

Working paper



30

L'accès aux données de santé en France : de la création des bases à leur ouverture 'en sourdine' par l'Institut des Données de Santé (période 1990-2010)

Chloé BÉRUT, chaire PARI

Mars 2023

PARI

PROGRAMME DE RECHERCHE
SUR L'APPRÉHENSION DES RISQUES
ET DES INCERTITUDES

L'accès aux données de santé en France : de la création des bases à leur ouverture 'en sourdine' par l'Institut des Données de Santé (période 1990-2010)

Chloé Bérut, chaire PARI, chercheuse associée au laboratoire Pacte
chloe.berut@umrpacte.fr

Résumé

Ce premier Working paper vise à décrire les évolutions du secteur de la régulation de l'accès aux données de santé en France des années 1990 jusqu'à la fin des années 2000. Durant cette période, on assiste à la mise en place des bases de données de santé centralisées par la caisse nationale d'assurance maladie et par le Ministère de la Santé. Ces bases de données, qui servent avant tout au pilotage (notamment financier) du système de santé, sont dans un premier temps essentiellement portées par des petites équipes d'acteurs administratifs. Certaines 'fenêtre d'opportunité' politiques vont ensuite permettre de stabiliser cet enjeu sur l'agenda institutionnel, tout en créant les conditions d'une ouverture 'en sourdine' des données de santé tout au long des années 2000. Pendant cette décennie, c'est en effet essentiellement l'Institut des Données de Santé, créé dans le cadre des négociations entre l'État, la caisse nationale d'assurance maladie, et les organismes d'assurance complémentaire, qui va contribuer à ouvrir l'accès aux données de santé. Cette ouverture concerne cependant essentiellement ses membres, qui sont des représentants de nombreux acteurs du système de santé : professionnels de santé, représentants de patients, organismes d'assurance complémentaire, établissements hospitaliers.

Sommaire

Introduction	2
I. La création des bases de données de santé dans un contexte de maîtrise des dépenses de santé (années 1990)	7
a) Les débuts des bases de données de santé : des projets ‘techniques’ soutenus par un petit nombre d’acteurs	8
Les débuts du Programme de Médicalisation des Systèmes d’Information PMSI (État, secteur hospitalier)	8
Les débuts du Système national d’information inter-régimes de l’Assurance maladie (CNAMTS, secteur ambulatoire)	9
Un projet essentiellement soutenu par ‘la secte’	9
b) Les Ordonnances Juppé de 1996 : une fenêtre d’opportunité	11
c) Les usages secondaires des données de santé : de l’impensé aux premières restrictions	12
Des bases de données peu orientées vers les usages secondaires	12
‘Mésusage des données’ et ‘Big Brother’ : les débats autour de l’usage secondaire des données de santé	13
La modification de la loi LIL et l’introduction de la CNIL dans la gouvernance opérationnelle des données de santé (1998)	17
II. Une ouverture ‘en sourdine’ portée par l’Institut des données de santé (années 2000)	20
a) La différenciation des règles d’accès applicables au SNIIRAM	20
La stabilisation des règles d’accès au PMSI	20
L’évolution contrastée des règles d’accès au SNIIRAM	21
La CNAMTS : une position de réticence face à l’ouverture des données	22
b) La loi du 13 août 2004 relative à l’Assurance Maladie et la création de l’Institut des Données de Santé	23
Un contexte de transformation des relations entre assurance maladie obligatoire et assurances complémentaires	23
L’enrichissement des missions de l’Institut des Données de Santé	25
La CNAMTS et l’Institut des Données de Santé : une relation conflictuelle	26
c) L’ouverture progressive des données du SNIIRAM	28
L’extension des finalités associées à l’usage du SNIIRAM	28
Les règles d’accès : une ouverture ambivalente	29
La gouvernances opérationnelle et stratégique : multiplication des acteurs et concurrences	31
Les enjeux pratiques : le développement de nombreux outils pour faciliter l’exploitation des données du SNIIRAM	32
Conclusion	36
Bibliographie	37
Annexe : Liste anonymisée des entretiens réalisés	38

Introduction

L'objectif de ce document de travail est de présenter les premiers résultats d'une enquête menée en 2022-2023 sur la régulation de l'accès aux données de santé en France. Plus précisément, il s'agit d'interroger la construction de cette régulation, en particulier en ce qui concerne l'accès aux bases de données de santé centralisées créées dans le cadre du remboursement de soins et du pilotage financier de l'Assurance Maladie et des hôpitaux. Ainsi, les deux grandes bases centrales analysées dans cette recherche sont le Système national d'information inter-régimes de l'Assurance maladie (SNIIRAM), qui rassemble les données pseudonymisées des remboursements des prestations aux assurés, et le Programme de médicalisation des systèmes d'information (PMSI) qui est constitué par les données relatives aux séjours en hôpital. Depuis 2016, ces deux bases sont rassemblées (et appareillées) au sein du Système National des Données de Santé (SNDS), qui contient également les données sur les causes médicales de décès (CépiDc).

Ces bases de données constituent une source d'informations très importante pour de nombreux acteurs, bien que les objectifs d'origine de ces bases soient essentiellement de nature gestionnaire. Avec le processus de digitalisation de la santé et le développement de l'intelligence artificielle, la question de l'usage secondaire¹ des données de santé devient de plus en plus prégnante. Parfois qualifiées 'd'or noir'² du XXIème siècle, les données de santé peuvent être utilisées à de nombreuses fins : améliorer le parcours du patient, aider à la formulation des politiques publiques, ou encore réaliser des recherches scientifiques, notamment en santé publique (Skovgaard *et al.* 2019). Pour certains acteurs privés, ces données peuvent présenter un intérêt commercial non négligeable.

En parallèle, la crise de la COVID-19 a contribué à renforcer l'attrait pour les données de santé, désormais perçues comme un outil crucial de lutte contre les épidémies. Ainsi, les plateformes nationales d'accès aux données de santé, pour la plupart mises en place très récemment, ont toutes 'accélééré les procédures d'accès aux données pour les recherches relatives à la COVID-19'³. Au niveau européen également, les acteurs s'organisent pour rendre les données des États membres compatibles et comparables ('interopérables') dans un 'Espace européen des

¹ L'usage secondaire des données correspond à un traitement des données différent de celui qui était envisagé au moment de la récolte des données. Par exemple, les données de la sécurité sociale sont récoltées à des fins administratives et de remboursement, et leur usage dans le cadre d'une recherche en santé publique correspond à un usage secondaire.

² Lors des entretiens menés pendant cette recherche, le vocabulaire lié à l'or est très souvent utilisé pour désigner les données de santé : 'trésor', 'mine d'or', 'coffre-fort'...

³ Health Data Hub (2021). *Les structures d'accès et de gouvernance des données de santé. Restitution du benchmark international 2020*. Paris : publications du Health Data Hub, p.34

données de santé' (*European health data space*, EHDS), dont l'objectif serait, entre autres, d'améliorer la réponse européenne en cas de crise sanitaire.⁴

Au vu des outils technologiques dont nous disposons (ou commençons à disposer) aujourd'hui, la question de l'accès à ces bases de données devient donc un enjeu majeur sur la scène politique, et c'est ainsi souvent avec le prisme de l'ouverture ou de la fermeture de cet accès que nous avons tendance à aborder le sujet de la régulation des bases de données de santé. Cela étant posé, répondre de manière simple à cette question (les bases de données sont-elles plutôt 'ouvertes' ou 'fermées' ?) n'a rien d'évident, et ce pour différentes raisons. En premier lieu, et comme nous le verrons dans ce document de travail, l'accès à ces bases de données de santé va être conditionné par un ensemble d'éléments, bien plus complexe que la simple question des règles d'accès. C'est ainsi tout un ensemble de 'blocs' interagissant les uns avec les autres, qui vont participer à la création de conditions d'ouverture ou de fermeture de l'accès aux bases de données de santé : les règles d'accès, mais également la nature de la gouvernance opérationnelle et de la gouvernance stratégique, ainsi d'autres d'enjeux plus 'techniques' ou 'pratiques' :

- **Les règles d'accès** : Les règles d'accès constituent sans doute l'aspect le plus visible de l'accès aux données de santé pour un observateur extérieur. Il est ainsi généralement possible de se faire une première idée des dynamiques d'ouverture et de fermeture d'une politique d'accès aux données en analysant la palette d'acteurs pouvant prétendre y avoir accès, et en observant les conditions dans lesquelles cet accès est susceptible d'être réalisé (objectifs à atteindre, procédure à suivre, moyens à mettre en œuvre etc.). En plus des règles établies sur un plan général, il s'agit également de s'intéresser à l'ensemble des règles d'exceptions qui sont très nombreuses dans ce secteur. Ces règles peuvent soit donner des accès différenciés selon les jeux de données considérés (on peut par exemple, créer plusieurs jeux de données thématiques de données à partir du SNIIRAM, au sein desquels les données peuvent être plus ou moins agrégées), soit établir des procédures d'accès spécifiques à certains acteurs. Aussi les règles d'accès, qui peuvent sembler relativement simples à appréhender aux premiers abords, sont-elles en réalité complexifiées du fait de l'existence d'une multitude d'exceptions.
- **La gouvernance opérationnelle** : Un autre aspect relativement visible de la gouvernance de l'accès aux données de santé pour un observateur externe est relatif à la mise en œuvre opérationnelle de la régulation. Cette gouvernance opérationnelle désigne les acteurs et organisations en charge de mettre en œuvre les règles d'accès adoptées au niveau national, c'est-à-dire dans le cas des données de santé le ou les organisme(s) participant à la délivrance d'autorisations d'accès aux données du SNIIRAM, du PMSI, ou à partir de 2016 du Système National des Données de Santé (SNDS).
- **La gouvernance stratégique** : Un autre type d'instrument central pour comprendre la régulation de l'accès aux données de santé a trait à la

⁴ Martins Henrique (2021). *EU health data centre and a common data strategy for public health*. Bruxelles : publications du Parlement européen.

‘gouvernance stratégique’ de ce secteur. Par ‘gouvernance stratégique’, nous entendons les dispositifs visant à donner des orientations spécifiques à la politique d’accès aux données de santé. Peu visible, cette gouvernance est pourtant cruciale ; la question est de savoir qui fait partie de cette gouvernance stratégique, dans la mesure où elle est susceptible d’influencer les Ministres et la rédaction des textes réglementaires en matière de données de santé.

- **Les enjeux pratiques** : Un des derniers aspects à prendre en compte dans l’analyse de l’ouverture ou de la fermeture des bases de données de santé est celui des enjeux proprement ‘pratiques’ ou ‘techniques’ liés à l’accès et à la réalisation des extractions. En effet, l’octroi d’une autorisation n’est que la première étape de l’accès, il s’agit ensuite de permettre aux demandeurs d’accéder aux données qu’ils souhaitent traiter. Dans ce cadre, plusieurs caractéristiques ou dispositifs peuvent contribuer à faciliter ou au contraire rendre plus difficile l’accès aux données de santé. Les bases de données telles que le SNIIRAM et le PMSI, créées avec un objectif comptable et d’administration, sont en effet particulièrement difficiles à manipuler et à interpréter. Elles requièrent l’acquisition d’outils techniques spécifiques pour pouvoir les explorer, mais également une maîtrise approfondie de la réglementation afin d’éviter les erreurs d’interprétation.⁵ Une des façons de faciliter cet accès est de créer des outils plus faciles à appréhender techniquement, grâce auxquels les données sont déjà en partie traitées (c’est-à-dire triées et lissées). Un autre aspect important à prendre en compte est celui des moyens mis en œuvre pour réaliser les extractions. Si le nombre de demandes est disproportionné par rapport aux moyens alloués, l’instruction des demandes peut alors se dérouler sur un temps extrêmement long, peu compatible avec les objectifs des demandeurs. Ce temps peut encore être allongé lorsque les données demandées font appel à des appariements entre plusieurs jeux de données, appariement qui peuvent être faits de manière systématique (objectif du SNDS ou du *Health Data Hub*). Enfin, le support de l’extraction lui-même peut constituer un enjeu important de l’accès : faut-il se rendre sur place pour analyser les données ? Ou est-il possible de créer une ‘bulle’ d’accès sécurisée à distance, qui permet de tracer les opérations réalisées sur les données ?

Du fait de ces multiples ‘blocs’, l’accès aux données de santé se caractérise la plupart du temps par son ambivalence plutôt que par son degré absolu d’ouverture ou de fermeture. Cette ambivalence se manifeste également par la diversité des acteurs à prendre en considération pour comprendre l’ouverture ou la fermeture des bases de données : à chaque période analysée, certains types d’acteurs vont connaître des restrictions d’accès, tandis que d’autres connaîtront dans le même temps une ouverture de leur accès aux données de santé. Pour comprendre l’ouverture et la fermeture des bases de données de santé en France, il faut ainsi

⁵ Pour prendre un exemple très simple, une décroissance abrupte de l’utilisation d’une spécialité pharmaceutique peut n’avoir aucun rapport avec la diminution de la fréquence d’une maladie, mais simplement être le reflet de retrait de cette spécialité de la liste des médicaments remboursés.

comprendre *qui* a accès aux bases de données, à quel *moment* et dans quelles *conditions*.

L'ambiguïté de ce secteur s'explique en grande partie par le fait que l'enjeu de l'accès aux données de santé s'inscrit dans des luttes entre les acteurs historiques du système de santé : assurance maladie, Ministères, professionnels de santé, mais également industries pharmaceutiques, représentants de patients et assureurs complémentaires. Plus encore, la thématique des données de santé a la particularité de rassembler l'ensemble de ces acteurs autour d'un même enjeu, alors que leurs activités tendent habituellement à se concentrer sur des points d'entrée différenciés : autorisation de mise sur le marché et fixation des prix pour les industries pharmaceutiques, négociations conventionnelles (ou question plus générale de la tarification des prestations) pour les professions libérales du secteur de la santé, construction de la 'démocratie sanitaire' pour les organisations de patients (ou 'représentants d'usagers'), ou encore organisation et répartition de la couverture du risque santé pour les assurances maladie obligatoire et complémentaires.

Encadré. Les acteurs du système de santé français

En France, le système de santé (dans sa version après-guerre) a été construit sur un modèle Bismarckien, dans lequel les professionnels de santé exercent majoritairement leur profession en libéral. Dans les systèmes dits Bismarckiens, le financement de la protection sociale (et dans notre cas, du risque santé) est principalement assuré par le biais de cotisations salariales et patronales. C'est ainsi le fait de travailler qui 'donne droit' à une protection sociale. En termes organisationnel, cela signifie que les caisses qui récoltent ces cotisations (en France l'Assurance maladie, ou la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés – CNAMTS) sont des acteurs centraux du système de santé, car ce sont elles qui assurent toute la gestion financière du système, en particulier en ce qui concerne les soins ambulatoires (médecine en dehors de l'hôpital). Ce sont par exemple les négociations entre la CNAMTS et les professionnels de santé libéraux qui sont à l'origine de la fixation des prix des consultations et autres actes médicaux. Par contraste, le secteur hospitalier (en dehors du secteur privé) est géré par l'État, et les professionnels de santé qui y travaillent sont généralement des salariés ('agents publics'). L'État a par ailleurs la charge d'assurer la protection sociale des personnes qui ne peuvent travailler. Cette protection est financée par l'impôt, et repose sur un principe de solidarité nationale (et non sur une solidarité entre 'cotisants').

