du SNDS s'est traduite par un assouplissement de l'accès aux données et une extension ambitieuse de son périmètre juridique. Les mesures prises ont favorisé un premier développement des usages des données, en particulier par les administrations.

1 - Un assouplissement limité de l'accès aux données

Pour les recherches interventionnelles n'impliquant pas la personne humaine⁴⁹¹, l'accès aux données du SNDS suppose, selon la procédure applicable aux données de santé, une autorisation de la Cnil, donnée après obtention d'un avis d'un comité éthique et scientifique⁴⁹², en pratique toujours suivi, sauf si le traitement demandé est conforme à un référentiel défini préalablement par la Cnil.

Afin de préserver leur confidentialité, les données ne peuvent être mises à disposition qu'en raison d'un motif d'intérêt public. Le SNDS ne

⁴⁹¹ Une autre procédure est prévue pour les recherches interventionnelles impliquant la personne humaine. L'article L. 1121-1 du code de la santé publique prévoit que ces dernières sont organisées et pratiquées sur l'être humain en vue du développement des connaissances biologiques ou médicales.

⁴⁹² L'avis du comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé porte sur la méthodologie scientifique du projet et sur le recours à des données à caractère personnel.

comporte pas de données nominatives⁴⁹³ et est conçu pour prévenir les risques de réidentification des personnes.

Le recours à ces données étant jugé insuffisant, les lois du 26 janvier 2016 et du 24 juillet 2019 ont assoupli les conditions pour y accéder et les exploiter.

Des accès permanents ont été accordés par la réglementation à vingt-cinq, puis, depuis juillet 2021, à trente organismes ou catégories d'organismes chargés d'une mission de service public. Cette évolution a permis un essor des usages des entités concernées. Toutefois, sept organismes ou catégories d'organismes⁴⁹⁴ n'utilisent pas ou quasiment pas ces accès, ce qui pose les questions de leur pertinence et de la nécessité d'une revue périodique des habilitations. Près de 2 500 personnes étaient habilitées à accéder au SNDS en 2023, dont 250 n'en avaient jamais fait usage.

Pour les organismes ne disposant pas de ces accès permanents, la Plateforme des données de santé a été constituée en 2019 sous forme de groupement d'intérêt public⁴⁹⁵ afin de faciliter leur accès aux données. Elle doit accompagner les organismes demandeurs, dits « porteurs de projets » dans la formulation de leur demande et, après autorisation par la Cnil, procéder aux opérations de traitement et d'extraction des données utiles issues du SNDS. Selon le plan « Innovation santé 2030 », elle a plus largement pour objet d'accélérer la recherche et l'innovation reposant sur les données de santé et elle vise à promouvoir l'utilisation de l'intelligence artificielle. Elle a accompagné 192 extractions de données et traité 266 projets déposés en 2022.

Par ailleurs, des procédures d'accès simplifié aux données ont été définies et certaines d'entre elles sont couramment utilisées.

En premier lieu, il n'est pas nécessaire de demander d'autorisation de la Cnil pour accéder aux bases de données de données agrégées relatives au suivi des dépenses et à l'analyse de l'offre de soins, à des tableaux de bord sur la biologie et sur la pharmacie, ainsi qu'à un échantillon du Sniiram d'informations sur les caractéristiques sociodémographiques et médicales des bénéficiaires de soins et leurs prestations perçues⁴⁹⁶.

⁴⁹³ Le SNDS ne contient ainsi pas les noms, prénoms, adresses et numéros d'inscription au répertoire national d'identification des personnes physiques.

⁴⁹⁴ Unions régionales des professionnelles de santé, Établissement français du sang, Autorité de sûreté nucléaire, Haut Conseil pour l'avenir de l'assurance maladie, Agence nationale d'appui à la performance des établissements de santé et médico-sociaux, Fonds de financement de la couverture maladie universelle, centres de lutte contre le cancer (à l'exception de deux d'entre eux), Observatoire français des drogues et tendances addictives (recrutement en février 2024 d'un agent susceptible d'utiliser le SNDS).

⁴⁹⁵ Il s'est substitué à l'Institut national des données de santé créé en 2016.

⁴⁹⁶ En application de la délibération n° 2020-072 du 16 juillet 2020 de la Cnil.

En deuxième lieu, de nouvelles méthodologies de référence, procédures simplifiées d'accès aux données de santé, ont été conçues pour dispenser les porteurs de projets d'une autorisation de la Cnil. En raison de leur souplesse, elles ont représenté près de la moitié des demandes d'accès aux bases du SNDS répertoriées par la Plateforme des données de santé entre 2017 et juin 2023.

En troisième lieu, une procédure simplifiée encadre les projets d'un même responsable de traitement qui s'inscrivent dans le cadre d'une décision unique de la Cnil pour des traitements répondant à une même finalité, portant sur des catégories de données identiques et ayant des catégories de destinataires identiques. Dans ces conditions strictes, seule une cinquantaine de décisions uniques a été mise en œuvre entre 2019 et 2023.

Enfin, un décret en Conseil d'État était auparavant nécessaire pour procéder à des appariements avec des données utilisant le numéro d'inscription au répertoire national d'identification des personnes physiques. La procédure d'autorisation de la Cnil, moins contraignante, l'a remplacé.

2 - Une extension juridique volontariste du contenu du SNDS

Le SNDS a été créé avec, pour composantes initiales, les données de liquidation des bénéficiaires de l'assurance maladie obligatoire du Sniiram, les données médico-économiques d'hospitalisation des patients du programme médicalisé des systèmes d'information, dont la finalité principale est la tarification des séjours hospitaliers, les données sur les causes médicales de décès, les données médico-sociales du système d'information commun des maisons départementales des personnes handicapées, et un échantillon représentatif des données de remboursement par bénéficiaire transmises par des organismes d'assurance maladie complémentaire.

