Université de POITIERS

Faculté de Médecine et de Pharmacie

2019 **Thèse n°**

THESE
POUR LE DIPLOME D'ETAT
DE DOCTEUR EN PHARMACIE

(arrêté du 17 juillet 1987)

présentée et soutenue publiquement
le 14 juin 2019 à POITIERS
par Monsieur BOIDIN Romain
né le 22 mars 1989

Protection et gouvernance des données dans la recherche en santé

Composition du jury :

Président :

Monsieur le Professeur OLIVIER Jean-Christophe, Docteur en pharmacie, Docteur ès Sciences Pharmaceutiques

Membres :

Monsieur le Docteur BERTOYE Pierre-Henri, Docteur en médecine, Président de la Commission Nationale des Recherches Impliquant la Personne Humaine

Monsieur le Docteur LACOSTE Louis, Docteur en médecine, Président de la Conférence Nationale des Comités de Protection des Personnes

Directrice de thèse :

Madame la Professeure RAGOT Stéphanie, PU-PH en Santé publique, Docteure en pharmacie

UNIVERSITÉ DE POITIERS

Faculté de Médecine et de Pharmacie

Année universitaire 2018-2019

PHARMACIE

Professeurs

- CARATO Pascal, Chimie Thérapeutique
- COUET William, Pharmacie Clinique
- DUPUIS Antoine, Pharmacie Clinique
- FAUCONNEAU Bernard, Toxicologie
- GUILLARD Jérôme, Pharmaco chimie
- IMBERT Christine, Parasitologie
- MARCHAND Sandrine, Pharmacocinétique
- OLIVIER Jean Christophe, Galénique
- PAGE Guylène, Biologie Cellulaire
- RABOUAN Sylvie, Chimie Physique, Chimie Analytique
- RAGOT Stéphanie, Santé Publique
- SARROUILHE Denis, Physiologie
- SEGUIN François, Biophysique, Biomathématiques

Maîtres de Conférences

- BARRA Anne, Immunologie-Hématologie
- BARRIER Laurence, Biochimie
- BODET Charles, Bactériologie (HDR)
- BON Delphine, Biophysique
- BRILLAULT Julien, Pharmacologie
- BUYCK Julien, Microbiologie
- CHARVET Caroline, Physiologie
- DEBORDE Marie, Sciences Physico-Chimiques
- DELAGE Jacques, Biomathématiques, Biophysique
- FAVOT Laure, Biologie Cellulaire et Moléculaire
- GIRARDOT Marion, pharmacognosie, botanique, biodiversité végétale
- GREGOIRE Nicolas, Pharmacologie (HDR)
- HUSSAIN Didja, Pharmacie Galénique (HDR)
- INGRAND Sabrina, Toxicologie
- MARIVINGT-MOUNIR Cécile Pharmaco chimie
- PAIN Stéphanie, Toxicologie (HDR)
- RIOUX BILAN Agnès, Biochimie
- TEWES Frédéric, Chimie et Pharmaco chimie
- THEVENOT Sarah, Hygiène et Santé publique
- THOREAU Vincent, Biologie Cellulaire
- WAHL Anne, Pharmaco chimie, Produits naturels

AHU

- BINSON Guillaume

PAST - Maître de Conférences Associé

- DELOFFRE Clément, Pharmacien
- HOUNKANLIN Lydwin, Pharmacien

Professeur 2nd degré

- DEBAIL Didier
- GAY Julie

Poste de Doctorant

- FREYSSIN Aline

Remerciements

Je tiens à remercier toutes les personnes qui ont soutenu ou nourri ce travail de thèse et notamment

Les membres du jury, Stéphanie Ragot, Jean-Christophe Olivier, Louis Lacoste et Pierre-Henri Bertoye pour l'intérêt qu'ils portent au sujet et leur disponibilité ;

Stéphanie Ragot pour avoir dirigé la rédaction de cette thèse ;

Pierre-Henri Bertoye pour sa confiance, son soutien et le temps qu'il m'a accordé pour travailler sur ces enjeux ;

Hafsa Boutabaa (Unicancer) et Pierre Malvoisin (Unicancer) pour les relectures communes des textes et les réflexions sur leur interprétation ;

Les délégués à la protection des données des Centres de Lutte Contre le Cancer (CLCC), la Coordination des Promoteurs Institutionnels (CPI), les personnels de l'Institut National de la Santé et de la Recherche Médicale (INSERM), de l'European Organisation for Research and Treatment of Cancer (EORTC), de la Ligue contre le cancer, de France Lymphome Espoir, de la Commission Nationale de l'Informatique et des Libertés (CNIL), de l'Institut National des Données de Santé (INDS), de la Direction Générale de la Santé (DGS) et de la Direction de la Recherche, des Études, de l'Évaluation et des Statistiques (DREES) pour les travaux communs sur les enjeux et applications pratiques de la protection des données personnelles ;

Et toutes les personnes avec qui alimentent régulièrement ces réflexions dans le cadre de nos échanges professionnels ou en dehors.