Si, historiquement, la CNAMTS et l'État sont deux acteurs nettement distincts (les décisions à la CNAMTS étant prises par les représentants syndicaux des travailleurs et des employeurs), les évolutions de ces dernières années tendent à renforcer le contrôle de l'État sur la CNAMTS, tandis que le financement de l'assurance santé passe de plus en plus par l'impôt. Malgré ces évolutions, la CNAMTS et l'État (et en particulier le Ministère de la Santé) restent deux 'poids lourds' du système de santé, dont les relations, plus ou moins conflictuelles selon les périodes et les enjeux, sont centrales dans la prise de décision en matière de politique de santé.

En parallèle, les professionnels de santé, partenaires centraux du système de santé, sont également un acteur crucial des réformes, de même que les laboratoires pharmaceutiques, qui sont en négociation constante avec la CNAMTS et l'État dans le cadre de la fixation des prix et des remboursements des médicaments. Plus récemment, des initiatives ont été prises afin d'institutionnaliser la participation d'autres acteurs à la gouvernance du système de santé ; c'est le cas par exemple des assureurs complémentaires avec la création de l'Union nationale des organismes complémentaires d'assurance maladie (UNOCAM) en 2004, ou encore des organisations de patients avec la création en 2016 de France Assos Santé (Union Nationale des Associations Agréées du Système de Santé).

Pour comprendre la régulation de l'accès aux bases de données de santé en France et ses évolutions, c'est précisément ces luttes d'acteurs qu'il s'agit d'analyser, car ce sont elles qui structurent les résultats de la prise de décision. En parallèle, ces luttes de pouvoir s'insèrent dans des développements plus généraux, qui vont directement affecter la capacité des acteurs à influencer le processus décisionnel : c'est le cas par exemple de l'imbrication de l'enjeu des données de santé au sein de problématiques telles que la réduction des dépenses de santé, le développement de l'*open data* ou de l'intelligence artificielle. Sur certaines périodes, les transformations du secteur de l'accès aux données de santé sont ainsi fortement influencées par les 'fenêtres d'opportunité' constituées par la mise sur agenda de ces enjeux plus globaux.

Dans ce premier Working paper, ces dynamiques seront explorées sur la période 1990-2010. Ce premier découpage temporel permet de comprendre la régulation de l'accès aux données de santé à un moment où cet enjeu ne constitue pas encore une catégorie d'action publique à part entière. En effet, de leur création jusqu'à la fin des années 2000, le développement et l'accès aux bases de données de santé est essentiellement subordonné à la réalisation d'objectifs en lien avec d'autres enjeux (réduction des dépenses de santé dans un premier temps, puis partenariat avec les assurances complémentaires dans un second temps). Ce n'est qu'à partir du début des années 2010 que l'enjeu des données de santé commence à se constituer de manière autonome (voir Working paper n°2).

Afin de retracer les changements intervenus au cours de cette période dans le secteur de la régulation de l'accès aux bases de données en France, ce document de travail revient sur les dynamiques et processus ayant donné lieu à des transformations de l'accès aux données de santé en France. Ces transformations seront par ailleurs décrites à l'aide de la typologie présentée ci-dessus (règles d'accès, gouvernance opérationnelle, gouvernance stratégique, enjeux pratiques). Du fait de la complexité de ce secteur, l'objectif de ce document de travail est ainsi d'identifier des 'mouvements' ou des dynamiques d'ouverture ou de fermeture, plutôt que de produire un constat, dans l'absolu, sur le 'degré d'ouverture' des bases de données en France.

Le matériau empirique de cette enquête qualitative a principalement été récolté entre janvier et septembre 2022. La collecte de données a mobilisé les stratégies classiques de l'analyse qualitative des politiques publiques : récolte et analyse documentaire (lois, rapports, documents de politique publique, articles de

presse...), et réalisation d'entretiens semi-directifs. Pour identifier les enquêtés, nous avons d'abord utilisé la méthode positionnelle (identifier les personnes ayant occupé une fonction institutionnelle ou politique en lien avec la régulation de l'accès aux données de santé), complétée occasionnellement par une méthode réputationnelle (personnes pertinentes à contacter selon les enquêtés). En tout, 21 entretiens semi-directifs ont été réalisés, dont la moitié en présentiel.⁶ Les personnes rencontrées occupaient généralement des fonctions au sein du Ministère de la Santé, de la Caisse Nationale d'Assurance Maladie des Travailleurs Salariés (CNAMTS), des cabinets ministériels ou au sein d'agences spécialisées dans la régulation de l'accès aux données de santé (comme l'Institut des données de santé par exemple). Les questions visaient à comprendre les représentations des acteurs en matière de données de santé (bénéfices et/ou risques associés à l'ouverture, positionnement), mais également à recueillir des récits relatifs au déroulé des événements sur une période donnée. Les informations, en particulier celles qui occupent un statut central vis-à-vis de l'administration de la preuve, ont été 'triangulées' autant que possible, c'est-à-dire comparées et confirmées à l'aide de différentes sources.

Dans les deux parties suivantes, les dynamiques intrinsèques à la régulation de l'accès aux données de santé sont décrites autour de deux périodes spécifiques : les années 1990 et les années 2000. Dans une première partie, nous nous intéresserons ainsi à la mise en place des bases de données de santé (PMSI et SNIIRAM) et aux débats relatifs à leur création et à leur accès par des acteurs tiers. Comme nous le verrons, à cette époque, c'est plutôt l'utilité de ces bases en matière de gestion et de pilotage financier qui est soulignée, ce qui permet aux Ordonnances Juppé de 1996 de constituer une réelle fenêtre d'opportunité assurant leur mise en place opérationnelle. Au cours des années 2000, une nouvelle configuration institutionnelle liée aux négociations entre l'État et les assurances maladie complémentaires va profondément transformer l'accès aux données de santé. Ainsi, l'Institut des Données de Santé est créé en 2004 et se révèle particulièrement proactif pour ouvrir l'accès aux données de santé à un plus grand nombre d'acteurs. Que ce soit dans les années 1990 ou 2000 cependant, il est intéressant de noter que l'enjeu des données de santé est subordonné à d'autres enjeux du système de santé et n'est pas traité en tant que secteur 'autonome'. C'est pourquoi les évolutions de ce secteur se réalisent principalement 'en sourdine', au gré des luttes institutionnelles qui caractérisent les relations entre les différents acteurs du système de santé.

I. La création des bases de données de santé dans un contexte de maîtrise des dépenses de santé (années 1990)

Les travaux administratifs sur la création de bases de données de santé débutent dans les années 1970-1980 au sein de la CNAMTS (données des assurés et de la médecine de ville) et du Ministère de la Santé (données des hôpitaux). Les acteurs

⁶ Certains de ces entretiens ont été réalisés antérieurement à cette étude dans le cadre de la préparation d'une thèse de doctorat en science politique sur les politiques d'e-santé.

qui s'en emparent forment alors un petit groupe relativement homogène de personnes convaincues (la 'secte') qui s'investissent pleinement dans la création de ces bases, dont le potentiel est encore largement méconnu. En 1995, dans le contexte du plan Juppé, les bases de données de santé sont perçues par le politique comme des instruments pouvant contribuer à un meilleur pilotage financier du système de santé. À partir de cette époque, la construction des bases de données de santé se stabilise sur l'agenda institutionnel. La question des usages secondaires des données de santé n'est cependant pas réellement conceptualisée ; c'est plutôt les premiers usages publics du PMSI par des journalistes, ainsi que les débats relatifs au risque de 'mésusage' des données de santé (en particulier par l'État) qui conduisent aux premières formes de régulation de ce secteur à la fin des années 1990.

a) Les débuts des bases de données de santé : des projets 'techniques' soutenus par un petit nombre d'acteurs

Les débuts du Programme de Médicalisation des Systèmes d'Information PMSI (État, secteur hospitalier)

Le Programme de Médicalisation des Systèmes d'Information (PMSI) est une base de données mise en place dans le cadre de la refonte du financement des hôpitaux. Dans le système hospitalier d'après-guerre, il n'existe en effet pas de données fiables relatives à l'activité hospitalière, ce qui rend le pilotage du financement des hôpitaux particulièrement délicat (Entretien, ex-fonctionnaire, 24 et 30/05/2018 ; voir également Moisdon 2013 sur le système de 'dotations annuelles' qui prévalait alors). Initialement présenté comme un simple 'enrichissement' des données disponibles sur le système hospitalier en 1983 (Moisdon 2013), le PMSI est ensuite généralisé suite à l'obligation pour les établissements hospitaliers de procéder à 'l'évaluation et l'analyse de leur activité'⁷. Contrairement à la CNAMTS, la culture de la gestion et du pilotage financier semble moins présente à cette époque au sein du Ministère et plus précisément de la Direction de l'hospitalisation et de l'organisation des soins (DHOS). Aussi le projet PMSI est-il relativement marginalisé à ses débuts (voir sous-section 'Un projet essentiellement soutenu par « la secte »'). Aujourd'hui, ce sont les données du PMSI qui sont utilisées dans le cadre de la tarification à l'activité (T2A) des hôpitaux.

Le PMSI contient des données produites à partir d'un document administratif appelé 'Résumé de sortie standardisé', qui, une fois, anonymisé (c'est-à-dire dépouillé de ses informations directement identifiantes : nom, numéro de sécurité sociale...), est rebaptisé 'résumé de sortie anonyme' (RSA). Les RSA contiennent les informations médicales et administratives des patients de passage dans un service hospitalier. Les diagnostics sont classés à l'aide de la Classification internationale des maladies (CIM) 10, une classification produite par l'Organisation Mondiale de la Santé. Les actes médicaux sont quant à eux référencés à partir de la Classification commune des actes médicaux (CCAM). Ces données permettent ensuite de classer les patients par groupe homogène de malades (GHM), qui, à

⁷ Section 2 de la Loi du 31 juillet 1991 portant réforme hospitalière

partir de données administratives (comme l'âge) et des diagnostics permettent de déterminer une rémunération forfaitaire pour la prise en charge par les hôpitaux. Les GHM peuvent ainsi regrouper plusieurs pathologies différentes, le critère de regroupement étant le montant attribué pour la prise en charge de l'individu.

Les débuts du Système national d'information inter-régimes de l'Assurance maladie (CNAMTS, secteur ambulatoire)

Au niveau de la CNAMTS, l'existence de bases de données de santé centralisées n'est pas récente, et c'est dès 1977 qu'est créé le Système National Inter-régime (SNIR), un système d'information géré par la CNAMTS dans le but d'agréger les données associées aux remboursements effectués dans le cadre de la pratique libérale. À cette époque cependant, le codage ne permet pas d'avoir une vue détaillée des actes réalisés par les médecins.⁸

'Au début des années 1990, l'Assurance Maladie utilise la **nomenclature générale des actes professionnels** qui est essentiellement faite **pour donner un prix**. Alors vous avez une lettre qui indique si c'est un acte de chirurgie ou un acte d'imagerie, par exemple les actes de chirurgie c'est K, les actes d'imagerie c'est Z, et puis derrière vous avez un chiffre qui en gros est un prix, l'appendicectomie par exemple c'est K50. Et puis donc, mais il y a plein de K50 qui ne sont pas tous des appendicectomie, et donc la NGAP [Nomenclature Générale des Actes Professionnels] **ne permet pas de faire le codage des actes**' (Entretien, ex-fonctionnaire, 24 et 30/05/2018)

Ce n'est en effet qu'à partir des années 1990 que naissent les bases de données de santé centralisées avec une captation d'informations suffisamment fine pour constituer de riches bases de données exploitables pour tout un ensemble de fins. Ainsi, sur une feuille de soins, l'ensemble des éléments est codé afin de donner une information précise sur les actes médicaux pratiqués, ainsi que sur les spécialités pharmaceutiques prescrites et remboursées. L'objectif est alors d'assurer un meilleur pilotage financier de l'assurance maladie. La centralisation de ces bases de données dans un unique entrepôt, le Système national d'information inter-régimes de l'Assurance maladie (SNIIRAM), n'intervient cependant qu'à la fin des années 1990 avec l'adoption des Ordonnances Juppé (voir section b) ci-dessous).

Un projet essentiellement soutenu par 'la secte'

Dans les années 1990, le potentiel de ces bases de données est loin d'être perçu par l'ensemble des acteurs politico-administratifs, dans un contexte où l'informatisation et la numérisation des données en est encore à ses débuts. Les architectes de ces bases forment alors des groupes relativement restreints (s'appelant eux-mêmes avec humour 'la secte', Entretien, ex-fonctionnaire, 06/04/2022) au sein du Ministère de la Santé ou de la CNAMTS, tandis que ce sujet tend à être marginalisé politiquement du fait de son aspect essentiellement 'technique'. Le potentiel de ces bases de données, tel qu'il peut être présenté par 'la secte', est alors associé pour de nombreux acteurs à des croyances, certains

⁸ CNIL (1995). *16^{ème} rapport d'activité 1995*. Paris : La documentation française, p.92, voir également Entretien, ex-agent CNAMTS, 22/09/2022

percevant ces projets comme relevant d'une forme de 'mégalomanie' (Entretien, ex-agent CNAMTS, 22/09/2022).

'Oh ben le PMSI ça a été longtemps le fait d'un **petit noyau de gens convaincus**, d'une **secte** quasiment, et SNIIRAM un peu pareil hein donc euh... C'était un peu moins, enfin PMSI c'était à la fois fabriquer des données et la base, SNIIRAM les données existaient, il s'agissait de les regrouper dans une même boîte, c'était un peu différent. Mais voilà c'est pas des dossiers qui ont bénéficié à l'origine d'une médiatisation, d'une communication politique intense, c'était un sujet voilà, de l'administration, petit, enfin un **petit sujet**.' (Entretien, ex-fonctionnaire, 06/04/2022).

'Je suis arrivé au Ministère de la santé, j'étais **passionné**, le milieu du PMSI était un truc que je connaissais que j'avais interviewé, l'ancien directeur des hôpitaux Jean Kersvadoué, lorsque j'étais à l'ENA, [?] le PMSI je me suis dit mais mon dieu mais bien sûr, voilà une manière de mesurer l'*output*, c'est vachement intéressant pour un économiste, et donc lorsque je sors de l'ENA, je demande à prendre la mission PMSI. Je vous passe les détails mais, et donc alors je suis nommé [...] donc je recrute des gens, je monte une **équipe de fous furieux** et on a monté, alors je considère qu'on a monté... Un peu dans notre coin une espèce **d'organisation militante pour mettre en place les systèmes d'informations hospitaliers** [...]. Et c'est un miracle [?]. Et donc on a monté ce truc-là, de manière très passionnée, et avec à l'époque un Ministre et surtout un directeur de cabinet, et un directeur [?], tous ces gens disent surtout **faites pas de vagues**, n'allez pas faire croire qu'on va financer les hôpitaux en partant du PMSI, n'y pensez-pas etc.' (Entretien, ex-fonctionnaire, 24 et 30/05/2018)

'[...] je travaillais avec eux, c'est vrai qu'ils étaient, voilà, ils étaient, y compris moi je pense, on était **des personnages un peu, un peu différents**, voilà. Aujourd'hui je dois paraître pour un vieux mais à l'époque on était des **jeunes et on en voulait quoi**.' (Entretien, agent CNAMTS, 21/04/2022)

'[...] le SNIIRAM, des débats sur cet énorme entrepôt qu'on était en train de constituer et qui [...] avait été considéré comme **un projet un peu fou** et assez **peu pertinent** par beaucoup de gens, notamment un peu par les gens de la statistique publique, c'est à dire disant qui, mais on peut travailler sur des échantillons, pourquoi on se met à faire des énormes entrepôts exhaustifs, ça part dans tous les sens, cette espèce **mégalomanie** et en fait, très honnêtement à l'époque on savait pas très bien ce qu'on allait en faire, et on était pas d'ailleurs certains qu'on en ferait tellement plus que ce qu'on faisait avec des échantillons des enquêtes ou quoi. [...] Donc les usages, oui, au début je vous disais beaucoup de **scepticisme**⁹ et beaucoup de voilà de **ricanements** sur, qu'est-ce c'est que cette **mégalomanie** de faire un énorme entrepôt avec tous les Français dedans à quoi ça va servir.' (Entretien, ex-agent CNAMTS, 22/09/2022)

L'intérêt politique relativement limité (en tout cas au sein des Ministères)¹⁰ pour la création de ces bases rend alors l'investissement dans leur construction

⁹ Ce terme de 'scepticisme' est également évoqué lors d'un entretien mené avec un ex-fonctionnaire le 06/04/2022.