Ainsi, le SNDS était constitué de données médico-administratives sur le recours aux soins, les hospitalisations, le handicap, les prestations sociales et l'activité professionnelle. Elles couvrent quasi exhaustivement la population et ont un grand degré de finesse, mais elles sont dénuées de finalité épidémiologique.

À la lumière de travaux, dont ceux de la Cour⁴⁹⁷, la loi du 24 juillet 2019 a élargi le périmètre du SNDS en distinguant, désormais, une base principale, dont les données couvrent l'ensemble de la population,

⁴⁹⁷ Cour des comptes, Les données personnelles gérées par l'assurance maladie, communication à l'Assemblée nationale, 2016.

et, sous le terme de catalogue, un ensemble de bases de données couvrant des champs plus restreints. Par le biais d'appariements avec d'autres bases contenant des informations complémentaires, l'alimentation du catalogue doit enrichir le SNDS avec des données à contenu clinique.



La législation prévoit l'inclusion dans la base principale des données destinées aux professionnels et aux organismes de santé recueillies à l'occasion de leurs activités. Même si cette extension ne s'est pas encore concrétisée, le SNDS permet déjà de publier des statistiques et indicateurs sur les coûts et la qualité des prises en charge et du système de soins, et d'alimenter ainsi le débat public et la prise de décision.

Avec ce contenu élargi, il est possible de réaliser des études de santé publique afin de mieux évaluer l'état de santé de la population, les facteurs de risque ou les inégalités territoriales ou sociales en santé⁴⁹⁸. Des écarts de pratiques non justifiés peuvent être mis en évidence, ce qui favorise la diffusion des meilleures pratiques. Enfin, le monde de la recherche et les industriels peuvent utiliser ces données à des fins de connaissance et d'innovation.

3 - Un recours croissant des organismes publics au SNDS

a) *Un usage régulier par les organismes d'assurance maladie*

En 2016, la Cour avait mis en évidence la sous-utilisation du système national d'informations interrégimes de l'assurance maladie (Sniiram) par les organismes d'assurance maladie et par les autres acteurs publics⁴⁹⁹. Depuis la mise en place du SNDS, la situation a évolué favorablement. Pour ses besoins propres, la Cnam a recours au SNDS pour alimenter certains de ses indicateurs orientés vers la recherche d'économies⁵⁰⁰. Le SNDS a permis de supprimer le décalage de trois mois dans le suivi des dépenses en dates de remboursement de soins du régime général.

En matière de gestion du risque et de lutte contre les abus et la fraude des professionnels de santé⁵⁰¹, les données du SNDS sont utilisées pour cibler les offreurs de soins susceptibles de faire l'objet d'investigations. Au plan national et au plan local, le SNDS permet de détecter des atypies dans les prescriptions de soins, dans les arrêts de travail ou dans les transports sanitaires. Les caisses primaires d'assurance maladie identifient ensuite les professionnels de santé concernés au moyen d'un système d'information distinct, qui comprend des données nominatives.

⁴⁹⁸ La Drees a créé un échantillon permanent qui apparie des données démographiques et des informations issues du SNDS. Croisé avec d'autres données, il permet de mettre en regard des données sanitaires les informations relatives à la vie professionnelle et aux niveaux de vie. Il permet d'étudier les inégalités sociales et territoriales de santé et contribue à l'élaboration de politiques publiques tendant à les réduire.

⁴⁹⁹ Cour des comptes, *Les données personnelles gérées par l'assurance maladie*, communication à l'Assemblée nationale, 2016.

⁵⁰⁰ Par exemple, ces indicateurs ont été relatifs aux économies sur le champ des médicaments dits de spécialités (2019), au recours aux chirurgiens-dentistes pour les classes d'âge concernées par l'examen bucco-dentaire (2021) ou à l'évolution des dépenses standardisées par bénéficiaire (2022).

⁵⁰¹ Cour des comptes, *L'avenir de l'assurance maladie. Assurer l'efficacité des dépenses, responsabiliser les acteurs*, 2017, p. 158.

b) Un développement des utilisations par le ministère chargé de la santé

Au sein du ministère chargé de la santé, la direction de la recherche, des études, de l'évaluation et des statistiques (Drees) a développé une expertise en matière d'utilisation des données de santé. Plus de la moitié des études produites s'appuie sur le SNDS⁵⁰², qui sert les 35 agents habilités de la direction (contre 12 en 2016), pour suivre la consommation de soin, la santé de la population et l'organisation du système de santé.

La direction générale de l'offre de soins dispose de six statisticiens habilités fin 2023 et produit des travaux originaux⁵⁰³ alors que la Cour avait relevé en 2016 qu'elle n'était que consommatrice d'études produites par d'autres directions.

La direction de la sécurité sociale utilise le SNDS pour chiffrer l'effet des mesures introduites dans la législation, en particulier celles des lois de financement de la sécurité sociale, pour la rédaction des rapports de la commission des comptes de la sécurité sociale, ou encore pour le suivi des dépenses d'assurance maladie.

Malgré ce développement des usages, le SNDS reste peu utilisé par le ministère pour concevoir et mettre en œuvre les politiques de santé et de protection sociale. Des indicateurs pourraient être définis à divers niveaux géographiques pour le pilotage et l'évaluation des politiques publiques.

c) Une mobilisation plus fréquente par certaines agences sanitaires

Les agences sanitaires recourent couramment au SNDS. À titre d'exemple, pour le suivi hebdomadaire des indicateurs de santé mentale, Santé publique France utilise la base OSCOUR, inscrite au catalogue et alimentée par plus de 600 services d'urgences en 2023. Cette agence a également créé un observatoire cartographique des indicateurs épidémiologiques⁵⁰⁴, dénommé Géodes, qui repose sur des données issues du SNDS mises à la disposition du grand public.