Sommaire

Abréviations

ATIH	Agence technique de l'information sur l'hospitalisation
BCR	*Binding Corporate Rules*, règles contraignantes d'entreprise
CCT	Clauses contractuelles types, publiées par la Commission européenne
CEPD	Comité européen de la protection des données
CépiDc	Centre d'épidémiologie sur les causes médicales de Décès
CEREES	Comité d'expertise pour les recherches, les études et les évaluations dans le domaine de la santé
CIL	Correspondant à la protection des données à caractère personnel, également appelé Correspondant informatique et libertés
CNIL	Commission nationale de l'informatique et des libertés
CSP	Code de la santé publique
DCP	Donnée à caractère personnel
DPIA	*Data Privacy Impact Assessment*, Analyse d'impact sur la protection des données
EEE	Espace économique européen, rassemblant les 28 Etats membres de l'Union européenne et trois des quatre États membres de l'Association européenne de libre-échange (AELE) que sont l'Islande, la Norvège et le Liechtenstein (la Suisse n'ayant pas ratifié l'accord).
EGB	Echantillon généraliste de bénéficiaires
G29	Groupe de l'Article 29
INDS	Institut national des données de santé
LIL	Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés modifiée, dite loi "informatique et libertés"
MR	Méthodologie de référence, publiée par la CNIL afin de simplifier les démarches réglementaires
NIR	Numéro d'inscription au répertoire de l'INSEE (numéro dit de sécurité sociale)
PMSI	Programme de médicalisation des systèmes d'information
RGPD	Règlement UE 2016/679 ou Règlement général sur la protection des données
SNDS	Système national des données de santé
SNIIRAM	Système national d'information inter-régimes de l'assurance maladie

I. Introduction

Avec le développement des bases de données du soin et de la recherche ainsi que la multiplication des flux de données par le biais des objets connectés, des outils informatiques et de la télémédecine, la problématique de la protection des données personnelles en santé est plus que jamais d'actualité. Il devient impératif pour tout professionnel, établissement ou organisme du secteur de la santé de maitriser la protection des données qu'ils détiennent et utilisent en établissant une véritable gouvernance des données.

Au-delà des obligations réglementaires, la protection des données face aux risques de piratage, d'utilisation à mauvais escient et de divulgation est essentielle à la relation de confiance entre les professionnels de santé et les patients.

Par ailleurs, la recherche dans le domaine de la santé est stimulée par la disponibilité de données. Celle-ci est rendue possible par les collaborations entre chercheurs et le partage de données, que ce soit sur support numérique ou sur le support organique que constituent les échantillons biologiques. La bonne préparation du cadre de collaboration et de partage des données est donc essentielle.

L'erreur de cantonner le pharmacien au domaine des médicaments voire à l'officine est souvent faite. Pourtant, il est avant tout un scientifique de haut niveau, généraliste, adaptable et pluri-compétent.

Dans ses diverses missions à l'officine, à l'hôpital, dans l'industrie des produits de santé ou dans les pouvoirs publics, le pharmacien joue un rôle clé de maitre d'ouvrage, d'interface ou de support, expert et pédagogue.

A l'heure de la transition digitale, le pharmacien a un rôle à jouer dans la mise en place des conditions qui permettront la pleine exploitation des nouvelles technologies de communication, de collecte et d'exploitation des données, de manière pertinente pour la prise en charge des patients et pour la recherche, tout en proposant un niveau élevé de protection des données.

Du pharmacien garant des connaissances, de la conservation des produits de santé et de leur flux de distribution, au pharmacien acteur de l'organisation de la gouvernance des flux et bases de données et de leur mise à disposition des équipes de soin et de recherche, il n'y a qu'un pas, que je vous propose de franchir dans cette thèse.

Nous aborderons dans un premier temps les notions essentielles et le contexte réglementaire encadrant la protection des données appliquée à la recherche en santé afin d'illustrer des principes parfois difficiles à cerner pour le non-spécialiste.

Dans un second temps, nous traiterons des clarifications qui restent à apporter et de l'opportunité de mettre en place une véritable gouvernance des données.

II. Contexte réglementaire

A. Textes encadrant la protection des données

Le cadre réglementaire relatif à la protection des données personnelles de santé des patients en France est composé principalement d'un Règlement européen, de la loi n°78-17 du 6 janvier 1978 modifiée, dite loi "Informatique et libertés", ainsi que de textes du Code de la santé publique.

1. Règlement (UE) n° 2016/679 (RGPD)

Le Règlement (UE) 2016/679 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, dit Règlement général sur la protection des données (RGPD) est applicable dans l'Espace économique européen (EEE) depuis le 25 mai 2018.