¹⁰ Au sein de la CNAMTS, l'intérêt de la constitution de telles bases de données pour le pilotage de l'assurance maladie est plus nettement perçu, notamment par certains acteurs occupant de hautes fonctions décisionnelles (Entretien, ex-agent CNAMTS, 19/04/2022 ; Entretien, ex-agent CNAMTS, 22/09/2022).

particulièrement dépendant des appétences personnelles des différents Ministres (Entretien, ex-fonctionnaire, 06/04/2022), du moins jusqu'aux Ordonnances Juppé.

b) Les Ordonnances Juppé de 1996 : une fenêtre d'opportunité

Au milieu des années 1990 cependant, le déficit croissant de la branche maladie de la sécurité sociale et les Ordonnances Juppé contribuent à renforcer la place de l'enjeu des bases de données de santé sur l'agenda politique. Ces ordonnances peuvent ainsi être interprétées comme de véritables 'fenêtre d'opportunité' (Kingdon 2002), qui permettent aux bases de données de santé de s'inscrire sur l'agenda et d'assurer leur mise en œuvre dans la durée.

Encadré. Les Ordonnances Juppé de 1996

Le plan Juppé, initialement présenté sous la forme d'une Déclaration de Politique Générale en 1995 et adopté via plusieurs ordonnances en 1996, est une réforme emblématique de la Sécurité Sociale. Cette réforme est alors présentée par le gouvernement comme le seul moyen de 'sauver' la Sécurité Sociale. Si ce n'est pas la première fois que le gouvernement entreprend de rééquilibrer les comptes de l'Assurance Maladie (Palier 1999), la particularité de cette réforme réside dans l'ampleur des mesures adoptées. En effet, en plus des traditionnelles méthodes d'augmentation des recettes et de réduction des dépenses, le plan Juppé a la particularité d'introduire l'enjeu de la contrainte financière sur l'agenda institutionnel. C'est ainsi via le plan Juppé que sont créés la Loi de Financement de la Sécurité Sociale, votée tous les ans au Parlement, ainsi que l'Objectif National des Dépenses de l'Assurance Maladie (ONDAM), qui prévoit le budget 'maximal' de l'Assurance maladie. Le plan Juppé introduit également la Contribution pour le remboursement de la dette sociale (CRDS, impôt) et augmente la Contribution sociale généralisée (CSG). Enfin, le plan Juppé marque un tournant dans les relations entre la CNAMTS et l'État, dans la mesure où différentes dispositions permettent au gouvernement d'exercer un contrôle accru sur l'Assurance maladie (Hassenteufel and Palier 2005).

Le fait que des bases telles que le PMSI ou le SNIIRAM puissent fortement contribuer (ou du moins apparaître comme des outils pouvant contribuer) au 'sauvetage' de la sécurité sociale permet de stabiliser l'enjeu de leur construction sur l'agenda :

'[...] Mais en fait il y a eu à l'époque, vous voyez c'était bah c'était le problème des ordonnances Juppé, c'est-à-dire c'était le problème de **sauver la sécurité sociale**, donc on a des déficits de l'Assurance maladie, trouver d'urgence des mesures, grand plein de réformes de la sécurité sociale, trouver des mesures pour réformer l'assurance maladie, une mission IGAS [Inspection Générale des Affaires Sociales], IGF [Inspection Générale des Finances], inspection générale de l'administration, je me souviens très bien, missionnée pour regarder, pour proposer des outils, et ces gens viennent, ils arrivent à la DOS [Direction de l'Offre de Soins], la DH [Direction des Hôpitaux] de l'époque je sais plus, **ils aboutissent dans mon bureau**, quelqu'un les envoie chez moi, et ils connaissaient pas du tout, et ils disent vous faites quoi et **dans je leur explique ce qu'on fait, et ça leur a vachement plu** quoi. Et ensuite ils m'ont reconvoqué, ils étaient tous là [...], et vraiment dans leur rapport ils

ont joué un rôle extrêmement important **pour dire ce truc-là il faut absolument le développer, c'est un des outils essentiels du dispositif.** (Entretien, ex-fonctionnaire, 06/04/2022)¹¹

‘ [...] on a un système de d'assurance maladie éclaté hein, à 17 régimes d'assurance maladie et à partir de 1995 donc loi Juppé, bon **il faut voter l'ONDAM** [Objectif National de Dépenses de l'Assurance Maladie] et l'ONDAM, c'est *un* chiffre, tout régime de l'assurance maladie, et **on était incapable**, chaque mois et donc de **de faire des prévisions à partir de chiffres réels** des 12 régimes d'assurance maladie, donc on pouvait agréger les 3 principaux, c'est un régime général, MSA [Mutualité Sociale Agricole] et enfin, à l'époque, donc RSI [Régime Social des Indépendants] maintenant. Et puis pour le reste, eh bien, on extrapolait. [...] donc une première critique, c'était de dire mais qu'est-ce que, vous n'êtes pas foutu de de sortir un chiffre complexe, c'est le bazar et cetera. Ben le bazar, c'était qu'on avait des régimes complètement indépendants, et donc et c'est là que, après que la loi Juppé est sortie donc en décembre 1998 pour 1999, l'idée de de faire le SNIIRAM donc une méga base de données. Qui rassemble les données de tous les régimes en temps réel.’ (Entretien, ex-agent CNAMTS, 19/04/2022)

Les ordonnances Juppé apparaissent donc comme une véritable ‘fenêtre d’opportunité’ pour les premiers architectes des bases de données de santé, qui pour certains se réfèrent à un ‘concours de circonstances’ particulièrement propice (Entretien, ex-agent CNAMTS, 19/04/2022). Bien que les bases de données ne naissent pas en 1996, le plan Juppé permet de transformer leur positionnement relativement marginal en enjeu majeur de l’agenda politique. En ce qui concerne les hôpitaux, le PMSI est ainsi perçu comme un réel outil de maîtrise des dépenses, et sa construction sera dès lors plus étroitement associée à la gestion financière des hôpitaux pour aboutir, *in fine*, au système de tarification à l’activité que nous connaissons aujourd’hui. Côté CNAMTS, les ordonnances Juppé permettent surtout d’entreprendre un travail de centralisation et d’harmonisation des bases de données existantes, bases alors dispersées entre les différents régimes de l’Assurance maladie. Ce travail de longue haleine nécessite la création d’un immense entrepôt de données ainsi que mise en conformité du codage entre tous les organismes d’assurance maladie, notamment via l’adoption de ‘normes d’échange’ (Entretien, ex-agent CNAMTS, 22/09/2022).

c) Les usages secondaires des données de santé : de l’impensé aux premières restrictions

Des bases de données peu orientées vers les usages secondaires

En parallèle, la question de l’ouverture de ces bases de données à d’autres acteurs ne se pose pas encore. Côté hôpitaux et PMSI, comme nous l’avons vu, l’objectif affiché est ‘d’améliorer la connaissance et l’évaluation de l’activité et des coûts et de favoriser l’optimisation de l’offre de soins’¹², et non de conduire des études de santé publique. Par ailleurs, dans la loi qui crée le SNIIRAM en 1998, seules deux finalités sont explicitement mentionnées : la ‘connaissance des dépenses de l’ensemble des régimes d’assurance maladie’ et ‘la transmission en retour aux

¹¹ Voir également Entretien, ex-agent CNAMTS, 19/04/2022

¹² Section 2 de la Loi du 31 juillet 1991 portant réforme hospitalière

prestataires de soins d'informations pertinentes relatives à leur activité et leurs recettes, et s'il y a lieu à leurs prescriptions'.¹³ À ce moment-là, les usages secondaires en matière de santé publique, d'analyse du parcours de soins, ou encore d'études des effets des médicaments ne sont pas envisagés :

'Bah on... on avait oui enfin parce que c'était quand même une réflexion largement collective, oui on avait quand même une idée de l'intérêt, non pas pour la recherche, pas pour **la recherche honnêtement nous c'était pas notre problème**, on avait une idée de ce que ça pouvait représenter comme **intérêt dans le pilotage du système de santé**, et pas simplement le pilotage financier, voilà, y compris dans une logique de **planification**, comparaison des besoins, d'analyse épidémiologique mais **d'épidémiologie, administrative** si je puis dire. Voilà ça on voyait bien qu'il y avait un large usage possible, dans le champ du **pilotage du système de santé**. Les chercheurs honnêtement ce n'était pas notre problème quoi [rires], voilà la tâche était déjà suffisamment importante donc voilà.' (Entretien, ex-fonctionnaire, 06/04/2022)¹⁴

À ce titre, les destinataires envisagés des informations issues de ces bases de données sont les acteurs en charge du pilotage du système de santé (CNAMTS et Ministères), ainsi que les professionnels de santé, alors considérés comme les principaux 'producteurs' de la donnée.

'Mésusage des données' et 'Big Brother' : les débats autour de l'usage secondaire des données de santé

Au-delà de la question des finalités, la création du SNIIRAM et du PMSI génère des débats axés sur les risques associés à l'existence et aux potentiels 'mésusages' de ces bases de données. Cependant, ces deux bases ne semblent pas être soumises au même niveau de controverse. Plusieurs éléments permettent d'expliquer ces différences : la temporalité de création de ces deux bases, la nature même de leurs données, mais également le fait que ces bases de données ne soient pas gérées par les mêmes organismes (État pour le PMSI, CNAMTS pour le SNIIRAM).

Temporalité

Pour commencer, le PMSI, mis en place dans la première moitié des années 1990 (et donc avant la SNIIRAM), est relativement épargné par les polémiques présentant les initiatives étatiques en matière de données personnelles sous l'angle du 'Big Brother', qui pourtant avaient conduit en 1978 à l'adoption de la Loi Informatique et Libertés (LIL) et à la création de la Commission Nationale de l'Informatique et des Libertés (CNIL).

Encadré. L'affaire SAFARI et la loi Informatique et Liberté (1978)

La France est souvent présentée comme un pays pionnier dans le domaine de la protection des données personnelles. En effet, la première directive européenne

¹³ Article 21 de la Loi n° 98-1194 du 23 décembre 1998 de financement de la sécurité sociale pour 1999

¹⁴ Sur les objectifs attachés à la construction des bases de données, voir également : Entretien, agent CNAMTS, 21/04/2022 ; Entretien, ex-agent CNAMTS, 19/04/2022

en la matière, adoptée en 1995, serait directement inspirée de la Loi Informatique et Liberté (loi LIL) adoptée en 1978. Cette loi fait alors directement suite à l'affaire SAFARI. Le projet SAFARI (Système automatisé pour les fichiers administratifs et le répertoire des individus) désigne une initiative lancée par le gouvernement en 1973 dans le but de croiser plusieurs bases de données contenant des fichiers de renseignements nominatifs. Dénoncé par un article du Monde publié le 21 mars 1974 ('Safari ou la chasse aux français') le projet est finalement abandonné. À cette époque, c'est bien l'État, et non les plateformes numériques qui n'existent pas encore, qui est principalement associé à un risque de mésusage des données personnelles. L'État est alors dépeint comme un acteur susceptible d'utiliser les données personnelles à des fins de fichage et de surveillance des individus ('Big Brother').

Dans un premier temps, la CNIL ne s'estime pas compétente en ce qui concerne le PMSI, puisque les données sont réputées anonymes, et donc ne relevant pas d'un régime de protection des données personnelles (Entretien, ex-fonctionnaire, 06/04/2022). La notion d'anonymisation évolue cependant avec le temps, et ce que l'on considère comme anonyme en 1990 ne trouve pas forcément d'équivalent quelques années plus tard. Ainsi, suite à la mise en œuvre généralisée du PMSI, certains acteurs découvrent les risques de réidentification indirecte associés aux Résumés de Sortie Anonymes, risques passés inaperçus au démarrage du PMSI (Entretien, ex-fonctionnaire, 24 et 30/05/2018).¹⁵ En revanche, l'enjeu qui prévaut à la mise en place du PMSI est davantage axé sur la pratique des professionnels de santé : il s'agit en effet de leur permettre de faire remonter des informations relatives à leurs patients sans trahir juridiquement le secret médical auquel ils sont astreints.¹⁶ La mise en place du PMSI à un moment où les enjeux d'anonymisation des données de santé n'ont pas encore complètement émergé permet ainsi dans un premier temps au PMSI de 'passer sous le radar' et d'être adopté sans polémique majeure (en dehors des critiques relatives à l'introduction d'outils de contrôle dans la gestion hospitalière).

Par contraste, les débats associés à la construction du SNIIRAM, qui intervient à la fin des années 1990, font davantage référence aux risques associés à la création de ces bases de données pour les individus. En effet, l'initiative est très tôt associée à la volonté des pouvoirs publics de se poser en 'Big Brother' :

'Alors le gros débat, ça avait été, est-ce qu'on fait l'exhaustif individuel, ou est-ce qu'on fait de l'exhaustif agrégat et échantillon individuel. Ça, c'était le gros débat, parce que le toujours, la première chose qui est sortie, c'est, **vous êtes en train de faire un Big Brother**, l'idée de rassembler, dans un seul lieu, les données, de tous les Français, même anonymes, on a dit

¹⁵ Les données du PMSI, de même que celles du SNIIRAM, sont dites aujourd'hui 'pseudonymisées' et non 'anonymisées'. Cette distinction est cependant relativement récente. La pseudonymisation correspond aux situations où les données directement identifiantes (nom, date de naissance etc.) ont été supprimées ou pseudonymisées, mais où il existe un risque de réidentification indirecte dû à la nature (souvent cumulative) des données. Une donnée anonyme est une donnée pour laquelle toute réidentification est impossible.