⁵⁰² La Drees a analysé les causes médicales des décès survenus en 2021 (2023), la survenue des accidents vasculaires cérébraux (2022), leur fréquence, la prise en charge et les séquelles selon le niveau de vie des personnes (2022).

⁵⁰³ Les cas d'usages, variés, peuvent concerner le taux de recours aux urgences, le nombre d'exams biologiques par département, etc.

⁵⁰⁴ Ces indicateurs concernent par exemple les cancers, la covid 19, les valvulopathies ou le suicide.

L'Agence nationale de sécurité du médicament et des produits de santé s'appuie sur le SNDS au titre de sa mission de surveillance sanitaire. En 2018, elle a constitué avec la Cnam un groupement d'intérêt scientifique (EPI-PHARE) qui mène des études de pharmaco-épidémiologie s'appuyant essentiellement sur les données du SNDS. Ses études éclairent les pouvoirs publics sur la balance bénéfice-risque associée à chaque médicament, par exemple en estimant la fréquence des événements indésirables liés à la prise d'un traitement. Fin 2023, EPI-PHARE avait publié une trentaine d'articles dans des revues internationales à comité de lecture.

C - Un potentiel encore sous-exploité

Le SNDS se caractérise par deux lacunes principales. D'une part, nombre de ses composantes restent indisponibles, parfois sans perspectives d'amélioration. D'autre part, les délais de mise à disposition des données auprès des porteurs de projets restent trop élevés, ce qui conduit certains d'entre eux à privilégier des bases de données étrangères.

1 - Un enrichissement du SNDS encore incomplet

Au début de 2024, aucune des nouvelles composantes de la base principale du SNDS prévues par la loi du 24 juillet 2019 n'a été encore rendue disponible⁵⁰⁵. Cette absence de concrétisation peut s'expliquer en partie par la crise sanitaire mais elle est d'autant plus décevante qu'il n'existe pas de perspective quant à leur intégration au SNDS. En particulier, malgré une expérimentation menée au début des années 2010, les données des organismes d'assurance maladie complémentaire sont manquantes et devraient le rester à moyen terme, faute d'un recueil et d'une standardisation dans un format permettant leur intégration dans le SNDS.

⁵⁰⁵ Données relatives à la perte d'autonomie, données recueillies lors des visites médicales et de dépistage obligatoires, données recueillies par les services de protection maternelle et infantile, données recueillies lors des visites de prévention.

Concernant les bases inscrites au catalogue, certaines comportent peu de données et leur intérêt s'en trouve réduit en termes d'appariements avec la base principale. En outre, la disponibilité de ces données est tributaire de l'adoption de l'actualisation du référentiel de sécurité du SNDS qui date de 2017. Cette actualisation n'était toujours pas intervenue en mars 2024 et dans l'attente, les données restent non accessibles aux demandeurs.

En complément des bases d'ores et déjà prévues par les textes, deux catégories d'informations présentent un intérêt important pour le SNDS : d'une part, les registres et cohortes, d'autre part, les entrepôts de données hospitaliers et leur équivalent pour la médecine de ville.

Les premiers constituent des instruments pour le suivi épidémiologique. Certains sont d'ailleurs d'ores-et-déjà inscrits dans le catalogue, comme la cohorte Heparther, consacrée au suivi des patients atteints des hépatites B ou C, ou la cohorte Memento, portant sur les personnes atteintes de troubles de la mémoire. La cohorte Constances, notamment, serait utile à la recherche pour la richesse de ses données. Au Danemark, l'Autorité des données de santé est responsable d'une quarantaine de registres nationaux susceptibles d'être appariées aux données de santé générales, dans le cadre d'une nouvelle stratégie nationale pour une médecine personnalisée.

Les entrepôts de données de santé constituent une initiative plus récente, encouragée en 2022 par le ministère chargé de la santé dans le cadre de la stratégie nationale d'accélération en santé numérique. Un socle commun de données a été retenu en septembre 2023 par le comité stratégique des données de santé pour harmoniser le contenu des entrepôts hospitaliers en vue de les constituer en réseau d'ici à 2025.

L'incorporation des données de ce socle commun dans le catalogue enrichirait les informations cliniques du SNDS, en particulier dans le domaine de l'imagerie médicale ou des prescriptions médicamenteuses liées aux séjours hospitaliers. La mise en commun centralisée de ces données permettrait d'éviter des coûts de coordination dans la gouvernance⁵⁰⁶ et la gestion opérationnelle⁵⁰⁷ d'entrepôts, dont le fonctionnement dans la durée impliquera l'identification d'un mode de financement pérenne pour couvrir les investissements et les dépenses d'exploitation.

⁵⁰⁶ Analyse des demandes par les équipes opérationnelles des conseils scientifiques et éthiques des centres hospitaliers, négociation des conditions de valorisation, contractualisation locale.

⁵⁰⁷ Requêtage locaux supposant la définition de programmes informatiques adaptés à chaque environnement technique pour produire des requêtes similaires et sans biais, agrégation au niveau central avec des contrôles de qualité et de cohérence, réalisation de l'appariement avec la base principale pour chacun des projets.

D'autres types de données pourraient être incluses dans le SNDS, comme la plateforme de données en cancérologie développée par l'Institut national du cancer, les registres de santé, les cohortes en santé ou encore les bulletins de naissance collectés par l'Institut national de la statistique et des études économiques⁵⁰⁸.