Abrogeant et remplaçant la Directive 95/46/CE, il a pour objectif d'harmoniser les règles et droits fondamentaux relatifs à la protection des données à caractère personnel dans l'Espace Économique Européen.

En tant que Règlement, il est directement applicable dans les États Membres depuis le 25 mai 2018, bien qu'il possède la particularité de laisser à ces derniers une certaine latitude pour légiférer dans le droit national notamment sur les sujets des données réputées sensibles et dans le domaine de la recherche scientifique.

Dans sa genèse, il vise notamment à réguler l'utilisation des données personnelles par les géants du web tels que Google, Apple, Facebook, Amazon - les fameux GAFA -, etc. Par conséquent, et du fait de la facilité de circulation des données dans le monde, ce Règlement est applicable non seulement aux organismes traitant des données personnelles dans l'EEE, mais également à tout organisme traitant des données concernant des résidents européens, quel que soit la localisation dudit organisme.

2. Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés modifiée

La loi dite "Informatique et libertés" (LIL)(1) a été rédigée dans les années 1970 suite à la prise de conscience des possibilités d'entrave à la vie privée par la puissance de l'outil informatique.

Si d'autres réflexions étaient en cours sur ce thème depuis la fin des années 1960 en Europe, aux États-Unis et au Canada entrainant la publication de textes législatifs, le fait fondateur de cette loi est la révélation au public par le journal le Monde, dans un

article du 21 mars 1974 titré "SAFARI ou la chasse aux français", d'un système d'envergure visant à recouper, par le numéro INSEE de Sécurité Sociale (NIR), divers répertoires d'identification jusque-là régionaux et indépendants.(2)

Le chapitre de cette loi portant sur les données de santé, a été largement remanié par la loi n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles, cette dernière visant à mettre la LIL en adéquation avec le RGPD.

L'ordonnance n° 2018-1125 du 12 décembre 2018, prise en application de l'article 32 de la loi n° 2018-493, vise à clarifier la LIL et à la mettre en cohérence avec le RGPD. Elle entrera en vigueur à publication de son décret d'application et au plus tard le 1er juin 2019.(3)

Du fait de sa prochaine entrée en vigueur, c'est, sauf mention contraire, la numérotation des articles de la LIL résultant de cette ordonnance qui est utilisée dans cette thèse.

Notons que le législateur français s'est largement saisi, pour ce qui est des données de santé, du pouvoir de subsidiarité dont sont investis les États membres par l'article 9.4 du RGPD. Par conséquent, le Chapitre III Section 3 de la LIL introduit des dispositions spécifiques au traitement des données de santé en France. Alors que le RGPD prévoit la responsabilisation des acteurs qui doivent assurer la documentation de leur conformité en interne, la LIL maintien en France pour la recherche un principe d'autorisation préalable qui prévalait avant le Réglement.

3. La Commission nationale de l'informatique et des libertés

Créée par la Loi informatique et libertés, la CNIL est une autorité administrative indépendante ayant pour mission l'information des personnes et la vérification de la conformité des traitements de données à caractère personnel.

C'est donc auprès de cette autorité que sont effectuées les démarches de déclaration et de demande d'autorisation de ces traitements.

Afin de simplifier les démarches réglementaires, la CNIL développe des simplifications sectorielles et notamment des Méthodologies de référence qui permettent la mise en place plus rapide des recherches en santé (*cf* §III).

4. Comité européen de la protection des données

Afin de permettre une application cohérente de la réglementation, le RGPD institue un Comité européen de la protection des données (Chapitre VII, Section 3). Celui-ci

se compose du chef de l'autorité de contrôle de chaque État membre et du Contrôleur européen de la protection des données, ou de leurs représentants respectifs.

Le comité peut adopter des documents d'orientations générales afin de clarifier les dispositions de la réglementation et en fournir une interprétation cohérente.

Il remplace l'ancien Groupe de travail créé par l'article 29 de la Directive 95/46/CE, dit G29. A la différence de ce dernier, le nouveau comité a la capacité d'adopter des avis et des décisions contraignantes afin de trancher les différends entre autorités de contrôle.

B. Définitions essentielles

1. Donnée à caractère personnel (DCP)

Il convient de clarifier une confusion fréquente sur la notion de donnée à caractère personnel (souvent qualifiée de "donnée personnelle") et sur la notion d'anonymat.

L'article 4(1) du RGPD définit la donnée à caractère personnel telle que *toute information se rapportant à une personne physique identifiée ou identifiable (ci-après dénommée «personne concernée». Est réputée être une «personne physique identifiable» une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale.*

La Directive 95/46/CE et l'article 2 de la LIL, dans sa rédaction préalable à l'ordonnance n° 2018-1125 du 12 décembre 2018 clarifiaient que " Pour déterminer si une personne est identifiable, il convient de considérer l'ensemble des moyens en vue de permettre son identification dont dispose ou auxquels peut avoir accès le responsable du traitement ou toute autre personne".