¹⁶ Cet enjeu est ainsi spécifiquement traité par l'ajout du chapitre V bis dans la Loi Informatique et Libertés en 1994

pseudonymisées, **ça faisait peur à tout le monde**. Moi j'étais là, je je vous raconte pas les débats ! que j'ai dû subir à l'époque.¹⁷ (Entretien, ex-agent CNAMTS, 19/04/2022)

Ensuite, malgré un travail préalable important entre la CNIL et la CNAMTS, l'autorisation de la CNIL ne se joue finalement qu'à un vote, celui du Président de la CNIL, les votes pour et contre étant alors équitablement partagés (Entretien, ex-agent CNAMTS, 19/04/2022). Les relations interpersonnelles entre le Président de la CNIL de l'époque (Michel Gentot), et le directeur de la CNAMTS (Gilles Johannet) auraient été cruciales dans la décision finale d'autorisation de la CNIL (Entretien, ex-agent CNAMTS, 19/04/2022).

Le contexte de création du SNIIRAM, qui intervient seulement quelques années après celui du PMSI, semble donc radicalement différent : à la première réaction d'indifférence de la CNIL lors de la création du PMSI se substitue une approche beaucoup plus attentive aux risques posés pour les individus. Cette différence de cadrage autour de la construction de ces deux bases peut en partie s'expliquer par un effet de contexte dû à l'accumulation de différentes 'affaires' à la fin des années 1990 (et en particulier en 1998).

En premier lieu, la réutilisation des données du PMSI dès l'année 1998 contribue à mettre à l'agenda l'enjeu de l'usage (ou du 'mésusage' selon les points de vue) des données de santé. Ainsi, les données de la première année de fonctionnement du PMSI, analysées par des journalistes, donnent lieu à la publication d'un des premiers 'palmarès des hôpitaux' en septembre 1998 par le magazine Sciences et Avenir. À ce moment-là, la diffusion de ces informations ne nécessitait pas d'accord préalable de la CNIL, mais une saisine de la Commission d'accès aux documents administratifs – CADA (Pierru 2004; Malye and Vincent 2015).¹⁸ Les acteurs du système de santé, en particulier les professionnels de santé et le pouvoir politique, réalisent alors que ces données peuvent être utilisées pour remettre en cause ou dénoncer certaines pratiques, ce qui provoque d'importantes levées de bouclier, notamment par rapport à la méthodologie utilisée dans la construction de ces palmarès (Entretien, ex-fonctionnaire, 06/04/2022 ; voir également Pierru 2004). Le risque souligné ici est celui de la mauvaise interprétation des données (et

¹⁷ Afin de convaincre les fonctionnaires du Ministère du bien-fondé de la création d'une base de données individuelle et exhaustive pour les patients, cet enquêté nous raconte avoir stratégiquement lié l'enjeu de la construction de cette base de données à celle des données relatives aux professionnels de santé. En effet, une nouvelle loi obligeait alors la CNAMTS à fournir au service des impôts les données relatives à l'activité des professionnels de santé (notamment les honoraires perçus). Dans ce contexte, la base de donnée créée devait contenir des informations individuelles et exhaustives sur les professionnels de santé. Si la décision avait été prise de créer une base de données seulement agrégée concernant les données des patients, il aurait fallu créer deux bases de données avec deux systèmes différents de récolte des données sur les feuilles de soins. L'argument technique de la nécessité de ne créer qu'une seule base avec un unique mode de fonctionnement aurait alors été crucial pour convaincre certains fonctionnaires (Entretien, ex-agent CNAMTS, 19/04/2022)

¹⁸ Plus de 30 saisines de la CADA auraient été nécessaires pour accéder aux données du PMSI cette année-là (Pierru 2004)

non le fait que des données potentiellement confidentielles aient été utilisées par des journalistes).

‘Ah oui alors après sur le classement, mais le problème qui était posé **c'était pas de dire pour fabriquer ce classement on traite des données potentiellement confidentielles**, là la question qui a très vite été **la question c'était est-ce que ce classement a du sens**, est-ce qu'il est pertinent pas pertinent, est-ce qu'il va pas mettre à l'annexe justement des gens des hôpitaux, voilà ça a été ça le débat, le fait de dire mais pourquoi on a fait ce classement, il y a des gens qui ont accès à des données dans quelles conditions, zéro débats. Vraiment les premiers débats sur les palmarès ce n'était absolument pas ça.’ (Entretien, ex-fonctionnaire, 06/04/2022)

En parallèle et à la même période, la question de la déclaration obligatoire des patients atteints du SIDA et de l'informatisation de cette déclaration provoque également des tensions importantes entre les associations de représentants de patients, les agences sanitaires (notamment l'Institut de Veille Sanitaire, InVS) et l'État. L'affaire fait alors grand bruit dans les médias et dans l'écosystème des acteurs des politiques de la santé (Antoniotti *et al.* 2002 ; Entretien, représentant, association d'usagers, 19/07/2022). Enfin, l'année 1998 signe également le déploiement national de la Carte Vitale, dont l'objectif est d'informatiser massivement les feuilles de soins papier. Dans le cadre de cette mise en œuvre, la CNIL est consultée, et c'est d'ailleurs suite aux considérations liées à la Carte Vitale que l'on retrouve le positionnement de la CNIL relatif à la Loi Informatique et Libertés et à l'accès aux données de santé dans son rapport d'activité de 1998.¹⁹ Une recherche menée sur Europresse pour l'année 1998²⁰ montre également que le déploiement de la Carte Vitale est effectivement associé à des discours sur les risques de perte de confidentialité et à un risque de ‘dérive commerciale’, au profit des assureurs complémentaires, des laboratoires pharmaceutiques et des éditeurs de logiciels.²¹ Ces risques auraient été notamment mis en évidence par Maurice Viennois et Jean-Pierre Michel dans un rapport remis à la CNIL la même année.²²

Enfin, l'informatisation et la création de fichiers nationaux dans d'autres secteurs de politique publique contribuent à cristalliser l'idée d'un État désireux de ‘ficher’ l'intégralité de sa population. C'est par exemple le cas des débats autour du STIC (Système de traitement des infractions constatées) de la police nationale ou de l'utilisation du NIR (Numéro d'inscription au répertoire, autrement appelé ‘Numéro de sécurité sociale’) par le service des impôts pour centraliser les données des individus. La fin des années 1990 semble donc marquée par de nombreux débats relatifs à l'informatisation et au droit des personnes, ce qui a pu contribuer la CNIL à repenser son positionnement initial sur les bases de données de santé, que ce soit au regard du PMSI ou du SNIIRAM.

¹⁹ CNIL (1998). *19^{ème} rapport d'activité 1998*. Paris : La Documentation Française, p.111

²⁰ Mot clé ‘CNIL’, recherche effectuée sur la période du 01/01/1998 au 31/12/1998, sur les sources de presse française répertoriées sur Europresse (158 résultats)

²¹ Libération. ‘La carte à puce santé, sous surveillance. La Cnil souligne le risque de "piratage" des données’. 9 juillet 1998 ; Le Progrès. ‘Carte de santé : vitales questions’. 28 mars 1998 ; Le Figaro (économie). ‘Carte Vitale : bouleversements et risques selon la Cnil’. 28 mars 1998

²² Le Progrès. ‘Carte de santé : vitales questions’. 28 mars 1998

Des données différentes, détenues par des acteurs distincts

Au cours de l'enquête de terrain, il est également apparu que les données du PMSI étaient régulièrement présentées comme étant moins 'sensibles' que celles du SNIIRAM. Cela est peut-être dû au fait que ces données sont moins nombreuses et concernent une petite fraction de la population (celle qui se rend à l'hôpital), et non l'entièreté de la population française comme c'est le cas pour le SNIIRAM. Le fait que les données du PMSI aient, pendant longtemps, été accessibles seulement sur une année a peut-être également contribué à diminuer le risque perçu associé à l'usage de ces données.

Enfin, il ne faut pas oublier que le SNIIRAM est avant tout un produit de la CNAMTS, tandis que le PMSI est un produit issu des directions du Ministère de la Santé, dont la gestion est ensuite confiée à une agence publique à partir de 2000. Le risque associé à l'usage des données de santé peut ainsi largement différer selon les convictions des personnalités en fonction, et il n'est alors pas étonnant d'observer des divergences de points de vue entre ces deux entités. Cette fragmentation entre les deux systèmes reflète les oppositions classiques de la gouvernance du système de santé, avec médecine de ville et CNAMTS d'un côté, et hôpital et État de l'autre. Pour ces deux acteurs (CNAMTS et État), c'était 'chacun sa base' et 'charbonnier maître chez soi' (Entretien, fonctionnaire, 14/04/2022).

Dans un premier temps cependant, ces différences entre le PMSI et le SNIIRAM ne vont pas donner lieu à une réglementation distincte, qui se matérialisera plutôt au début des années 2000 avec les premières mises en œuvre effectives du SNIIRAM. L'ensemble des affaires de la fin des années 1990 (palmarès des hôpitaux et débats associés à la construction du SNIIRAM) conduit plutôt à développer le premier niveau de réglementation de l'accès aux données de santé, qui met sur le même plan le SNIIRAM et le PMSI.

La modification de la loi LIL et l'introduction de la CNIL dans la gouvernance opérationnelle des données de santé (1998)

À la suite de ces différents événements, la CNIL se penche de nouveau sur la question de l'accès aux données du PMSI et du SNIIRAM (bien que celui-ci ne soit pas encore fonctionnel), et c'est alors qu'est prise la décision de soumettre cet accès à l'octroi d'une autorisation préalable par la CNIL (Entretien, juriste spécialiste des données de santé, 15/06/2022). C'est précisément à cette occasion qu'est instituée la doctrine juridique actuelle de l'accès aux données de santé, qui fait son apparition via la création d'un Chapitre V ter intitulé 'Traitement des données personnelles de santé à des fins d'évaluation ou d'analyse des activités de soins et de prévention' dans la Loi Informatique et Liberté,²³ supposé s'appliquer plus spécifiquement aux accès au SNIIRAM et au PMSI.²⁴ Cet article est

²³ Ce chapitre est introduit par l'article 41 de la Loi n° 99-641 du 27 juillet 1999 portant création d'une couverture maladie universelle

²⁴ Bras, Pierre-Louis et André Loth (2013). *Rapport sur la gouvernance et l'utilisation des données de santé*. Paris : publications de la DREES, p.31

essentiellement centré sur la régulation de l'accès aux données de santé, avec l'édiction de règles relativement strictes :

Article 40-12

Les données issues des systèmes d'information visés à l'article L. 710-6 du code de la santé publique, celles issues des dossiers médicaux détenus dans le cadre de l'exercice libéral des professions de santé, ainsi que celles issues des systèmes d'information des caisses d'assurance maladie, ne peuvent être communiquées à des fins statistiques d'évaluation ou d'analyse des pratiques et des activités de soins et de prévention que sous la forme de statistiques agrégées ou de données par patient constituées de telle sorte que les personnes concernées ne puissent être identifiées.

Il ne peut être dérogé aux dispositions de l'alinéa précédent que sur autorisation de la Commission nationale de l'informatique et des libertés dans les conditions prévues aux articles 40-13 à 40-15. Dans ce cas, les données utilisées ne comportent ni le nom, ni le prénom des personnes, ni leur numéro d'inscription au Répertoire national d'identification des personnes physiques.

Article 40-13

Pour chaque demande, la commission vérifie les garanties présentées par le demandeur pour l'application des présentes dispositions et, le cas échéant, la conformité de sa demande à ses missions ou à son objet social. Elle s'assure de la nécessité de recourir à des données personnelles et de la pertinence du traitement au regard de sa finalité déclarée d'évaluation ou d'analyse des pratiques ou des activités de soins et de prévention. Elle vérifie que les données personnelles dont le traitement est envisagé ne comportent ni le nom, ni le prénom des personnes concernées, ni leur numéro d'inscription au Répertoire national d'identification des personnes physiques. En outre, si le demandeur n'apporte pas d'éléments suffisants pour attester la nécessité de disposer de certaines informations parmi l'ensemble des données personnelles dont le traitement est envisagé, la commission peut interdire la communication de ces informations par l'organisme qui les détient et n'autoriser le traitement que des données ainsi réduites.

La commission détermine la durée de conservation des données nécessaires au traitement et apprécie les dispositions prises pour assurer leur sécurité et la garantie des secrets protégés par la loi. (Loi Informatique et Libertés, version du 28 juillet 1999)

Le fait que la CNIL doive désormais donner son autorisation ne signifie pas que les données ne sont plus accessibles ; d'ailleurs, les journalistes du palmarès de 1998 continuent d'y avoir accès (Malye and Vincent 2015). À ce moment-là les données annuelles du PMSI sont transmises sur disque dur ou sur disquette, directement remis au demandeur, moyennant finance. Quant à l'autorisation de la CNIL, elle était délivrée après vérification de l'adéquation entre les finalités annoncées de l'étude et les données demandées. Il n'existe cependant pas de doctrine claire quant aux finalités autorisées pour l'usage de ces données, la LIL faisant simplement référence aux 'fins statistiques d'évaluation ou d'analyse des pratiques et des activités de soins et de prévention'. Un entretien avec une employée de la CNIL ayant travaillé sur ces dossiers dans les années 2000 indique cependant que la notion 'd'intérêt public' influait sur la décision d'autorisation :

'Alors donc qu'est-ce qu'on vérifiait, la finalité, l'organisme qui voulait les mener pour **éviter les finalités qui n'étaient pas d'intérêt public**... Le fait donc que la donnée était très indirectement, pas de nom, pas de prénom, pas de numéro de Sécu, pas de date de naissance complète, donc c'était ces

données-là qu'on vérifiait, et il n'y avait pas d'avis scientifique... antérieur. Alors [...] la notion d'intérêt publique, c'était plutôt dans le contrôle de la mission de l'organisme, il y avait des finalités quand même qui faisaient qui faisaient bondir, enfin, qui pouvaient effrayer. Donc oui, y avait quand même un regard, dans l'esprit, c'est comme on protégeait les données de santé on avait un regard très particulier sur la sécurité depuis le départ. **C'était écrit nulle part, mais on le faisait.**' (Entretien, juriste spécialiste des données de santé, 15/06/2022)

La création et la mise en place des bases de données de santé centralisées qui s'opèrent au cours des années 1990 semblent ainsi s'accompagner, suite à l'année 1998, d'une volonté d'encadrement relativement stricte, puisque des dispositions précises sont prises en lien avec la protection des données de santé (modification de la Loi Informatique et Libertés et autorisation préalable de la CNIL). En parallèle, la question de l'ouverture de ces données à des tiers externes et de leur potentiel en matière de recherche et d'analyse ne semblent pas avoir une place prépondérante sur l'agenda, puisque l'usage de ces données est surtout pensé en lien avec la gestion financière de l'assurance maladie et des hôpitaux.

Tableau. La régulation de l'accès aux données de santé dans les années 1990

Règles d'accès	Gouvernance opérationnelle	Gouvernance stratégique	Enjeux pratiques
Evolution par rapport au début des années 1990			
Fermeture	-	-	-
Fermeture : Nécessité d'obtenir l'autorisation de la CNIL pour accéder au PMSI et au SNIIRAM, notion 'd'intérêt public'	Avant la CNIL, la CADA se charge par défaut de réguler l'accès aux données du PMSI. La loi de 1998 érige la CNIL en unique acteur de la gouvernance opérationnelle	-	-

Tableau. Les accès aux bases de données de santé selon les types d'acteurs (années 1990)

	PMSI (CNIL)	SNIIRAM (CNIL) ²⁵	SNIIRAM accès de droit (jeux de données)	SNIIRAM accès de droit (total)
Acteurs publics (Ministères et services)	Oui	Oui	N/A	N/A

²⁵ À la fin des années 1990, le SNIIRAM n'est pas encore mis sur pied ; néanmoins, ce sont théoriquement les règles de la LIL qui s'appliquent.