Plus largement et dans une perspective plus lointaine, une source d'enrichissement du SNDS en données cliniques résiderait dans le contenu des dossiers médicaux partagés, qui recueillent des éléments diagnostiques et thérapeutiques liés aux actes et consultations, afin de favoriser la prévention, la coordination, la qualité et la continuité des soins⁵⁰⁹. Il en va de même pour le contenu des dossiers pharmaceutiques, en particulier s'agissant des informations sur la dispensation de médicaments.

2 - Des conditions de mise à disposition des données encore complexes et insatisfaisantes

Le maintien d'un régime d'autorisation préalable par la Cnil pour la mise à disposition des données en l'absence de conformité à des référentiels se démarque des choix retenus par d'autres pays européens, moins exigeants. La responsabilité de l'octroi de l'accès repose sur le responsable des bases de données requêtées, souvent après avis d'un comité scientifique. En Finlande ou au Royaume-Uni, un même organisme autorise les projets et met à disposition les données, ce qui raccourcit les délais de traitement.

La Cnil respecte le plus souvent le délai maximal qui lui est imparti de deux mois, renouvelable une fois. Elle rend ses décisions dans un délai moyen de 74 jours, qui peut être beaucoup plus long pour certains projets.

Ce délai s'ajoute à celui, en amont, de l'analyse par le comité éthique et scientifique et, en aval, par les opérations de ciblage et d'extraction des données. La Cnam intervient à ce dernier stade puisque c'est sur elle que repose, encore à ce jour, la plupart des mises à disposition de données. Elle s'attache à mettre en œuvre un principe de minimisation, qui consiste à ne délivrer que les données strictement nécessaires à l'accomplissement du projet approuvé par la Cnil. Ce principe découle du

⁵⁰⁸ Cour des comptes, rapport au Parlement, *La politique de périnatalité, une mobilisation à amplifier*, 2024.

⁵⁰⁹ Cour des comptes, *Mon espace santé : des conditions de réussite encore à réunir*, Ralfss chapitre X, 2024.

règlement général européen sur la protection des données (RGPD)⁵¹⁰, dont l'article 5 prévoit que les données à caractère personnel doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités de traitement.

Néanmoins, et comme l'indique le rapport récent d'une mission interministérielle coordonnée par M. Marchand-Arvier⁵¹¹, une partie des porteurs de projets considère que la Cnam examine de manière trop minutieuse le périmètre et la profondeur historique des données sollicitées, alors que ceux-ci ont déjà été analysés dans le cadre de la procédure d'autorisation par la Cnil. En pratique, cette étape nécessite des échanges souvent longs avant la signature des conventions de mise à disposition des données.

Il revient certes à la Cnam de cibler le champ des données nécessaires à la réalisation des projets. Mais un assouplissement des modalités d'analyse pourrait être envisagé et précisé dans ses procédures internes qui seront actualisées en 2024 afin de réduire le délai de mise à disposition des données.

3 - Une mise à disposition des chercheurs et des bureaux d'études qui reste insuffisante

La loi du 24 juillet 2019 a été guidée par la volonté de développer l'utilisation des données du SNDS pour favoriser la recherche clinique mais aussi pour permettre l'émergence de nouveaux usages, notamment liés au développement des méthodes d'intelligence artificielle. Il devait s'en suivre une augmentation du recours au SNDS par les milieux de la recherche et par les acteurs privés, notamment par les bureaux d'études agissant pour le compte de firmes pharmaceutiques ou de fabricants de dispositifs médicaux.

En dehors des recherches impliquant la personne humaine, telles que les essais cliniques, qui relèvent d'une législation spécifique, les demandes de données du SNDS émanant des travaux des organismes de recherche et des bureaux d'études concernent surtout trois champs :

- les études en vie réelle ou apparentées, comme le suivi des parcours des patients postérieurement aux soins ;

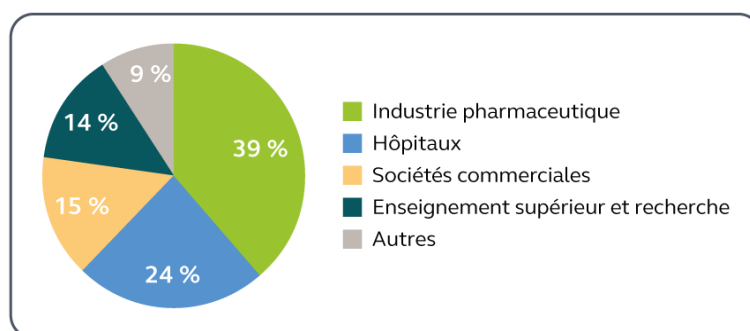
⁵¹⁰ Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016.

⁵¹¹ Jérôme Marchand-Arvier et al., *Fédérer les acteurs de l'écosystème pour libérer l'utilisation secondaire des données de santé*, décembre 2023.

- des études sur des facteurs de risque pour la santé en vue d'améliorer les politiques de prévention, de veille et de surveillance sanitaires ;
- des études qui analysent les dimensions économiques et sociales des problèmes de santé.

Le nombre de nouveaux projets est passé de 37 en 2018 à 165 en 2022. Cette quantification reste néanmoins incomplète, dans la mesure où certaines équipes publiques⁵¹² bénéficient d'un accès permanent aux données. Sous cette réserve, il apparaît que de 2019 à septembre 2023, les bureaux d'études ont représenté plus de la moitié des études, dont 39 % pour les seules industries pharmaceutiques.

Graphique n° 39 : nombre de projets soumis à la Cnam nécessitant la mise à disposition de données du SNDS (2019 à sept. 2023)



Source : Cnam

Or le nombre de demandes de la part des acteurs privés, de l'ordre d'une centaine par an entre 2018 et 2020, a diminué de manière importante jusqu'en 2022 (54 demandes) avant de revenir en 2023 à près d'une centaine. Parallèlement, le nombre de projets de recherche publique reste loin en 2023 du niveau constaté avant la crise sanitaire (71 projets contre 123 en 2019). Si la crise sanitaire a conduit à une réduction du nombre de projets, il demeure que l'augmentation attendue du recours au SNDS pour la production d'études et de recherches, notamment issues des acteurs privés, ne s'est pas, à ce jour, matérialisée.