La donnée anonyme, elle, s'oppose par définition à la donnée à caractère personnel (cf. Fig. 1).

On ne peut considérer une donnée comme anonyme que dans deux situations

- lorsque tout lien permettant la ré-identification a été détruit et que les données sont suffisamment vagues pour ne pas concerner une personne ou un groupe réduit de personne ;
- ou dans le cas des données agrégées.

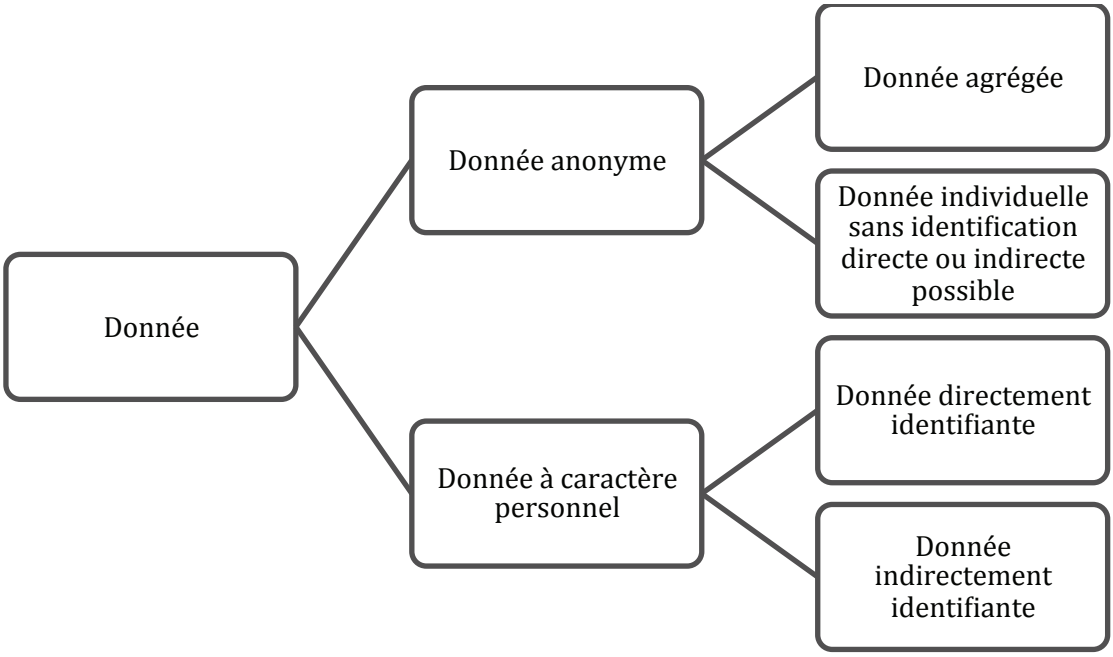

Figure 1 : Distinction des différents types de données selon le RGPD

Aux États-Unis, seules les données dites identifiantes entrent dans le champ de la loi HIPAA relative à la protection de ces données. La loi définit une liste de 18 identifiants, tels que le nom, adresse, dates d'hospitalisation, etc. Lorsque ces identifiants sont supprimés d'une base de données, celle-ci est alors considérée comme dé-identifiée et non sujette à la réglementation.

Face à la définition américaine reposant sur des critères précis, le Groupe de l'article 29, a considéré en 2014 qu'aucune technique de minimisation des données ne permet d'atteindre un niveau permettant un réel anonymat.(4)

Le considérant 26 du RGPD clarifie que la possibilité d'identification des personnes s'interprète au regard des moyens raisonnablement susceptibles d'être utilisés en prenant en considération l'ensemble des facteurs objectifs tel que notamment le coût et le temps nécessaires à l'identification ainsi que les technologies disponibles au moment du traitement mais également de leur évolution. La notion d'anonymat doit donc être raisonnablement robuste dans le temps.

Afin de concilier impératifs de transparence et exigence réglementaire de protection des données personnelles, l'Agence européenne du médicament considère acceptable pour les données publiées un seuil de risque de ré-identification de 0,09. C'est-à-dire, que les données caractéristiques d'une personne dans le set considéré puissent concerner au moins 11 autres participants de l'essai. Ce seuil est actuellement utilisé pour la publication des bases de données d'essais portant sur le médicament sur la base de l'Agence européenne.

Cette notion représente un enjeu des discussions autour de la réglementation car travailler sur une donnée anonyme permettrait aux chercheurs de s'affranchir des dispositions du RGPD, telles que l'information des personnes, le maintien d'un niveau de sécurité adéquat, et la formalisation de tout traitement de données. Toutefois, les méthodes d'anonymisation sont complexes et leur robustesse dans le temps difficile à apprécier. Il convient donc de considérer tout donnée de recherche en santé comme une donnée personnelle, jusqu'à preuve du contraire.