Agences publiques	Oui	Oui	N/A	N/A
Professionnels de santé	Oui	Oui	N/A	N/A
Organisations de patients	Oui	Oui	N/A	N/A
Assurances complémentaires	Oui	Oui	N/A	N/A
Laboratoires pharmaceutiques	Oui	Oui	N/A	N/A

II. Une ouverture ‘en sourdine’ portée par l’Institut des données de santé (années 2000)

Au cours des années 2000, plusieurs dynamiques ambivalentes et parfois contradictoires vont mener à de profondes transformations de l'accès aux données de santé. Ces dynamiques sont de deux ordres. On observe, tout d'abord, une différenciation nette de la régulation entre le PMSI et le SNIIRAM, cette dernière base de données se voyant doter d'une gouvernance spécifique et de règles d'accès plus strictes. Cette première dynamique de restriction de l'accès au SNIIRAM va néanmoins être contrariée dès 2004 par la création de l'Institut des Données de Santé (IDS), qui intervient dans le cadre d'un compromis plus global entre l'État et les assurances complémentaires. Dès lors, ce sont les relations de pouvoir entre la CNAMTS, l'IDS et l'État qui vont structurer la régulation de l'accès aux données de santé jusqu'à la fin des années 2000. Malgré les réticences de la CNAMTS, l'IDS parviendra à se positionner en figure de proue de l'ouverture des données de santé, en contribuant à une ouverture lente mais progressive des accès au SNIIRAM.

a) La différenciation des règles d'accès applicables au SNIIRAM

Alors que le PMSI est mis en œuvre dans les années 1990, il faut plutôt attendre le début des années 2000 pour voir apparaître les premiers usages du SNIIRAM, qui n'interviennent qu'à l'année 2004.²⁶ Avec le développement du SNIIRAM, la régulation de l'accès aux données de santé se différencie, et l'on distingue alors les règles régissant le SNIIRAM d'un côté, et celles régissant le PMSI de l'autre.

La stabilisation des règles d'accès au PMSI

Concernant le PMSI, les règles d'accès restent celles établies en 1998 au moment de la modification de la loi LIL. L'accès aux données du PMSI est ainsi conditionné par l'obtention d'une autorisation CNIL, généralement délivrée dans la mesure où les données sont bien anonymisées et où l'accès aux données semble justifié au vu de l'objectif du demandeur. Concrètement, de nombreux acteurs peuvent ainsi

²⁶ Suhard, Véronique (2021). *L'ouverture des données de santé à la recherche en France. Bibliographie thématique*. Paris : Publications de l'IRDES, Synthèses et Bibliographies, p.10 ; Article 21 de la Loi n° 98-1194 du 23 décembre 1998 de financement de la sécurité sociale pour 1999.

accéder au PMSI, à condition de pouvoir en payer le prix, l'accès aux données étant payant pour les acteurs privés. Les demandes d'accès se font à partir de l'année 2000 auprès de l'Agence technique de l'information sur l'hospitalisation (ATIH), établissement public à caractère administratif qui rassemble les différents organismes précédemment en charge du PMSI, dont la 'mission PMSI' du Ministère de la Santé.²⁷ L'ATIH a alors explicitement la mission de s'assurer de l'accessibilité au tiers' du PMSI, et doit également contribuer à la réalisation de certains traitements de données.²⁸

L'évolution contrastée des règles d'accès au SNIIRAM

La régulation du SNIIRAM prend quant à elle une tout autre direction. En effet, des règles spécifiques d'accès au SNIIRAM sont adoptées par arrêté, et connaissent des évolutions tout au long des années 2000. L'instance qui alimente directement ces évolutions (et donc la rédaction des arrêtés SNIIRAM) est le comité d'orientation et de pilotage de l'information inter-régimes (COPIIR), créé au moment de la mise en place du SNIIRAM.²⁹ C'est ainsi ce comité qui occupe la fonction de 'gouvernance stratégique' des données de santé au début des années 2000. Sa mission officielle est d'orienter le Ministère de la santé en matière de gestion de la donnée de santé, en particulier au regard du SNIIRAM (Entretien, ex-agent CNAMTS, 22/09/2022). Ce COPIIR est composé des 'producteurs de données' du SNIIRAM, à savoir : les régimes d'assurance obligatoire, l'État et les professionnels de santé, qui possèdent chacun un tiers des voies.³⁰

Le premier arrêté de mise en œuvre du SNIIRAM est publié le 11 avril 2002.³¹ Si certains organismes (principalement à caractère public) se voient octroyer un droit d'accès aux données agrégées issues du SNIIRAM,³² toute demande d'accès aux données individuelles anonymisées doit faire l'objet d'un accord délivré par la CNIL, de manière similaire au PMSI.³³

²⁷ Les deux autres organismes sont le pôle d'expertise et de référence nationale des nomenclatures de santé (Perms), et le Centre de traitement des informations du PMSI (CTIP).

²⁸ Décret n° 2000-1282 du 26 décembre 2000 portant création de l'Agence technique de l'information sur l'hospitalisation et modifiant le code de la santé publique

²⁹ Comité d'Experts de l'IDS (2010). *Livre blanc 2010 du Comité d'Experts. 7 propositions au service de la recherche*. Charenton-le-Pont : publications de l'IDS, p.18

³⁰ Bras, Pierre-Louis et André Loth (2013). *Rapport sur la gouvernance et l'utilisation des données de santé*. Paris : publications de la DREES., p.48

³¹ Bras, Pierre-Louis et André Loth (2013). *Rapport sur la gouvernance et l'utilisation des données de santé*. Paris : publications de la DREES, p.31

³² Agences Régionales de l'Hospitalisation (ARH), Ministère de l'économie, Unions régionales de médecins libéraux

³³ En théorie, les autorisations d'accès devaient également dans certains cas faire l'objet d'un avis du Comité consultatif pour le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS), créé en 1994 lors de l'ajout du chapitre V bis à la Loi Informatique et Libertés. Il semble cependant qu'en pratique ce Comité n'ait pas toujours été consulté à hauteur de ses attributions (Bonaïti-Pellié *et al.* 2009). Une de nos enquêtées indique par ailleurs que la consultation de ce comité était obligatoire pour les données non anonymisées, mais pas pour les données pseudonymisées du SNIIRAM et du PMSI (Entretien, juriste spécialiste des données de santé, 15/06/2022). Enfin, lorsque les recherches impliquent la personne humaine, celles-ci devaient (et doivent toujours) également faire l'objet d'un avis d'un comité de protection des personnes.

Cependant, trois ans plus tard, et alors que les données du SNIIRAM commencent tout juste à être disponibles, un autre arrêté est adopté (arrêté du 20 juin 2005). Si celui-ci élargit la liste des potentiels utilisateurs, il interdit dans le même temps l'accès au SNIIRAM aux organismes à but lucratif (qui peuvent néanmoins disposer d'un accès via un intermédiaire public). Avec l'ajout de cette interdiction spécifique, l'accès au SNIIRAM devient *de facto* plus difficile que pour le PMSI. En effet, il ne s'agit plus simplement de poursuivre un objectif relevant de l'intérêt général (condition 'non officielle' de l'accès au PMSI) mais d'être un organisme appartenant à la sphère publique ou associative pour pouvoir accéder aux données du SNIIRAM.

La CNAMTS : une position de réticence face à l'ouverture des données

Un de nos enquêtés indique que cette interdiction d'accès visait en particulier les industries pharmaceutiques et les assureurs complémentaires (Entretien, ex-agent CNAMTS, 19/04/2022), avec qui la CNAMTS refuse systématiquement toute coopération (Entretien, agent CNAMTS, 21/04/2022). Ce positionnement est par ailleurs renforcé par ce que certains appellent un 'réflexe propriétaire'³⁴ de la CNAMTS sur les données du SNIIRAM. Ce 'réflexe propriétaire', qui désigne la réticence de principe au partage des données, trouve sans doute son origine dans plusieurs sources. En effet, comme nous l'avons vu précédemment, la mise en place du SNIIRAM a provoqué d'importants débats avec la CNIL au regard de la protection des données personnelles ; il est donc probable que dans ce cadre la CNAMTS soit particulièrement prudente quant au partage de ces mêmes données avec d'autres acteurs (Entretien, ex-agent CNAMTS, 22/09/2022). Ensuite, le 'réflexe propriétaire' est également corrélé à l'investissement dans la construction de la base et aux possibilités d'exploitation qui y sont associées. Le travail fastidieux et de longue haleine de construction d'une base de données peut ainsi pousser les acteurs à en restreindre l'accès, au moins le temps de pouvoir profiter eux-mêmes des bénéfices de l'exploitation de cette base (Entretien, ex-agent CNAMTS, 22/09/2022). Ce type d'argument prend par ailleurs une dimension particulière au regard du positionnement plus global de la CNAMTS dans le système de santé : cette organisation est en effet en constante lutte avec les autres acteurs du système (industries pharmaceutiques, professionnels de santé, assurances complémentaires), et ses agents anticipent ainsi le risque que leurs propres données soient utilisées 'contre' la CNAMTS dans le cadre de négociations (Entretien, ex-agent CNAMTS, 22/09/2022).

Ces différentes raisons poussent la CNAMTS à adopter une position de réticence vis-à-vis de l'ouverture des données de santé, qu'elle souhaite limiter à certains acteurs publics, en particulier les agences sanitaires et certains acteurs de la recherche (Entretien, ex-agent CNAMTS, 22/09/2022 ; Entretien, ex-agent CNAMTS, 19/04/2022 ; Entretien, agent CNAMTS, 21/04/2022). Ce

³⁴ Entretien, ex-agent CNAMTS, 22/09/2022 ; Entretien, fonctionnaire et agent CNAMTS, 20/07/2022 ; Entretien, représentant, secteur des assurances complémentaires, 16/06/2022

positionnement sera celui tenu par la CNAMTS tout au long de l'évolution de la régulation de l'accès aux données de santé.

b) La loi du 13 août 2004 relative à l'Assurance Maladie et la création de l'Institut des Données de Santé

Au début des années 2000, la mise en place du SNIIRAM s'accompagne donc de l'adoption d'une réglementation spécifique à cette base de données, qui reflète le positionnement de la CNAMTS en faveur d'une restriction de l'accès aux données de santé.

Un contexte de transformation des relations entre assurance maladie obligatoire et assurances complémentaires

En 2004, un évènement va bousculer cet équilibre initial. En effet, cette année est marquée par la création d'une nouvelle organisation dans le secteur des données de santé, à savoir l'Institut des données de santé (IDS). Contrairement à ce que l'on pourrait penser, la création de cet Institut n'est pas liée à l'adoption par l'État d'une politique volontariste en matière d'ouverture des données de santé. L'IDS représente plutôt une mesure parmi d'autres dans un contexte plus général de négociations entre l'État, l'Assurance Maladie obligatoire et les assurances complémentaires. Dès la fin des années 1990 et au début des années 2000, on observe en effet de nouvelles formes d'articulation entre assurance maladie obligatoire et assurance maladie complémentaire, parfois qualifiées de 'privatisation' de l'assurance maladie (Benamouzig 2012). Ces changements traduisent, en partie, le principe de 'co-régulation' du risque maladie par l'assurance obligatoire et les assurances privée, principe qui prévaut au début des années 2000 suite à la publication d'un rapport sur ce sujet par Jean-François Chadelat (Benamouzig 2012), inspecteur général des affaires sociales et 'proche du pouvoir' (Entretien, agent CNAMTS, 21/04/2022).³⁵ Dans ce cadre, l'État cherche à organiser ces nouvelles relations entre assurance maladie obligatoire et assurances complémentaires, et c'est précisément dans le cadre de ces négociations que s'inscrit l'IDS.

Cet institut fait en effet partie d'un paquet de mesures plus générales autour de la loi du 13 août 2004, portée par Philippe Douste Blazy alors Ministre de la Santé. Dans ce cadre, la création de l'IDS apparaît comme le résultat d'un compromis passé avec les assurances complémentaires et leurs représentants, en particulier la Fédération Nationale de la Mutualité Française – FNMF (Entretien, agent CNAMTS, 21/04/2022, voir également pour un compte-rendu plus général : Benoît and Coron 2018). On observe ainsi que la création d'un 'établissement public chargé de rassembler les données de santé [...] d'en assurer le traitement, d'en garantir l'accès à tous les acteurs impliqués dans la gestion du risque' fait effectivement l'objet de la 12^{ème} proposition de la résolution adoptée par la FNMF

³⁵ Groupe de travail de la Commission des comptes de la Sécurité sociale présidé par J-F. Chadelat (2003). *La répartition des interventions entre les assurances maladie obligatoires et complémentaires en matière de dépenses de sante*. Rapport au Ministre de la Santé J-F. Mattei, Paris : publications de la Commission des Comptes de la Sécurité Sociale, 33p.

lors de son congrès de Toulouse du 14 juin 2003.³⁶ Le souci des assureurs complémentaires est alors de ne plus être des ‘payeurs aveugles’³⁷ et de participer à la ‘cogestion’ du système de santé. La création de l’IDS, et de sa mission d’ouverture de l’accès aux données de santé, est ainsi à ses débuts très ‘tournée organismes complémentaires’ (Entretien, agent CNAMTS, 21/04/2022) : c’est pour permettre aux assurances complémentaires de participer à la ‘co-régulation’ du risque maladie que cet organisme est créé.

‘Et donc ensuite, l’Institut des données de santé, alors oui l’Institut des données de santé c’est, quelque part né, d’une **proposition de la Mutualité Française**, la fameuse, les fameuses résolutions du congrès de Toulouse, donc 2003, on est à un moment donné finalement, oui c’est ça 2003, qui proposait à la fois la création des différentes unions nationales, donc pour les complémentaires santé, pour les professionnels de santé, et la création d’un établissement public des données de santé. Donc ces propositions ont été retenues, dans la loi, mais on n’en avait pas défini les contenus à ce moment-là on avait juste posé les principes.’ (Entretien, organisme assurance complémentaire, 16/06/2022)

‘Dès le départ la **FNMF revendique d’avoir accès aux données**, parce qu’elle veut, la revendication historique de la FNMF c’était de **cogérer le système**, et donc pour cogérer il faut avoir les données’ (Entretien, agent CNAMTS, 21/04/2022)

C’est d’ailleurs Jean-François Chadelat, l’auteur du rapport emblématique sur la notion de ‘co-régulation’, qui s’occupe directement de la mise en œuvre de l’IDS ; il en choisit notamment le directeur, réputé inspirer ‘confiance’ aux assurances complémentaires (Entretien, agent CNAMTS, 21/04/2022). En parallèle, le Président de l’IDS, Christian Babusiaux, est également étroitement connecté au monde des assurances complémentaires, ayant auparavant dirigé une mission sur l’accès des assureurs complémentaires aux données des assurés pour le compte du Ministre de la Santé Jean-François Mattei (Entretien, fonctionnaire, 14/04/2022 ; Entretien, représentant, secteur des assurances complémentaires, 16/06/2022).³⁸ L’IDS se met donc en place dans un contexte où l’État cherche à se positionner en ‘acteur central’ et à ‘tenir un équilibre’ entre la CNAMTS d’un côté et les assureurs complémentaires de l’autre ; lors de la création de l’IDS, l’équilibre du pouvoir entre ces différents acteurs était d’ailleurs l’une des consignes du cabinet ministériel de Xavier Bertrand, Ministre de la santé de 2005 à 2007 (Entretien, fonctionnaire et agent CNAMTS, 20/07/2022).