⁵¹² CHU, centres de lutte contre le cancer, Centre national de la recherche scientifique, Institut national de recherche en informatique et en automatique, Institut national d'études démographiques, etc.

II - Un nouvel élan nécessaire pour permettre l'essor des usages du SNDS

Outre les différents facteurs déjà décrits qui ont freiné le développement des usages du SNDS, le fonctionnement du dispositif mis en place par la loi du 24 juillet 2019 a buté depuis cinq ans sur l'impossibilité pour la Plateforme des données de santé de disposer d'une copie des données du SNDS.

En conséquence, la Plateforme des données de santé n'a assuré que très partiellement le rôle que les textes lui ont confié, reportant sur la Cnam une importante charge de travail. Alors que les pouvoirs s'attachent à définir une stratégie pour développer l'utilisation des données de santé au premier semestre 2024, des mesures doivent être prises pour affiner les attendus de la solution cible de long terme, mais aussi pour remédier à ce blocage à court et à moyen terme.

A - Un blocage institutionnel persistant

Dès sa création, la Plateforme des données de santé a utilisé une solution technologique d'hébergement des données qui a suscité des réserves de la Cnil motivées par des risques d'atteinte à la confidentialité des données. Dans ce contexte, l'autorisation de la Cnil pour lui permettre de disposer de la copie de la base principale du SNDS a été sollicitée puis retirée. Il en a résulté une remise en cause du schéma institutionnel envisagé par les pouvoirs publics, qui se traduit en particulier par des délais excessifs de mise à disposition des données aux porteurs de projets.

1 - Un obstacle important à la remise des données du SNDS à la Plateforme des données de santé

En 2019, le ministère chargé de la santé a choisi de recourir à la solution technologique proposée par la filiale européenne d'un groupe américain pour héberger la copie de la base principale des données du SNDS⁵¹³ et assurer les fonctionnalités en permettant le traitement. Ce choix a été motivé notamment par le fait qu'aucune entreprise de l'Union européenne n'était en mesure de répondre techniquement aux besoins requis pour mettre en œuvre les fonctionnalités attendues d'apprentissage automatique pour les usages nécessitant le recours à l'intelligence artificielle.

⁵¹³ L'article R. 1461-3 du code de la santé publique prévoit que la Plateforme des données de santé détient une copie des bases du SNDS.

Or, le contexte juridique concernant la question des transferts de données a évolué en parallèle dans le cadre du RGPD, entré en vigueur en mai 2018, ce qui a conduit les autorités de contrôle de l'Union européenne⁵¹⁴ à renforcer leur vigilance à l'égard du type de solution retenue par la Plateforme des données de santé.

Un contexte juridique évolutif concernant les transferts de données en dehors de l'Union européenne

Au sens du RGPD, constituent des transferts de données toute communication, copie ou déplacement de données personnelles ayant vocation à être traitées dans un pays tiers à l'Union européenne. Ces opérations, issues de l'utilisation des services en ligne, sont encadrées par plusieurs dispositifs, de manière à garantir aux habitants de l'Union un niveau de protection adéquat de leurs données.

La décision d'adéquation constitue l'un de ces dispositifs. Elle consiste, pour la Commission européenne, à constater que le niveau de protection des données dans un pays extérieur à l'Espace économique européen est adéquat et que des transferts de données peuvent être effectués vers cet État.

Compte tenu de l'importance des États-Unis dans la mise à disposition de services numériques en ligne, la décision d'adéquation les concernant a présenté un enjeu particulier. Or, la Cour de justice de l'Union européenne a annulé en 2020 la décision du 12 juillet 2016 par laquelle la Commission européenne avait constaté que les États-Unis avaient apporté les garanties nécessaires. La Cour de justice a relevé que la réglementation américaine permettait l'accès, par les services de renseignement aux données traitées par tout opérateur américain, qu'elles soient stockées aux États-Unis ou non, remettant en cause le dispositif de protection des données sur laquelle la décision d'adéquation était fondée. Il en est résulté une impossibilité juridique de transferts de données vers les États-Unis.

À la suite d'une évolution de la réglementation américaine (Executive Order 14086 du 7 octobre 2022), une nouvelle décision d'adéquation a été prise par la Commission européenne en juillet 2023⁵¹⁵ : la collecte de données n'est dorénavant possible qu'à des fins de protection de la sécurité nationale et dans le respect du principe de proportionnalité ; les citoyens européens ont désormais accès à un mécanisme de recours auprès d'une Cour de contrôle de la protection des données. Depuis lors, les transferts de données en direction des États-Unis sont redevenus licites mais plusieurs recours contentieux ont été exercés à l'encontre de cette décision.

⁵¹⁴ Soit la Cnil et ses homologues des autres États membres de l'Union européenne dans le cadre du Comité européen de protection des données.

⁵¹⁵ Décision d'exécution 2023/1795 de la Commission européenne du 10 juillet 2023.

La Cnil considère que la décision d'adéquation de juillet 2023 qui autorise à nouveau les transferts de données vers les États-Unis ne traite pas la question des données hébergées par la Plateforme des données de santé au moyen de sa solution technologique actuelle. Selon elle, la question n'est pas tant celle des transferts que de l'accès par les autorités d'un pays tiers à des données hébergées sur le territoire de l'Union européenne. Or, celui-ci est toujours possible aux termes du contrat qui lie la Plateforme des données de santé avec son prestataire. La Plateforme conteste le bien-fondé de cette distinction entre transfert et accès, et estime que la décision d'adéquation donne les garanties nécessaires à une copie de la base principale du SNDS. Pour autant, elle n'est pas en mesure de l'obtenir en l'absence d'autorisation de la Cnil.