1. Traitement de données à caractère personnel

Le RGPD définit le traitement comme *toute opération ou tout ensemble d'opérations effectuées ou non à l'aide de procédés automatisés et appliquées à des données ou des ensembles de données à caractère personnel, telles que la collecte, l'enregistrement, l'organisation, la structuration, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, la diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, la limitation, l'effacement ou la destruction* (art.2.2 RGPD).

Ainsi, il convient de considérer toute activité relative aux données à caractère personnel, y compris la simple conservation comme soumise à la réglementation afférente.

2. Responsable de traitement

Le responsable de traitement, tel que défini à l'Article 2(7) du RGPD, est la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui, seul ou conjointement avec d'autres, détermine les finalités et les moyens du traitement. Lorsque les finalités et les moyens de ce traitement sont déterminés par le droit de l'Union ou le droit d'un État membre, le responsable du traitement peut être désigné ou les critères spécifiques applicables à sa désignation peuvent être prévus par le droit de l'Union ou par le droit d'un État membre.

L'article 26 du Règlement développe par ailleurs la notion de responsabilité conjointe. En effet, lorsque plusieurs entités déterminent conjointement les finalités et moyens du traitement, cette définition trouve à s'appliquer. Ce peut être le cas des consortiums de recherche par exemple, qui n'ont pas eux même de personnalité juridique, lorsque plusieurs parties indépendantes juridiquement siègent au sein d'une même instance décisionnaire pour la conduite de la recherche.

En recherche clinique, il est important de noter que la notion de responsable de traitement pourrait ne pas se superposer exactement à celle de promoteur. Prenons par exemple la définition de promoteur proposée par l'Article 2-2-14 du Règlement (UE) n° 536/2014 portant sur les essais cliniques de médicament :

« promoteur », une personne, une entreprise, un institut ou une organisation responsable du lancement, de la gestion et de l'organisation du financement de l'essai clinique.

Alors que la définition de responsable de traitement est factuelle (basée sur les faits), celle de promoteur pourrait être formelle (tel que le définirait un contrat). La définition de promoteur semble par ailleurs d'un niveau d'organisation plus bas, liée à la logistique de l'essai (sans jugement de valeur, la définition de promoteur emportant par ailleurs de lourdes responsabilités envers les participants aux essais). Ainsi il pourrait apparaitre des situations où le promoteur est un responsable conjoint du traitement ou un sous-traitant du responsable de traitement.

Cette distinction possible n'est pas actuellement prise en compte par les Méthodologies de référence portant sur les recherches impliquant la personne humaine (RIPH) que sont les MR-001 et MR-003.

Notons toutefois que le Règlement (UE) n° 536/2014 stipule dans son article 71 qu'un essai clinique peut avoir plusieurs promoteurs, ce qui pourra permettre dans certaines situations d'aligner les définitions.

3. Sous-traitant

Défini par l'article 2(8) RGPD, le sous-traitant est *la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui traite des données à caractère personnel pour le compte du responsable du traitement.*

En matière de protection des données à caractère personnel, il convient donc de considérer cette définition, comme pour le responsable de traitement, de manière factuelle. Ainsi il ne s'agit pas ici de jugement de valeur mais bien la définition d'un rôle au regard du traitement de données effectué (*data processor* en anglais). Un « partenaire de recherche » pourra donc être un sous-traitant au sens du RGPD dès lors qu'il ne définit pas la finalité ou le moyen de la recherche.

4. Délégué à la protection des données

L'article 37 du RGPD rend la désignation d'un délégué à la protection des données (DPD ou DPO pour *Data Protection Officer* en anglais) au sein des organismes suivants :

- Les autorités ou les organismes publics,
- Les organismes dont les activités de base les amènent à réaliser un suivi régulier et systématique des personnes à grande échelle,
- Les organismes dont les activités de base les amènent à traiter à grande échelle des données dites « sensibles » ou relatives à des condamnations pénales et infractions.

Le Groupe de l'article 29 a clarifié que la notion de grande échelle s'apprécie au regard du nombre de personnes concernées, du volume de données, de la durée du traitement et de son étendue géographique. Ainsi, les traitements réalisés par un médecin libéral seul n'entreraient pas dans cette définition alors que ce serait le cas pour un hôpital.

Dès lors, la plupart des recherches utilisant des données de santé, définies comme une catégorie de données sensibles, semble nécessiter la nomination d'un DPO au sein de l'organisme responsable du traitement.