³⁶ Fédération Nationale de la Mutualité Française. *Vingt-cinq mesures pour réduire les inégalités de santé*. Résolution adoptée le 14 juin 2003 au Congrès de Toulouse.

³⁷ Babusiaux Christian, Louis Bréas, Laurence Eslous et Dominique Thouvenin (2003). *L'accès des assureurs complémentaires aux données de santé des feuilles de soins électroniques*. Rapport remis au Ministre de la Santé, de la Famille et des Personnes handicapées, p.14.

³⁸ Babusiaux Christian, Louis Bréas, Laurence Eslous et Dominique Thouvenin (2003). *L'accès des assureurs complémentaires aux données de santé des feuilles de soins électroniques*. Rapport remis au Ministre de la Santé, de la Famille et des Personnes handicapées.

L'enrichissement des missions de l'Institut des Données de Santé

L'Institut des données de santé est donc créé par la loi du 13 août 2004, avec les objectifs suivants :

‘assurer la cohérence et de veiller à la qualité des systèmes d'information utilisés pour la **gestion du risque maladie** et de veiller à la **mise à disposition de ses membres**, de la Haute autorité de santé, des unions régionales des médecins exerçant à titre libéral ainsi que d'organismes désignés par décret en Conseil d'Etat, à des **fins de gestion du risque maladie** ou pour des **préoccupations de santé publique**, des données issues des systèmes d'information de ses membres, dans des conditions garantissant l'anonymat fixées par décret en Conseil d'Etat pris après avis de la CNIL³⁹

Bien que la formulation de l'article de loi ne soit que peu détaillée, on comprend que le rôle principal de l'IDS, qui doit prendre la forme d'un groupement d'intérêt public (GIP), est de permettre la ‘mise en commun’ des données de ses membres, et de favoriser le traitement et l'usage de ces données par ces mêmes membres à des fins de ‘gestion du risque maladie’ ou pour des ‘préoccupations de santé publique’. Dans la loi du 13 août 2004, la liste des membres de l'IDS est relativement resserrée, et reflète le compromis établi avec la FNMF : État, CNAMTS, Union nationale des organismes d'assurance maladie complémentaire (UNOCAM) et Union nationale des professionnels de santé (UNPS).⁴⁰

Si la forme que doit prendre l'IDS suscite quelque débats,⁴¹ un accord entre les différents acteurs pose l'IDS comme une plateforme de facilitation de l'accès aux données de santé pour ses membres, avec l'objectif de créer et de mettre en œuvre des procédures efficaces permettant cet accès.

Dès l'arrêté de mise en œuvre de l'IDS, on peut noter un enrichissement des missions de cette organisation.⁴² En effet, en plus d'assurer et de favoriser la mise en commun des données de santé de ses membres, le bureau de l'IDS a la charge d'émettre un avis sur les demandes d'accès aux données de santé du SNIIRAM,⁴³

³⁹ Article 64, Loi n° 2004-810 du 13 août 2004 relative à l'assurance maladie

⁴⁰ En parallèle, la loi indique la mise à disposition des données doit également s'effectuer à la faveur de la Haute Autorité de Santé (HAS), de l'Union Régionale des Professionnels de Santé (URPS) et éventuellement d'autres organismes devant être désignés en Conseil d'État.

⁴¹ Au début des discussions est proposée l'idée de créer un organisme qui disposerait des copies des bases de données et serait capable de produire lui-même des études, à l'instar de l'INSEE ou de l'actuel Health Data Hub. Ce projet est cependant rapidement retoqué à la faveur d'une proposition plus modeste de plateforme facilitant le partage des données. Sources : Entretien, agent CNAMTS, 21/04/2022 ; Entretien, représentant, secteur des assurances complémentaires, 16/06/2022

⁴² Arrêté du 30 avril 2007 portant approbation du groupement d'intérêt public « Institut des données de santé ». Il semblerait que cet arrêté ait été rédigé par l'équipe de préfiguration du GIP, celle-là même qui sera à la suite à la tête de l'IDS.

⁴³ Cet avis est adopté par le bureau de l'IDS qui représente ses membres, après avis du ‘comité d'experts’ de l'IDS, dont les membres sont, en 2016, les suivants (à une ou deux exceptions près – notamment la présence de Marcel Goldberg – composition similaire à celle de 2010, voir : Comité d'Experts de l'IDS (2010). *Livre blanc 2010 du Comité d'Experts. 7 propositions au service de la recherche*. Charenton-le-Pont : publications de l'IDS, 61p.) : Annick Alperovitch, directrice de recherche à l'INSERM, membre du Comité Consultatif National

avis préalable à l'autorisation de la CNIL. L'IDS se voit ainsi assigné une fonction de *gatekeeper* de l'accès aux données de santé, ce qui ajoute *de facto* une étape supplémentaire au processus d'autorisation d'accès aux données de santé du SNIIRAM. En parallèle, l'IDS se voit attribuer tout un ensemble de missions liées à son statut de 'facilitateur' du partage de données, comme par exemple la possibilité de créer des 'services pour étudier les données mises en commun'.

Avec cet arrêté, le rôle de l'IDS dans l'ouverture de l'accès aux données de santé est plus affirmé que ce qui avait été annoncé initialement dans les 'quatre lignes' (Entretien, agent CNAMTS, 21/04/2022) de la loi de 2004. Ceci est d'autant plus vrai que cet arrêté donne la possibilité à l'IDS 'd'accepter de nouveaux membres par décision de l'assemblée générale' (article 6). Or, la loi, qui ne porte pas de politique spécifique d'ouverture des données de santé, prévoyait plutôt un nombre limité de membres, dans le cadre du compromis arrêté avec les organismes d'assurance complémentaire. La possibilité donnée à l'IDS d'élargir le nombre de ses membres (qui sera largement utilisée, voir II.c) va permettre à cet organisme de conforter son statut de figure de proue de l'ouverture des données de santé en France.

Malgré cela, il faut garder à l'esprit que l'IDS demeure une petite structure (surtout par contraste avec l'actuel Health Data Hub), avec peu d'employés et un budget annuel légèrement inférieur à 1 million d'euros ('Cinq piou-pious, six piou-pious hein, six piou-pious pour faire tout ça', Entretien, agent CNAMTS, 21/04/2022, voir également Entretien, représentant professionnels de santé, 23/05/2022)

La CNAMTS et l'Institut des Données de Santé : une relation conflictuelle

De toute évidence, la création de l'IDS n'a pas été sans provoquer des remous institutionnels parmi les acteurs du système de santé et de la régulation de l'accès aux données de santé. En particulier, ce positionnement stratégique de chef de file de l'ouverture des données de santé entre en contradiction avec le pilotage stratégique du SNIIRAM, alors assuré par le COPIIR, et au sein duquel l'assurance maladie est, avec l'État, un décisionnaire central (Entretien, ex-agent CNAMTS, 22/09/2022 ; Entretien, agent CNAMTS, 21/04/2022).

d'Éthique ; Pascal Auquier, professeur à la faculté de Médecine de Marseille, directeur de l'EA2279, Responsable du service de santé publique et de l'information médicale, CHU de Marseille ; Brigitte Dormont, professeur à l'université Paris Dauphine, responsable de la Chaire Santé de la Fondation du Risque ; Martial Fraysse, Président du conseil régional des pharmaciens d'Ile de France, membre du Conseil National de l'Ordre des Pharmaciens ; Jean-Pierre Le Gléau, inspecteur général de l'INSEE, chef du département de la coordination statistique ; Claude Le Pen, professeur, université Paris-Dauphine ; Jacques Lucas, Vice-Président du Conseil National de l'Ordre des Médecins ; René Padiou, inspecteur général honoraire de l'INSEE, président de la commission de déontologie de la société française de statistiques ; Gérard de Pouvourville, professeur titulaire de la chaire Économie et Gestion de la Santé à l'ESSEC ; Catherine Quantin, professeur à la faculté de Médecine de Dijon, responsable du service de bio statistique et informatique médicale, CHU de Dijon ; Roxane Silberman, directrice de recherche au CNRS (Ecole d'Économie de Paris), Secrétaire générale du Comité de Concertation pour les Données en Sciences Humaines et Sociales (CCDSHS)

‘[...] on a eu des très **mauvais rapports** entre la CNAMTS et l’IDS, très très mauvais, ça a été la **guerre** [...]’ (Entretien, ex-agent CNAMTS, 22/09/2022)

Comme nous l’avons vu, ce COPIIR avait par ailleurs adopté une posture relativement stricte en matière d’accès aux données de santé, en interdisant l’accès au SNIIRAM pour les organismes à but lucratif (voir II.a). Le principe même de l’IDS entre alors en contradiction directe avec cette posture, puisque l’objectif est précisément de permettre aux assureurs complémentaires d’accéder aux données du SNIIRAM (tout en partageant également leurs propres données). La création de l’IDS agit donc comme un élément perturbateur, en donnant une base institutionnelle à des acteurs en faveur de l’ouverture des données de santé, qui étaient auparavant dispersés et minoritaires dans le système décisionnel.⁴⁴

Cette opposition se manifeste au quotidien dans les relations entre l’IDS et la CNAMTS. Ainsi, le COPIIR ne se réunit jamais plus de deux fois par an, ce qui contribue à ralentir fortement le processus décisionnel et limite le potentiel d’ouverture du SNIIRAM (Entretien, ex-agent CNAMTS, 22/09/2022 ; Entretien, agent CNAMTS, 21/04/2022). Les blocages sont tels que la mise en place de l’IDS ne sera finalement effective qu’en 2006, après le 38^{ème} congrès de la FNMF, dont la tenue incite le Ministère de la Santé à revenir à la charge auprès de la CNAMTS pour finaliser la mise en œuvre de l’IDS (Entretien, agent CNAMTS, 21/04/2022). La CNAMTS est ainsi régulièrement ‘accusée’ d’être ‘assise sur un trésor’ sur lequel elle voudrait garder un contrôle exclusif (Entretien, ex-agent CNAMTS, 19/04/2022 ; Entretien, ex-agent CNAMTS, 22/09/2022).

À la fin des années 2000 et au début des années 2010, un point de débat finit de cristalliser les tensions entre la CNAMTS et l’IDS. En effet, l’IDS se positionne en faveur de l’accès aux données de santé par les industries pharmaceutiques dans le cadre des études post-Autorisation de Mise sur le Marché (Entretien, agent CNAMTS, 21/04/2022 ; Entretien, fonctionnaire et agent CNAMTS, 20/07/2022 ; voir également les rapports de l’IDS de l’année 2010⁴⁵), tandis que la CNAMTS est farouchement opposée à un tel accès (Entretien, ex-agent CNAMTS, 19/04/2022). Le débat porte alors plus précisément sur l’interdiction de l’accès à des organismes à but lucratif (arrêté SNIIRAM de 2005), que l’IDS trouve trop restrictif et voudrait remplacer par une condition de poursuite de l’intérêt général (qui peut s’appliquer indépendamment du type d’acteur considéré).⁴⁶

⁴⁴ Ces défenseurs de l’ouverture des données de santé étaient par exemple certaines personnes travaillant sur la construction des bases de données (la ‘secte’, Entretien, ex-fonctionnaire, 06/04/2022), les journalistes notamment ceux utilisant les données de santé pour leur palmarès, mais aussi certains acteurs (alors minoritaires) qui se positionnent en faveur d’un système de santé davantage libéral, managérial et ‘moderne’ comme Gilles Johannet, ancien directeur de la CNAM (Pierru 2004: 259–262)

⁴⁵ Institut des données de santé (2010). *Rapport au Parlement 2010. Enjeux et étapes*. Charenton-le-Pont : publications de l’IDS, 63p.

⁴⁶ Comité d’Experts de l’IDS (2010). *Livre blanc 2010 du Comité d’Experts. 7 propositions au service de la recherche*. Charenton-le-Pont : publications de l’IDS, pp.22-23

c) L'ouverture progressive des données du SNIIRAM

Malgré ces tensions et le pouvoir dont dispose la CNAMTS via le COPIR, l'attitude proactive de l'IDS (qui a par ailleurs pris soin d'embarquer une large palette d'acteurs à son bord : représentants de patients, professionnels de santé, fédérations hospitalières) permet, au cours des années 2000, d'engager une ouverture progressive des données du SNIIRAM. Cette ouverture est notamment permise par le soutien occasionnel d'acteurs hauts placés au sein du Ministère de la santé.

'En fait on a eu des **conjonctions positives** avec des gens au **gouvernement** et dans les **cabinets ministériels** qui étaient plutôt ouverts, et donc on a eu toute une phase au moment où on a construit l'IDS effectivement avec une **ouverture**, un vrai effort de dialogue, entre les différentes parties prenantes.' (Entretien, représentant, secteur des assurances complémentaires, 16/06/2022)

Le soutien de l'État n'est cependant pas uniforme ni constant (Entretien, secteur des assurances complémentaires, 15/06/2022), obligeant l'IDS à jouer en permanence sur les rapports de force entre la CNAMTS et l'État pour mettre en œuvre son agenda d'ouverture. Tout au long des années 2000, ce sont donc essentiellement des dynamiques institutionnelles, peu visibles dans le débat public, qui structurent la régulation de l'accès aux données de santé.

L'extension des finalités associées à l'usage du SNIIRAM

En analysant les différents arrêtés SNIIRAM adoptés au cours du temps, on observe en premier lieu un net enrichissement des finalités associées à l'usage des données de santé. Au-delà de l'attitude proactive de l'IDS, cet enrichissement témoigne également, pour tous les acteurs, de la 'découverte' progressive du potentiel d'une telle base de données (Entretien, ex-agent CNAMTS, 22/09/2022 ; Entretien, ex-fonctionnaire, 06/04/2022 ; Entretien, ex-agent CNAMTS, 19/04/2022). Non seulement les principaux objectifs sont plus nombreux, mais les sous-objectifs qui y sont rattachés tendent également à se multiplier, comme en témoigne la comparaison entre l'arrêté SNIIRAM de 2002 et celui de 2005.

Encadré : L'évolution des objectifs associés au traitement des données du SNIIRAM

Objectifs du SNIIRAM dans l'arrêté de 2002

Les traitements mis en œuvre dans le cadre du système national d'information inter-régimes de l'assurance maladie ont pour finalités :

1° *D'améliorer la qualité des soins, notamment par la comparaison des pratiques aux référentiels, au sens de l'article L. 162-12-15 du code de la sécurité sociale, et moyennes professionnelles ;*

2° *De contribuer à une meilleure gestion de l'assurance maladie, notamment par :*

- la connaissance des dépenses de l'ensemble des régimes d'assurance maladie par circonscription géographique, par nature de dépense, par catégorie de professionnels responsables de ces dépenses et par professionnel ou établissement ;

- l'évaluation des transferts entre enveloppes correspondant aux objectifs sectoriels de dépenses fixés, en fonction de l'objectif national de dépenses d'assurance maladie, dans le cadre de la loi annuelle de financement de la sécurité sociale ;

- l'analyse quantitative des déterminants de l'offre de soins et la mesure de leurs impacts sur l'évolution des dépenses d'assurance maladie ;

3° De transmettre aux prestataires de soins les informations pertinentes relatives à leur activité, à leurs recettes et, s'il y a lieu, à leurs prescriptions.