En conséquence, hormis pour les projets de recherches qu'elle accompagne, la Plateforme des données de santé n'est pas à même, depuis 2019, de mettre à disposition des porteurs de projets les données qu'ils sollicitent. Seule la Cnam peut répondre à leurs besoins à partir de la base de données principale, ce qui ne correspond pas au schéma institutionnel légal et qui s'accompagne de délais excessifs.

Enfin, l'impossibilité dans laquelle se trouve la Plateforme des données de santé de recevoir une copie du SNDS contribue à entraver l'extension de ses bases prévue par la législation, de manière à y intégrer de nouvelles composantes et à les apparier avec les composantes existantes.

2 - Des freins à lever à la mise à disposition rapide des données

Les conditions dans lesquelles la Cnam répond actuellement aux demandes d'études et de recherches ne sont pas satisfaisantes. En lien avec la Plateforme des données de santé, elle a certes engagé de réels efforts. Mais elle éprouve des difficultés pour assurer cette mission, qui devrait relever de la Plateforme des données de santé et qui entraîne pour elle une lourde charge administrative.

La Cnam ne procède pas à un suivi précis et régulier des demandes entrantes et des délais de traitement des demandes de mise à disposition de données du SNDS. Ces derniers ont atteint, en moyenne, 228 jours pour la recherche privée et 284 jours pour la recherche publique entre août 2017 et juillet 2023⁵¹⁶. La situation récente est encore plus alarmante : en décembre 2023, un total de 110 demandes pour lesquelles l'autorisation de la Cnil

⁵¹⁶ Source : Plateforme des données de santé. Un rapport d'information de la commission des affaires sociales du Sénat sur les données de santé (juillet 2023) estime le délai de traitement au niveau de la Cnam entre 10 et 12 mois.

avait été donnée en 2022 ou antérieurement restait à honorer, avec un délai supérieur à deux ans et demi depuis la date de soumission des projets.

Selon la Cnam, cette situation trouve sa source dans la complexification des demandes qui lui sont adressées, dans les délais de réponse des demandeurs pour l'expression de leurs besoins et dans la capacité insuffisante de ses systèmes d'information pour assurer rapidement les traitements nécessaires sans dégrader leur qualité de service.

De tels délais compromettent l'aboutissement des projets de recherche, notamment ceux effectués dans le cadre de la préparation de thèses ou pour répondre à des études commandées par les industriels de santé. En Finlande, l'accès aux données personnelles de santé varie de six à onze mois et en Allemagne, le centre de données de recherche a prévu de mettre à disposition des porteurs de projets les données qu'ils sollicitent dans un délai de trois mois. Au Royaume-Uni, le délai moyen entre le dépôt de demande d'accès et l'extraction des données par UK Biobank⁵¹⁷ était inférieur à quatre mois en 2022.

Il résulte de cette situation que l'utilisation des données du SNDS reste nettement inférieure à ce que son riche potentiel permettrait d'envisager. Même s'il n'est pas possible de quantifier ce phénomène, des équipes de recherche en France recourent à des bases étrangères en raison de leur plus grande facilité d'accès (UK Biobank, Findata).

Accroître l'utilisation du SNDS, conformément à ce que prévoient les pouvoirs publics, implique donc non seulement qu'il comporte davantage de données cliniques mais aussi que les délais de mises à disposition des données soient mieux maîtrisés. La situation présente est d'autant plus insatisfaisante que la Plateforme des données de santé a été précisément constituée à cette fin.

3 - Une absence durable de décision sur la question de l'hébergement des données

Les sujets relatifs aux données personnelles de santé font intervenir une pluralité d'acteurs publics qui se réunissent au sein d'un comité stratégique.

⁵¹⁷ UK Biobank est une cohorte permettant de mettre à disposition des données comprenant le génotypage de plus de 500 000 volontaires, le séquençage complet de l'exome de plus de 400 000 Britanniques et le séquençage complet du génome de plus de 200 000 volontaires.

La gouvernance des données de santé

Au sein du ministère de la santé, la Drees « est chargée de la gouvernance des données de santé et de la définition des règles de leur mise à disposition »⁵¹⁸. De surcroît, une délégation du numérique en santé a été constituée au sein de l'administration centrale pour animer la gouvernance du numérique en santé⁵¹⁹. Outre le ministère chargé de la santé, la gouvernance du SNDS fait intervenir la Cnam et la Plateforme des données de santé, dont les statuts ont ceci d'atypique que l'État ne dispose que de 16,5 % des voix au conseil d'administration.

Pour améliorer le pilotage des questions relatives aux données de santé, un comité stratégique a été constitué en 2021. Sa mise en place s'est avérée tardive puisque la création de cette instance avait été prévue en 2017, même si la gestion de la crise sanitaire explique en partie ce retard.

Ce comité stratégique, à la composition très large (25 membres)⁵²⁰, a mis en place plusieurs groupes de travail, notamment sur le financement des entrepôts de données, la tarification de la mise à disposition des données, la définition d'un socle de données pour les entrepôts de données de santé hospitaliers, l'identification de bases susceptibles d'intégrer le catalogue du SNDS et la gouvernance de l'accès aux données de santé.

Dans un contexte où la priorité a été accordée à l'hébergement des données de santé par des solutions techniques relevant exclusivement du droit de l'Union européenne, aucune décision publique n'a été prise pour résoudre la difficulté créée par le choix d'un hébergeur soumis à des lois extra-européennes selon un horizon compatible avec les attentes des porteurs de projets.