L'article 39 du RGPD dresse une liste non exhaustive des missions du DPO :

- informer et conseiller le responsable du traitement ou le sous-traitant ainsi que les employés qui procèdent au traitement sur les obligations qui leur incombent en matière de protection des données;
- contrôler le respect du RGPD, et des autres dispositions du droit de l'Union en matière de protection des données et le respect des règles internes du responsable du traitement en matière de protection des données à caractère personnel, y compris en ce qui concerne la répartition des responsabilités, la sensibilisation et la formation du personnel participant aux opérations de traitement, et les audits s'y rapportant ;
- dispenser des conseils, sur demande, en ce qui concerne l'analyse d'impact relative à la protection des données et vérifier son exécution ;
- coopérer avec l'autorité de contrôle ;
- faire office de point de contact pour l'autorité de contrôle sur les questions relatives au traitement, y compris la consultation préalable visée à l'article 36 du RGPD, et mener des consultations, le cas échéant, sur tout autre sujet.

– Être au cœur de l’analyse d'impact relative à la protection des données.

C. Principes relatifs au traitement des données personnelles

Le RGPD institue avant tout des principes permettant de garantir une maitrise des traitements de données au sein des organismes ainsi qu'une transparence envers les personnes concernées.

Ces principes, définis à l'article 5, sont :

- la licéité, loyauté et transparence du traitement,
- la définition d'une finalité du traitement
- la minimisation des données,
- l'exactitude des données,
- la limitation de la durée de conservation des données,
- la sécurité technique et organisationnelle du traitement.

La documentation du respect de ces principes incombe au responsable de traitement. Par ailleurs, afin de garantir le respect de ces principes, les personnes concernées par le traitement de leurs données sont investies de droits qu'elles peuvent exercer auprès du responsable de traitement et de possibilités de recours auprès de l'autorité de contrôle (c'est à dire la CNIL, en France).

1. Finalité

Les données doivent tout d'abord être traitées pour une ou des finalités déterminées, explicites et légitimes.

Ces finalités devront à la fois être documentées en interne et faire l'objet d'une information des personnes concernées (cf. § III.B).

2. Licéité et fondements juridiques

Le traitement de données, pour être licite, doit être justifié et répondre à au moins un fondement juridique précis, tel que défini à l'article 6 RGPD.

Le traitement de catégories particulières de données, dites sensibles, dont font partie les données de santé, doit de plus répondre à une condition listée à l'article 9(2) afin de justifier de leur utilisation.

Fondement juridique de l'article 6 et exception de l'article 9 sont ainsi à considérer comme la justification de la légitimité du traitement de données personnelles, rendant ainsi illégaux les traitements qui ne répondraient pas à l'une de ces justifications considérées comme valables.

Les choix du fondement juridique et de la condition permettant le traitement de données sensibles peuvent s'avérer importants pour le responsable de traitement. En effet, de ceux-ci découlent les conditions d'exercice des droits des personnes et la robustesse juridique du traitement effectué. Par ailleurs, une fois déterminé, toute modification du fondement juridique nécessite une nouvelle information des personnes concernées (article 13 et 14 RGPD).

Dans le cadre de la recherche en santé, les fondements juridiques les plus pertinents semblent être le consentement éclairé de la personne (article 6(1)(a) du RGPD), l'exécution d'une mission d'intérêt public (article 6(1)(e)) et l'intérêt légitime du responsable de traitement (article 6(1)(f)). En addition, pour les fins de conservation des données pour la durée réglementaire, et dans la mesure où la conservation des données est définie comme une opération de traitement, le fondement du respect d'une obligation légale doit également être invoqué par le responsable de traitement (article 6(1)(c)).

a) Consentement

Tout d'abord, dans le contexte de la recherche en santé, rappelons que l'obtention du consentement éclairé de la personne préalablement à sa participation à une recherche impliquant la personne humaine (RIPH) de catégorie 1°, 2° ou à une RIPH nécessitant l'examen de ses caractéristiques génétiques est une obligation réglementaire découlant du Code de la santé publique (article L. 1122-1-1 CSP) et du Code civil (article 16-10) et visant avant tout à être un garde-fou à la réalisation de ces recherches sans l'accord de la personne concernée.

Cette obligation est à distinguer du fondement juridique sur lequel va reposer le traitement des données personnelles. En effet, pour une RIPH qui nécessiterait le recueil préalable du consentement du participant pour des raisons méthodologiques (RIPH de catégorie 1° ou 2° qui présente par définition des risques ou contraintes, ou recherche nécessitant l'analyse des caractéristiques génétiques), le responsable de traitement peut faire reposer le traitement des données personnelles sur un autre fondement que celui du consentement explicite.

Lorsque la personne retire son consentement, il convient de distinguer s'il s'agit d'un consentement requis pour participer à la recherche en raison de sa méthodologie (recherche impliquant la personne humaine comportant des risques ou contraintes pour la personne ou nécessitant l'examen des caractéristiques génétiques) ou s'il s'agit également du consentement en tant que fondement juridique sur lequel repose éventuellement le traitement des données.