Objectifs du SNIIRAM dans l'arrêté de 2005

Les traitements mis en œuvre dans le cadre du système national d'information interrégimes de l'assurance maladie ont pour finalités :

1. *D'améliorer la qualité des soins, notamment par :*

- la comparaison des pratiques aux référentiels, accords de bons usages ou contrats de bonne pratique, au sens des articles L. 162-12-15, L. 162-12-17, L. 162-12-18, et L. 162-12-20 du code de la sécurité sociale, et moyennes professionnels ;
- l'évaluation des comportements de consommation de soins ;
- l'analyse des caractéristiques et des déterminants de la qualité des soins ;

2. *De contribuer à une meilleure gestion de l'assurance maladie, notamment par :*

- La connaissance des dépenses de l'ensemble des régimes d'assurance maladie par circonscription géographique, par nature de dépense, par catégorie de professionnels responsables de ces dépenses et par professionnel ou établissement ;
- l'évaluation des transferts entre enveloppes correspondant aux objectifs sectoriels de dépenses fixés, en fonction de l'objectif national de dépenses d'assurance maladie, dans le cadre de la loi de financement de la sécurité sociale ;
- l'analyse quantitative des déterminants de l'offre de soins et la mesure de leurs impacts sur l'évolution des dépenses d'assurance maladie ;

3. *De contribuer à une meilleure gestion des politiques de santé, notamment par :*

- l'identification des parcours de soins des patients ;
- le suivi et l'évaluation de l'état de santé des patients et leurs conséquences sur la consommation de soins ;
- l'analyse de la couverture sociale des patients ;
- la surveillance de la consommation de soins en fonction de différents indicateurs de santé publique ou de risque ;

4. *De transmettre aux prestataires de soins les informations pertinentes relatives à leur activité, à leurs recettes et, s'il y a lieu, à leurs prescriptions. »*

On note cependant que la question de l'ouverture de l'accès aux données ne représente pas encore un objectif à part entière, il s'agit plutôt de lister de manière la plus exhaustive possible les finalités pour lesquelles les données de santé peuvent être utilisées.

Au-delà de l'extension de la liste des objectifs associés au SNIIRAM, on observe durant cette période de profondes transformations de l'accès aux données de santé, que nous allons détailler à l'aide de la typologie présentée en introduction (règles d'accès, gouvernance stratégique, gouvernance opérationnelle, enjeux pratiques).

Les règles d'accès : une ouverture ambivalente

Au regard des règles d'accès tout d'abord, nous avons vu que le SNIIRAM se ferme aux organismes à but lucratif dès 2005, tandis qu'une nouvelle étape est ajoutée dans la procédure d'accès dès 2007 (nécessité d'obtenir l'avis de l'IDS avant autorisation de la CNIL). Ce faisant, le contrôle de l'accès aux données de santé tend à se renforcer (dynamique de fermeture). Cependant, dans le même temps, la création de l'IDS donne la possibilité à ses membres de bénéficier d'un accès 'de droit' à certaines données issues du SNIIRAM, contribuant ainsi à l'ouverture des données de santé. Ces accès de droit ne concernent cependant pas l'intégralité du SNIIRAM, mais dans un premier temps certaines bases de données 'thématisées'

extraites du SNIIRAM, réputées plus faciles à manipuler (les *datamarts*), et dans lesquelles les données sont agrégées au sein de catégories rendant difficile la réidentification des individus.⁴⁷ Dès 2005-2006, certains membres auront également accès à d'autres outils (Echantillon Généraliste de Bénéficiaires, tableaux de bord), que nous décrirons ci-après dans la partie dédiée aux enjeux pratiques.

Les membres de l'IDS sont par ailleurs, comme nous l'avons mentionné, relativement nombreux en comparaison de ceux qui étaient initialement désignés dans la loi du 13 août 2004 (État, caisses nationales d'assurance maladie, UNOCAM et UNPS). Dès les débuts de l'IDS sont ainsi associés les représentants de patients via le Collectif inter-associatif en santé (CISS), qui à l'époque est la structure la plus 'institutionnalisée' en matière de représentants des patients.⁴⁸ De même, on retrouve parmi les membres de l'IDS les fédérations hospitalières (Fédération des Établissements Hospitaliers et d'Aide à la Personne privés solidaires – FEHAP, Fédération hospitalière de France – FHF, la Fédération de l'Hospitalisation Privée – FHP), et la Fédération nationale des centres de lutte contre le cancer (FNCLCC).⁴⁹ Par contraste, les industries pharmaceutiques (et notamment le syndicat Les Entreprises du Médicament – LEEM) ne sont pas représentés dans cette gouvernance.

Parallèlement à cet accès de droit des membres de l'IDS, plusieurs arrêtés SNIIRAM sont adoptés au cours des années 2000 qui élargissent l'accès de certaines bases de données issues du SNIIRAM à différents acteurs (qui n'ont alors pas besoin de passer par la CNIL pour obtenir une autorisation).

L'arrêté de 2005 notamment, signé par Dominique Libault alors Directeur de la Sécurité Sociale et soutien de l'IDS (Entretien, agent CNAMTS, 21/04/2022), ouvre l'accès aux données agrégées (*datamart*) et à l'Echantillon Généraliste de Bénéficiaires à plusieurs organismes publics, tels que l'Institut de Veille sanitaire (InVS), l'Agence française de sécurité sanitaire des produits de santé (AFSSAPS), et le Haut Conseil pour l'avenir de l'assurance maladie (HCAAM). L'accès est également ouvert à certains membres nommés d'organismes de recherche, tels que l'Institut de recherche et de documentation en économie de la santé (IRDES), l'Institut national de la santé et de la recherche médicale (INSERM) et le Centre National de la Recherche Scientifique (CNRS). D'après Pierre-Louis Bras et André Loth, auteurs d'un rapport sur les données de santé publié en 2013, ce mouvement se poursuit jusqu'au début des années 2010, avec des nouveaux arrêtés SNIIRAM adoptés en 2008, 2011 et 2013.⁵⁰

⁴⁷ Pour assurer l'absence de possibilité de réidentification, différents pare-feux ont été mis en place, comme l'interdiction de croiser plusieurs variables sensibles (date de naissance, commune...), ou encore l'instauration d'un nombre minimal d'individus pour que le résultat soit visible par l'utilisateur.

⁴⁸ Bien que doté d'une structure très souple, le CISS est alors le seul organisme ayant vocation à représenter plusieurs associations de patients à la fois.

⁴⁹ Institut des données de santé (2008). *Rapport au Parlement : 2008 les premières réalisations*. Charenton-le-Pont : publications de l'IDS, p.38

⁵⁰ Bras, Pierre-Louis et André Loth (2013). *Rapport sur la gouvernance et l'utilisation des données de santé*. Paris : publications de la DREES, p.31

Parmi ces nombreux acteurs, le cas des organismes d'assurance complémentaires (notamment de l'UNOCAM) demeure ambigu. En effet, bien que membres de l'IDS, les fédérations d'assureurs ne disposent dans un premier temps que de l'accès aux données agrégées issues du SNIIRAM (*datamarts*, dont les données sont réputées anonymes). En ce qui concerne l'accès à l'Echantillon Généraliste de Bénéficiaires (EGB), qui contient un échantillon de données individuelles anonymisées, l'accès n'est autorisé pour les fédérations d'assureurs qu'au début des années 2010, après la mise en place des expérimentations MONACO dont l'objectif est de chaîner des données de l'assurance maladie obligatoire aux données de l'assurance maladie complémentaire.⁵¹ L'idée est alors de conditionner l'accès à l'EGB aux fédérations d'assureurs contribuant à enrichir le SNIIRAM avec leurs propres données de santé. Pour les organismes d'assurance eux-mêmes,⁵² l'accès est possible seulement aux données agrégées et à condition qu'ils participent au projet MONACO.⁵³ De plus, ces différents accès sont conditionnés par la signature d'une 'charte d'engagements entre l'assurance maladie obligatoire et les assureurs maladie complémentaires pour le partage et la mise en commun des données de remboursement agrégées ou individuelles anonymisées issues de leurs systèmes d'information'.⁵⁴

En termes de règles d'accès, on observe donc un mouvement d'ouverture assez net, avec l'élargissement progressif de la palette d'acteurs ayant un accès 'permanent' à certains jeux de données extraits du SNIIRAM via leur adhésion à l'IDS. En parallèle, l'interdiction d'accès aux organismes à but lucratif et l'ajout de la nécessité d'obtenir un avis de l'IDS préalable à l'autorisation de la CNIL pour les accès ponctuels contribue au renforcement du contrôle de l'accès aux données du SNIIRAM. En termes de règles d'accès, les dynamiques observées sont donc relativement contradictoires, et traduisent l'état des rapports des forces entre deux blocs antagonistes vis-à-vis de l'ouverture des données de santé, la CNAMTS et l'IDS. Du côté du PMSI, les règles d'accès demeurent, comme nous l'avons vu, inchangées pendant cette période.

La gouvernance opérationnelle et stratégique : multiplication des acteurs et concurrence

Concernant la gouvernance opérationnelle et la gouvernance stratégique, les années 2000 sont surtout marquées par une multiplication des acteurs, résultant en une concurrence institutionnelle entre plusieurs organismes (en particulier au regard de la régulation de l'accès au SNIIRAM).

L'IDS semble se situer (ou aimerait se situer) à l'intersection de ces deux types de gouvernances, à la fois en participant au processus d'octroi de l'autorisation d'accès

⁵¹ Ce projet MONACO fait suite aux 'expérimentations Babusiaux' qui avaient été lancées après les rapports de Christian Babusiaux de 2003. Ces expérimentations avaient pour objectif de tester la faisabilité de la transmission anonymisée des données de l'Assurance maladie obligatoire aux complémentaires. Source : Entretien, représentant, secteur des assurances complémentaires, 16/06/2022

⁵² En l'espèce : Axa, AG2R, Allianz, Groupama, Malakoff-Mederic, Pro-BTP, Swiss Life

⁵³ Institut des données de santé (2010). *Rapport au Parlement 2010. Enjeux et étapes*. Charenton-le-Pont : publications de l'IDS, p.9

⁵⁴ Charte disponible dans le rapport de l'IDS de 2010 : Institut des données de santé (2010). *Rapport au Parlement 2010. Enjeux et étapes*. Charenton-le-Pont : publications de l'IDS, p.48

aux données du SNIIRAM (gouvernance opérationnelle), mais également en se posant en figure de proue de l'ouverture des données de santé (gouvernance stratégique). Le 'flou' institutionnel qui subsiste autour de l'IDS amène ainsi cet organisme à revendiquer dès 2009 des modifications réglementaires ou législatives afin de lui permettre d'occuper une place plus importante dans la gouvernance stratégique. L'IDS avance alors l'argument selon lequel il serait déjà, de fait, une instance de concertation centrale dans l'alimentation des débats autour de l'ouverture du SNIIRAM, et que les textes devraient refléter ce positionnement.⁵⁵ Cette revendication fera même l'objet d'une proposition d'amendement déposée par le Collectif Inter-associatif en santé (CISS) et MG France lors du vote de la loi Hôpital Patients Santé Territoires et du projet de financement de la Sécurité Sociale de 2009.⁵⁶

Si, au niveau de la gouvernance opérationnelle, le positionnement de l'IDS ne semble pas poser de problèmes avec la CNIL (qui demeure l'organisme responsable de la délivrance et perçoit sans doute l'avis de l'IDS comme un allègement de sa charge de travail), l'aspect stratégique pose quant à lui plus de difficultés, notamment par rapport à la CNAMTS et au COPIIR, au sein duquel l'IDS n'a qu'un rôle consultatif (sur les relations entre la CNAMTS et l'IDS, voir II.b).

Les enjeux pratiques : le développement de nombreux outils pour faciliter l'exploitation des données du SNIIRAM

Enfin, les années 2000 ont également été marquées par la mise en œuvre de tout un ensemble d'instruments techniques visant à favoriser l'accès aux données du SNIIRAM, qui est une base riche mais également extrêmement complexe et difficile à manipuler pour des non-spécialistes. C'est dans un premier temps la fonction attribuée aux *datamarts*, entrepôts de données thématiques et agrégées qui permettent une navigation plus facile dans les données issues du SNIIRAM (Entretien, ex-agent CNAMTS, 22/09/2022, Entretien, agent CNAMTS, 21/04/2022).⁵⁷ Entre février 2008 et mars 2010, environ 7100 requêtes sont réalisées dans les *datamarts* par une trentaine d'utilisateurs des membres de l'IDS.⁵⁸

De même, dès 2005, la CNAMTS crée l'Echantillon Généraliste de Bénéficiaires (EGB), dont le but est également de faciliter la réutilisation des données du SNIIRAM (Entretien, ex-agent CNAMTS, 19/04/2022). L'EGB est ainsi constitué d'un échantillon au centième de la population avec un historique de 20 ans. Ces données, anonymes, individuelles, et représentatives de la population

⁵⁵ Institut des données de santé (2009). *Rapport au Parlement 2009. Les progrès*. Charenton-le-Pont : publications de l'IDS, pp.38-39

⁵⁶ Comité d'Experts de l'IDS (2010). *Livre blanc 2010 du Comité d'Experts. 7 propositions au service de la recherche*. Charenton-le-Pont : publications de l'IDS, pp.42-43

⁵⁷ Pour éviter les risques de réidentifications, certains garde-fous étaient associés à l'usage des *datamarts* : les requêtes dont le résultat donnait lieu à un groupe de moins de 10 individus n'étaient pas affichées, et il était interdit de croiser plus de deux des trois variables jugées 'sensibles' pour la réidentification : date de naissance, adresse, lieu/dates des soins (en particulier d'hospitalisation).

⁵⁸ Institut des données de santé (2010). *Rapport au Parlement 2010. Enjeux et étapes*. Charenton-le-Pont : publications de l'IDS, p.13

française, constituent une source d'information précieuse pour mener des recherches scientifiques. Dans le même temps, le fait que l'EGB ne représente qu'un échantillon permet, selon certains acteurs, de diminuer le risque de réidentification de chaque patient (Entretien, fonctionnaire, 17/06/2022).