Le ministère du travail, de la santé et des solidarités a prévu de publier d'ici à la fin du premier semestre de 2024 une feuille de route sur l'usage secondaire des données de santé prenant en compte les travaux du comité stratégique, notamment en matière de gouvernance et de tarification, et les conclusions du rapport de la mission confiée à M. Marchand-Arvier. Il apparaît aujourd'hui indispensable que des décisions soient prises au plus tôt et intègrent différentes échéances temporelles.

⁵¹⁸ Article 8 du décret n° 2000-685 du 21 juillet 2000.

⁵¹⁹ Article 2 du décret n° 2019-1412 du 20 décembre 2019 portant diverses dispositions relatives à l'administration centrale des ministères chargés des affaires sociales.

⁵²⁰ Placé sous la présidence du ministre ou du directeur de la Drees par délégation, il inclut la Plateforme des données de santé, la Cnam, le ministère chargé de la recherche, des institutions utilisatrices du champ de la santé et de la recherche, ainsi que des représentants des hôpitaux.

B - Un nécessaire horizon de long-terme et des mesures palliatives à court et moyen terme

Comme l'a rappelé le rapport de la mission Marchand-Arvier, remédier à la situation de blocage est un préalable à l'engagement d'une stratégie visant à développer l'usage des données de santé.

Il ressort des propositions issues de ce rapport qu'il y a lieu de fixer des perspectives de long terme pour faire émerger une solution répondant aux attentes technologiques les plus avancées tout en relevant exclusivement du droit de l'Union européenne. Il conviendra aussi d'adopter des mesures pour faire évoluer les conditions de mise à disposition des données à moyen et améliorer la qualité de service à court terme.

1 - Des perspectives à tracer pour le long terme

Les études menées par le ministère chargé de la santé aboutissent au constat selon lequel aucun prestataire n'est actuellement à même à la fois d'apporter les garanties nécessaires à l'égard des risques d'accès par des États extérieurs à l'Union européenne et de proposer les fonctionnalités les plus avancées en matière d'intelligence artificielle. De son côté, le rapport de la mission Marchand-Arvier estime qu'une telle offre pourra être fonctionnelle d'ici 2026, ce qui apparaît optimiste car elle est tributaire de paramètres industriels complexes. Les ambitions initiales de la loi du 24 juillet 2019 ne pourront donc être mises en œuvre, au mieux, que sept ans après sa promulgation.

Au plus tard à cet horizon, il faudra également mettre en place un dispositif de redevances pour les utilisateurs privés des données de santé afin de couvrir tout ou partie du coût de mise à disposition⁵²¹. La Cour des comptes avait déjà appelé de ses vœux en 2016 la tarification de l'usage des bases de données de santé. Cette tarification contribuera, en outre, au financement pérenne des entrepôts de données de santé hospitaliers.

Une réflexion, conduite par le comité stratégique, a porté sur l'identification d'un outil d'estimation des coûts, qui constitue un prérequis au calcul des redevances. Les réflexions en cours pourraient s'inspirer des modalités mises en place par le centre d'accès sécurisé aux données (CASD).

⁵²¹ Bénéficient aujourd'hui d'une gratuité de droit les traitements de données concernant la santé demandés par l'autorité publique et les recherches réalisées exclusivement pour les besoins de services publics administratifs. Les autres recours au SNDS devraient être, en principe, payants.

Le Centre d'accès sécurisé aux données

Le CASD est un groupement d'intérêt public créé par arrêté du 29 décembre 2018 rassemblant l'État, représenté par l'Institut national de la statistique et des études économiques, le Groupe des écoles nationales d'économie et statistique, le Centre national de la recherche statistique, l'école polytechnique, HEC Paris et la Banque de France.

Il agit en tant que tiers de confiance en mettant à disposition un équipement conçu pour permettre aux utilisateurs de travailler sur des données fournies par des producteurs dans des conditions de sécurité élevées. Après avoir signé un contrat avec le CASD, le producteur transmet une copie de ses données afin qu'elles soient stockées dans l'infrastructure sécurisée du CASD en France, dite « bulle sécurisée ». À la différence de ce qui est proposé par les autres opérateurs « cloud », les données sont stockées dans un lieu unique et identifié. Le producteur des données en reste le propriétaire et définit leurs conditions d'utilisation et les personnes habilitées à y accéder.

Le CASD facture une partie de ses coûts de fonctionnement sous forme d'abonnement, dont le montant dépend notamment de la puissance de calcul. Dans le domaine de la santé, le CASD permet déjà l'accès aux données des séjours hospitaliers publics et privés en France (données du programme médicalisé des systèmes d'information), ainsi qu'à certaines cohortes.

2 - Une solution intermédiaire à identifier pour décharger la Cnam des tâches de mise à disposition des données

Compte tenu des délais actuels de mise à disposition des données par la Cnam, celle-ci n'est pas en mesure de répondre à une demande croissante de projets. Parallèlement à la concrétisation de la solution de long terme, des actions alternatives doivent être mises en œuvre. Plusieurs pistes sont envisageables.

La première, privilégiée par la mission Marchand-Arvier, consiste à identifier un acteur relevant exclusivement du droit de l'Union européenne européen à même de traiter les demandes d'accès aux données dans les conditions de sécurité requises, en renonçant temporairement aux fonctionnalités technologiques avancées liées à l'intelligence artificielle. La mission préconise de procéder à un audit flash du CASD. Si les conclusions de cet audit sont favorables, elle envisage une dévolution dans les six mois. Les délais pourraient être toutefois plus importants, compte tenu de la procédure d'autorisation de la Cnil et de l'opération technique de transfert de la copie de la base principale du SNDS, et surtout pour respecter les délais de passation d'un appel d'offres.