Le choix d'un autre fondement pourrait s'expliquer par la fragilité du fondement juridique que représente le consentement. En effet, lorsqu'un traitement ne repose que sur celui-ci, le retrait de consentement retire au responsable de traitement toute légitimité pour traiter ces données et l'oblige à les effacer (article 17(1)(b) du RGPD). Ceci peut notamment avoir pour conséquence d'introduire un biais de sélection dans la recherche si l'on considère que les personnes exercent généralement leurs droits de retrait de consentement dans des situations de rejet vis à vis de la recherche, suite par exemple à des effets indésirables ou des contraintes trop importantes.

Par ailleurs, le G29, dans des lignes directrices sur le consentement endorsées par son successeur, le Comité européen de la protection des données, indique que " les projets de recherche scientifique ne peuvent inclure des données à caractère personnel sur la base du consentement de la personne concernée que si leur finalité est décrite avec précision".(5) Lorsque les finalités ne peuvent être précisées d'entrée de jeu, G29 évoque notamment la possibilité pour le responsable de traitement de mettre en place un système s'apparentant au *smart-contract*, visant à revenir vers la personne durant la progression du programme de recherche afin d'obtenir un consentement qui sera modifié de façon flexible et progressive. Si l'intention est louable, elle parait difficilement compatible avec la difficulté opérationnelle parfois rencontrée pour revenir aux personnes concernées ainsi qu'avec les préoccupations éthiques qui se posent en revenant trop fréquemment ou trop tardivement vers des personnes atteintes de pathologies graves ou incurables.

Notons que lorsqu'un fondement repose sur le consentement, la personne concernée dispose d'un droit de portabilité sur les données (article 20.1(a) RGPD).

Le responsable de traitement pourrait donc, dans l'intérêt de la recherche scientifique, ne pas faire reposer la recherche sur le fondement du consentement, ou pas uniquement, dès lors que d'autres fondements juridiques du RGPD sembleraient plus spécifiquement adaptés à ce type de traitement.

En pratique, en cas de retrait du consentement d'une personne, le responsable de traitement arrête par conséquent la collecte des données auprès de cette personne et

- Si le fondement juridique du traitement n'est pas le consentement, il n'y a pas d'effet sur la possibilité de poursuivre le traitement des données. Ce point est par ailleurs rappelé par l'article L. 1122-1-1 CSP et l'article 28 du Règlement (UE) n° 536/2014 relatif aux essais cliniques de médicament.

- Si le seul fondement du traitement était le consentement de la personne, alors les données devraient être effacées (article 17(1)(b) du RGPD), à moins que le responsable de traitement ait pour obligation légale de conserver les données, auquel cas le responsable de traitement
    2. informe la personne de ce fondement juridique (article 6(1)(c) et 9(2)(b) du RGPD), si elle n'a pas déjà reçu cette information,
    3. ne conserve ces données qu'aux seules fins de remplir cette obligation légale.

b) Intérêt public

La réalisation de traitements par une autorité publique ou par tout opérateur dès lors que l'intérêt public de la mission est reconnu dans la loi est également un fondement juridique proposé par le RGPD. Ce fondement peut trouver à s'appliquer pour les RIPH, voire pour d'autres recherches dans le domaine de la santé dans la mesure où la garantie de normes élevées de qualité ou de sécurité des soins de santé et des médicaments ou des dispositifs médicaux est reconnue comme relevant de l'intérêt public dans le droit européen (article 9 RGPD, article 28 du Règlement 536/2014 relatif aux essais cliniques de médicament) ainsi que dans le droit français (Chapitre III Section 3 de la LIL) et que la réalisation de ces activités entre dans le mandat du responsable de traitement.

Toutefois les interprétations sur les conditions d'application de ce fondement ne s'accordant pas sur ce point, il sera traité plus en détails au §IV.A.2 de la présente thèse.

c) Intérêt légitime

L'intérêt légitime du responsable de traitement apparait comme un fondement relativement flexible. Toutefois, l'autorité de contrôle britannique (*Information Commissioner's Office*, ICO) indique que lors de l'application de ce fondement, le responsable de traitement doit être particulièrement vigilant à l'équilibre avec l'intérêt des personnes concernées par le traitement des données les concernant. Par ailleurs, ce fondement ne devrait être appliqué qu'en cas de nécessité, conformément à la formulation du RGPD, c'est à dire lorsqu'il n'y a pas de fondement plus approprié et moins invasif.(6)

3. Droits des personnes

Le RGPD confère aux personnes plusieurs droits quant aux données les concernant, afin notamment de garantir que leur volonté est respectée. Plusieurs de ces droits existaient préalablement au RGPD, notamment dans la loi française.

Conformément au principe de transparence, les personnes reçoivent les informations relatives au traitement ainsi qu'à l'exercice de ces droits.

Ces droits s'exercent dans la limite des obligations réglementaires que peuvent par ailleurs avoir le responsable de traitement et le sous-traitant.