En parallèle, l'IDS s'efforce de faciliter l'accès aux données de santé pour ses membres en mettant à leur disposition des tableaux de bord interactifs constitués de données agrégées, afin d'éviter aux utilisateurs d'avoir à effectuer des requêtes dans les *datamarts*, qui requièrent une certaine expertise technique et une formation par la CNAMTS (Entretien, représentant professionnels de santé, 23/05/2022 ; Entretien, représentant, association d'usagers, 30/06/2022). Ces tableaux de bord sont produits directement par la CNAMTS et l'IDS :

'[...] l'IDS a eu ce côté qu'a eu l'INDS après, d'essayer quelque part de **servir les besoins des parties prenantes**, parties prenantes au sens des syndicats des professionnels, des fédérations hospitalières, des associations de patients etc. Et du coup ils avaient fabriqué toute une, avec leur aide, toute une série de **tableaux de bord, qui étaient aussi supposés mieux que les datamarts, répondre à leurs besoins**, et qu'on demandait aussi à la CNAMTS d'alimenter puis à d'autres aussi parce que il n'y avait pas tout à fait que les données de la CNAMTS, le PMSI aussi, donc chacun mettait au pot le fait de rentrer régulièrement ces tableaux de bord en faisant des extractions de ses données.' (Entretien, ex-agent CNAMTS, 22/09/2022)

'[...] on avait pris un outil qui s'appelait Clickview je sais pas si vous connaissez, très intéressant parce que pour l'utilisateur, **il suffit de cliquer sur les icônes** pour avoir les données, sauf qu'il y a un gros travail de classement de données, en amont. Et on avait ouvert ça et on avait 30 tableaux de bord qui permettaient... le reste-à-charge, la démographie, voilà on avait fait ça nous.' (Entretien, agent CNAMTS, 21/04/2022)

Enfin, concernant l'aspect technique de l'accès aux données de santé, c'est l'ATIH qui est chargée de partager les données annuelles du PMSI et de les transmettre via un disque dur, tandis que la CNAMTS mobilise ses équipes pour permettre les extractions lorsqu'une demande d'accès au SNIIRAM leur parvient. On note ici la différence entre les deux bases de données : alors que l'ATIH est une agence spécifiquement dédiée à la mise en place de l'accès aux données de santé, la CNAMTS assure cette mission du fait de son rôle de dépositaire de la base de données. Ce n'est, de toute évidence, qu'une mission minoritaire parmi les nombreux attributs de la CNAMTS.

En conclusion, les années 2000 sont finalement marquées par deux éléments qui peuvent paraître contradictoires à première vue. En effet, il n'existe pas à l'époque de réelle politique volontariste d'ouverture de la donnée de santé, qui serait portée par des acteurs élus. Un de nos enquêtés témoigne ainsi d'un fonctionnement qui se serait pendant longtemps réalisé 'sans injonction' ou 'sans implication' des gouvernements successifs (Entretien, ex-fonctionnaire, 06/04/2022). De fait, les notions de 'Big Data', ou la qualification de la donnée en 'or noir' (Skovgaard *et al.* 2019), ne se développeront que plus tard au cours des années 2010. Or, lorsqu'on s'intéresse aux instruments de la régulation de l'accès aux données de santé, on observe, parallèlement, un réel mouvement d'ouverture 'progressif mais continu' (Entretien, représentant, secteur des assurances complémentaires, 16/06/2022).

Celui-ci est alors permis par une dynamique institutionnelle particulière, dont le point de départ se trouve dans les négociations avec les organismes d'assurance complémentaire entreprises autour de la loi du 13 août 2004 et qui créent l'Institut des Données de Santé. Durant les années 2000, l'ouverture des données de santé s'opère donc 'en sourdine', sans être accompagnée par une réelle politique d'ouverture des données de santé portée stratégiquement au niveau politique.⁵⁹

Tableau. La régulation de l'accès aux données de santé dans les années 2000

Règles d'accès	Gouvernance opérationnelle	Gouvernance stratégique	Enjeux pratiques
Evolution par rapport à la période précédente			
Fermeture/Ouverture	Fermeture	Ouverture	Ouverture
<p>Fermeture : Interdiction de l'accès au SNIIRAM pour des organismes à but lucratif ; Ajout d'une nouvelle étape au processus d'accès au SNIIRAM (avis de l'IDS) ;</p> <p>Ouverture : extension du nombre d'acteurs avec des accès 'de droit' à certains jeux de données issus du SNIIRAM (membres de l'IDS)</p>	<p>Ajout d'un nouvel acteur : l'IDS donne désormais un avis avant autorisation de la CNIL</p>	<p>Concurrence institutionnelle entre le COPIIR et l'IDS.</p> <p>Le rôle proactif de l'IDS, qui bénéficie de soutiens occasionnels mais cruciaux au sein du Ministère de la santé, permet d'adopter très progressivement des arrêtés ministériels actant l'ouverture des données du SNIIRAM</p>	<p>Mise en place d'instruments visant à faciliter l'accès aux données du SNIIRAM par l'IDS et la CNAMTS : datamarts, EGB, tableaux de bord</p>

Tableau. Les accès aux bases de données de santé selon les types d'acteurs (années 2000)

	PMSI (CNIL)	SNIIRAM (IDS +CNIL)	SNIIRAM accès de droit (jeux de données)	SNIIRAM accès de droit (total)
Acteurs publics (Ministères et services)	Oui	Oui	Oui (certains acteurs spécifiques)	Non

⁵⁹ Malgré tous les efforts de l'IDS, les accès au concret demeurent difficiles, la procédure étant complexe et le temps de réponse des différents acteurs (CNIL, CNAMTS...) étant particulièrement lent, ce qui s'explique en partie pour des raisons budgétaires et de moyens alloués à ce sujet.

Agences publiques	Oui	Oui	Oui (certains acteurs spécifiques)	Non
Professionnels de santé	Oui	Oui	Oui (certains acteurs spécifiques)	Non
Organisations de patients	Oui	Oui	Oui (certains acteurs spécifiques)	Non
Assurances complémentaires	Oui	Oui (organismes à but non lucratif)	Oui (certains acteurs spécifiques)	Non
Laboratoires pharmaceutiques	Oui	Non	Non	Non

Conclusion

Ce premier Working paper a permis de décrire les évolutions de la régulation de l'accès aux données de santé en France autour de deux périodes distinctes : les années 1990 et les années 2000. Bien qu'un tel découpage temporel demeure quelque peu sommaire, il permet de rendre compte de l'évolution de l'accès aux données de santé en prenant en compte le temps de formation et de mise en œuvre de ces bases et en analysant les premiers usages qui en sont faits. Ces deux périodes sont marquées par des dynamiques relativement similaires, au sens où l'enjeu des données de santé est systématiquement incorporé au sein d'autres enjeux plus globaux du système de santé. Autrement dit, les évolutions de ce secteur ne sont pas liées à une politique volontariste de l'État qui érigerait l'accès aux données de santé comme un objectif en soi. Au contraire, l'ouverture des données de santé évolue plutôt au gré des réformes plus globales du système de santé, et est souvent considérée comme une mesure 'parmi d'autres' dans le cadre de compromis plus généraux. C'est par exemple l'enjeu de la crise financière du système de sécurité sociale de 1995 qui va permettre au PMSI comme au SNIIRAM de devenir des projets stables sur un plan institutionnel. À ce titre, les Ordonnances Juppé adoptées en 1996 agissent comme de réelles fenêtres d'opportunité sur le développement des bases de données de santé centralisées en France. Ce sont ensuite les réformes de l'assurance maladie et l'institutionnalisation du partenariat entre l'État, l'assurance maladie obligatoire et les assurances maladie complémentaires qui vont venir transformer le secteur de la régulation de l'accès aux données de santé en créant l'Institut des Données de Santé. En opposition avec la CNAMTS, cet organisme proactif en matière d'ouverture des données de santé se mobilise, tout au long des années 2000, pour assurer une plus grande ouverture des données de santé à une palette diversifiée d'acteurs (professionnels de santé, représentants de patients, assurances complémentaires, fédérations hospitalières...). Au cours de cette décennie, ce sont donc essentiellement des luttes de pouvoir institutionnelles qui vont conditionner l'ouverture de l'accès aux données de santé, dont nous pouvons par ailleurs décrire les évolutions en trois temps.

Premièrement, on assiste, à la fin des années 1990, à la mise en place de la protection 'de base' des données de santé via la modification de la Loi Informatique et Libertés, qui oblige tout utilisateur du PMSI et du SNIIRAM à obtenir une autorisation de la CNIL avant d'accéder aux données de santé. Cette première régulation fait directement suite aux premiers usages 'non-anticipés' du PMSI par des journalistes, et prend place dans un contexte de méfiance envers les initiatives de l'État en matière de recueil et de centralisation des données personnelles. Dans un second temps, ces premières régulations sont renforcées par la mise en œuvre effective du SNIIRAM qui intervient au début des années 2000. En effet, la CNAMTS, responsable des données du SNIIRAM, ajoute des conditions supplémentaires à l'accès aux données de santé, puisque l'arrêté SNIIRAM interdit, dès 2005, l'accès aux organismes à but lucratif. Dans un troisième temps, cette dynamique est contrariée par la création de l'Institut des Données de Santé, qui vise à permettre l'accès aux données de santé aux assureurs complémentaires dans une perspective de 'co-régulation' du risque maladie. Malgré les réticences de la CNAMTS, l'IDS parvient, durant ces quelques années, à ouvrir l'accès aux données de santé à plusieurs acteurs, dont les fédérations d'assurance complémentaire.

Dès le début des années 2010, la question des données de santé va cependant devenir plus autonome, pour finir par constituer un enjeu à part entière qui sera traité dans la loi de modernisation de notre système de santé de 2016. Avec ce processus d'autonomisation, les enjeux relatifs aux données de santé seront redéfinis et les relations de pouvoir entre les différents acteurs profondément transformées. Ces évolutions font l'objet du second Working paper sur la régulation de l'accès aux données de santé.

Bibliographie

- Antoniotti, S., C. Manuel, C. Sapin, Pascal Auquier, and J.-L. San Marco (2002). 'Déclaration obligatoire de l'infection par le VIH. Justifications épidémiologiques et interrogations éthiques', *Santé Publique*, 14:1, 63–73.
- Benamouzig, Daniel (2012). 'Du grand soir au clair-obscur. Expertise économique et privatisation bureaucratique de l'assurance maladie', *Actes de la recherche en sciences sociales*, 193:3, 56–73.
- Benoît, Cyril, and Gaël Coron (2018). 'Les « mises en marché » du risque santé en France : une pluralité de logiques et de séquences de changement', *Revue française des affaires sociales*, 4, 259–279.
- Bonaïti-Pellié, Catherine et al. (2009). 'Traitement de l'information en matière de recherche dans le domaine de la santé - Nul n'est censé ignorer la loi !', *médecine/sciences*, 25:1, 93–98.
- Hassenteufel, Patrick, and Bruno Palier (2005). 'Les trompe-l'œil de la « gouvernance » de l'assurance maladie', *Revue française d'administration publique*, 113:1, 13–27.
- Kingdon, John W. (2002). *Agendas, Alternatives, and Public Policies*. 2nd ed. New York: Pearson Education Editions.
- Malye, François, and Jérôme Vincent (2015). 'Le palmarès des hôpitaux', *Les Tribunes de la santé*, 47:2, 65–74.
- Moisdon, Jean-Claude (2013). 'Une histoire de la T2A', *Journal de gestion et d'économie médicales*, 31:2–3, 107–120.
- Palier, Bruno (199). *Réformer la sécurité sociale. Les interventions gouvernementales en matière de protection sociale depuis 1945, la France en perspective comparative*. Thèse de doctorat Paris: Institut d'Etudes Politiques de Paris.
- Pierru, Frédéric (2004). 'La Fabrique des palmarès. Genèse d'un secteur d'action publique et renouvellement d'un genre journalistique : les palmarès hospitaliers', in Jean-Baptiste Legavre (ed.), *La presse écrite : objets délaissés, Logiques politiques*. Paris: L'Harmattan, 247–270.
- Skovgaard, Lea L., Sarah Wadmann, and Klaus Hoeyer (2019). 'A review of attitudes towards the reuse of health data among people in the European Union: The primacy of purpose and the common good', *Health Policy*, 123:6, 564–571.

Annexe : Liste anonymisée des entretiens réalisés

	Fonction anonymisée	Type entretien	Date entretien	Durée entretien
1	Entretien, ex-fonctionnaire, PMSI, DSSIS	Entretien en visioconférence	06/04/2022	01:02:51
2	Entretien, fonctionnaire, ex-IDS	Appel téléphonique	14/04/2022	00:27:32
3	Entretien, ex-agent CNAMTS	Entretien présentiel	19/04/2022	02:02:26
4	Entretien, agent CNAMTS	Entretien présentiel	21/04/2022	01:39:41
5	Entretien, représentant professionnels de santé	Entretien en visioconférence	23/05/2022	0:41:30
6	Entretien, ex-fonctionnaire, PMSI, Sesame-vitale	Utilisation secondaire entretien thèse	24/05/2018 30/05/2018	05:09:50
7	Entretien, agent CNIL	Utilisation secondaire entretien thèse	15/06/2018	01:15:26
8	Entretien, ex-conseillère de Marisol Touraine	Utilisation secondaire entretien thèse	24/11/2016	00:51:50
9	Entretien, secteur des assurances complémentaires	Entretien en visioconférence	15/06/2022	01:06:44
10	Entretien, juriste spécialiste des données de santé	Entretien présentiel	15/06/2022	01:35:40
11	Entretien, ex-représentant des professionnels de santé	Entretien présentiel	16/06/2022	00:40:17
12	Entretien, représentant, secteur des assurances complémentaires	Entretien en présentiel	16/06/2022	01:07:15
13	Entretien, fonctionnaire, DREES	Entretien en présentiel	17/06/2022	00:57:44
14	Entretien, représentant, association d'usagers	Entretien en visioconférence	30/06/2022	00:52:00
15	Entretien, ex-conseiller de Marisol Touraine	Entretien en visioconférence	19/07/2022	00:24:35
16	Entretien, fonctionnaire et agent CNAMTS	Entretien hybride	20/07/2022	01:25:56
17	Entretien, représentant, industries pharmaceutiques	Entretien en présentiel	21/07/2022	00:39:17
18	Entretien, industriel du numérique en santé	Entretien en visioconférence	22/07/2022	00:46:38
19	Entretien, industrie pharmaceutique	Entretien en visioconférence	24/08/2022	00:32:39
20	Entretien, fonctionnaire, Health Data Hub	Entretien en visioconférence	24/08/2022	00:56:24
21	Entretien, ex-agent CNAMTS	Entretien en présentiel	22/09/2022	01:18:57

PARI

PROGRAMME DE RECHERCHE
SUR L'APPRÉHENSION DES RISQUES
ET DES INCERTITUDES

PARI, placé sous l'égide de la Fondation Institut Europlace de Finance en partenariat avec l'ENSAE/Excess et Sciences Po, a une double mission de recherche et de diffusion de connaissances.

Elle s'intéresse aux évolutions du secteur de l'assurance qui fait face à une série de ruptures : financière, réglementaire, technologique. Dans ce nouvel environnement, nos anciens outils d'appréhension des risques seront bientôt obsolètes. PARI a ainsi pour objectifs d'identifier leur champ de pertinence et de comprendre leur émergence et leur utilisation.

L'impact de ses travaux se concentre sur trois champs :

- les politiques de régulation prudentielle dans un contexte où Solvabilité 2 bouleverse les mesures de solvabilité et de rentabilité (fin du premier cycle de la chaire);
- les solutions d'assurance, à l'heure où le big data déplace l'assureur vers un rôle préventif, créant des attentes de personnalisation des tarifs et de conseil individualisé ;
- les technologies de data science appliquées à l'assurance, modifiant la conception, l'appréhension et la gestion des risques.

Dans ce cadre, la chaire PARI bénéficie de ressources apportées par Addactis, la CCR, Generali, Groupama, la MGEN et Thémis.

Elle est co-portée par **Pierre François**, chercheur au CNRS, doyen de l'Ecole Doctorale de Sciences Po et **Laurence Barry**, chercheur à Datastorm, la filiale de valorisation de la recherche de l'ENSAE.

PARTENAIRES