La deuxième piste serait de confier la copie de la base principale au prestataire actuel de la Plateforme des données de santé, dont la maison-mère n'est pas européenne. La Cnil a déjà autorisé en décembre 2023, à titre dérogatoire, la Plateforme des données de santé à recourir à ce prestataire pour constituer un entrepôt de données de santé à des fins de pharmacovigilance. Cette dérogation, accordée en raison de sa durée limitée et au regard d'un engagement pris auprès de l'Agence européenne des médicaments⁵²², a été récemment contestée devant le Conseil d'État. Une dérogation étendue ne serait envisageable que si sa légalité était confirmée et si des assurances pouvaient être obtenues sur la mise en œuvre rapide de la solution-cible.

Une troisième piste serait de limiter le recours au prestataire actuel de la Plateforme des données de santé, après autorisation de la Cnil, aux seuls traitements de données liés à l'intelligence artificielle. Dans cette hypothèse, la Cnam puis le nouveau prestataire sélectionné prendrait en charge les demandes d'accès répondant à des besoins courants, qui sont les plus nombreux.

Tout en étant indispensable pour décharger la Cnam, une solution alternative occasionnerait des surcoûts par rapport au schéma initialement envisagé. En 2022, la Plateforme des données de santé avait estimé à 1,5 M€ le coût minimal induit par la migration vers un fournisseur de services autre que celui qu'elle avait sélectionné⁵²³. Une fois la migration effectuée, un surcoût d'exploitation par rapport à l'hébergeur actuel serait aussi à supporter.

3 - Des mesures à prendre rapidement pour réduire les délais de mise à disposition par la Cnam

Dans l'attente d'une solution alternative à mettre en place le plus rapidement possible, la mise à disposition des données issues de la base principale du SNDS continuera à reposer sur la Cnam. Dans cet intervalle, les délais de mise à disposition des données aux demandeurs doivent être réduits.

La Cnam prévoit d'actualiser en 2024 ses procédures internes de traitement des données et de conclusion des conventions de cession de données. À cette occasion, les délais inhérents à chaque étape du processus devraient être identifiés afin d'engager les mesures correctrices nécessaires pour les réduire.

⁵²² Délibération de la Cnil n° 2023-146 du 21 décembre 2023.

⁵²³ En revanche, une telle opération n'occasionnerait pas nécessairement le versement d'une indemnité de rupture contractuelle au prestataire actuel.

Par ailleurs, des moyens de la Plateforme des données de santé pourraient être redéployés temporairement vers la Cnam. Déjà, 3,6 équivalents temps plein ont été mis à disposition et des capacités de traitement de la part de bureaux d'études ont été financées. Des marges de manœuvre supplémentaires existent au regard des moyens dont dispose la Plateforme des données de santé, compte tenu de la sous-exécution en dépenses qui a marqué sa gestion budgétaire ces dernières années. Au 31 décembre 2022, son fonds de roulement atteignait 32,6 M€ et représentait plus de deux années de ses charges brutes d'exploitation.

Enfin, en suivant la logique retenue dans une convention de co-traitement conclue entre la Cnam et la Plateforme des données de santé⁵²⁴, une répartition du traitement des demandes pourrait être organisé.

Pour insatisfaisant que soit un tel dispositif au regard du schéma institutionnel prévu par la loi du 24 juillet 2019, il permettrait de réduire les délais de mise à disposition des données du SNDS, et ainsi d'encourager une plus vaste utilisation de son contenu par la recherche et par les acteurs privés.

⁵²⁴ Convention conclue le 30 novembre 2023 entre la Cnam et la Plateforme des données de santé pour organiser la répartition des tâches entre les deux organismes s'agissant d'un ensemble de 10 projets impliquant d'apparier le SNDS avec d'autres sources de données.

CONCLUSION ET RECOMMANDATIONS

Dans sa communication à l'Assemblée nationale de 2016 sur les données personnelles gérées par l'assurance maladie, la Cour avait souligné la richesse du système national d'informations interrégimes de l'assurance maladie (Sniiram), sans équivalent dans d'autres pays, et regretté le trop faible usage qui en était fait.

La création en 2016 du système national des données de santé (SNDS) puis son extension en 2019 devaient permettre d'y remédier en complétant ces bases de données par des informations cliniques. Toutefois, les intentions du législateur ne se sont traduites que par des progrès très en-deçà des ambitions affichées.

Une action plus résolue apparaît nécessaire. Le champ du SNDS doit être étendu aux entrepôts de données de santé hospitaliers et être enrichi avec plus de données cliniques. Surtout, cet enrichissement restera vain tant qu'il n'aura pas été remédié au blocage de l'hébergement des données, qui empêche la Plateforme des données de santé d'exercer ses missions.

La Cour formule donc la recommandation de politique publique et les deux recommandations de gestion suivantes :

- 39. dans l'attente de la mise en œuvre de la solution cible de cloud relevant exclusivement du droit de l'Union européenne, adopter une solution de moyen terme permettant la remise par la Cnam d'une copie de la base principale du SNDS à un hébergeur relevant exclusivement du droit de l'Union européenne (ministère du travail, des solidarités et de la santé, Plateforme des données de santé et Cnam) ;*
 - 40. dans l'immédiat, réduire les délais de mise à disposition des données en amplifiant les actions permettant d'accélérer le traitement des demandes par la Cnam (Cnam, Plateforme des données de santé) ;*
 - 41. intégrer dans le catalogue du SNDS le contenu du socle commun de données devant figurer dans les entrepôts de données de santé hospitaliers, en vue d'un hébergement sous la responsabilité de la Plateforme des données de santé (ministère du travail, de la santé et des solidarités, Plateforme des données de santé).*
-