Le délégué à la protection des données de l'organisme a notamment pour fonction d'être le point de contact des personnes pour l'exercice de leurs droits (article 38(4) du RGPD).

a) Droit d'accès

Le droit d'**accès** (article 15 RGPD) permet à la personne d'obtenir

- La confirmation que des données la concernant sont ou non traitées par le responsable de traitement ;
- la communication d'une copie de l'ensemble des données personnelles la concernant détenues par le responsable de traitement, sous réserve que ces données ne portent pas atteinte aux droits et libertés d'autrui.

b) Droit à la portabilité

Le droit à la **portabilité** (article 20 RGPD) permet à la personne d'obtenir les données qu'elle a fournies (de manière active ou passive, telle que par observation de cette personne). Ceci ne concerne donc pas les données déduites (tel que par exemple un diagnostic médical). Ce droit porte sur des données plus limitées que le droit d'accès mais permet d'obtenir ces données dans un format de travail, aisément transférable. C'est avant tout un droit visant à favoriser la concurrence entre prestataires de service auxquels peut avoir recours la personne concernée.

Le droit à la portabilité ne s'applique que lorsque le fondement juridique du traitement est le consentement (article 20(1) du RGPD).

c) Droit de rectification

Le droit de **rectification** (article 16 RGPD) permet à la personne de faire rectifier une information erronée la concernant, ou la faire compléter.

d) Droit d'opposition

L'exercice du droit d'**opposition** (article 21 RGPD) empêche le traitement des données.

Pour ce qui concerne les données de santé, nous retiendrons que le droit d'opposition peut être exercé, pour des raisons tenant à la situation particulière de la personne, sauf à démontrer qu'il existe des motifs légitimes et impérieux qui prévalent sur les intérêts et les droits et libertés de la personne concernée. En recherche scientifique, le responsable de traitement pourrait refuser l'exercice de ce droit notamment lorsque le traitement des données de cette personne est nécessaire à l'exécution d'une mission d'intérêt public (article 21(6) du RGPD).[1]

Dans ce cas, le responsable de traitement se doit de peser la nécessité de poursuivre le traitement de ces données dans l'intérêt public et, le cas échéant, ne poursuivre leur traitement qu'à la seule finalité répondant à l'intérêt public (en addition d'une conservation répondant à des obligations réglementaires).

e) Droit à l'effacement

Le droit à l'**effacement** ou "droit à l'oubli" (article 17 RGPD) permet à la personne d'obtenir l'effacement des données la concernant.

Le droit à l'effacement ne s'exerce que dans certaines conditions précises :

- Lorsque le traitement reposait uniquement sur le fondement juridique du consentement et que la personne retire celui-ci (article 17(1)(b) du RGPD) ;
- La personne qui exerce son droit d'opposition peut également obtenir l'effacement des données la concernant, sauf si les motifs légitimes poursuivis par le responsable du traitement prévalent sur ceux de la personne concernée (article 17(1)(c) du RGPD).
- Quel que soit le fondement juridique, le droit d'effacement ne s'applique pas lorsque l'effacement des données est susceptible de rendre impossible ou compromettre gravement la réalisation des objectifs de la recherche scientifique (article 17(3)(d) du RGPD).

[1] Notons que la version anglaise du RGPD, article 21.6 mentionne la finalité d'intérêt

En pratique, à chaque demande d'opposition ou d'effacement, le responsable de traitement doit donc, pour en déterminer les effets sur la conservation et la poursuite de l'utilisation des données,

- évaluer si l'effacement est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs de la recherche, auquel cas les données ne sont pas effacées (article 17(3)(d) du RGPD) ;
- dans le cas contraire, évaluer si la poursuite du traitement répond à un motif légitime impérieux (article 17(1)(c) du RGPD).

Dans le cadre de la recherche en santé, il est en effet essentiel de disposer de bases de données non tronquées afin de ne pas biaiser les résultats de l'essai clinique ou de recherches ultérieures menées sur la base de données. L'équilibre entre cet intérêt public de l'exploitation des bases de données et l'effectivité du droit à l'opposition et à l'effacement reste à clarifier.

f) Droit à la limitation du traitement

Le droit à la **limitation** du traitement, qui permet à la personne de suspendre l'utilisation active des données la concernant dans une situation où le traitement est remis en cause :

- en cas d'usage illicite des données ;
- lorsque la personne conteste l'exactitude des données ;
- lorsque les données ne sont plus nécessaires au traitement mais sont encore nécessaires pour la constatation, l'exercice ou la défense de droits en justice (la limitation est alors une alternative à leur effacement immédiat) ;
- lorsque la personne s'est opposée au traitement, pendant la vérification si des motifs légitimes permettant de poursuivre tout de même le traitement prévalent sur les droits de la personne.

g) Autres droits

En complément, l'article 85.I de la LIL dispose que les personnes peuvent définir des directives relatives à la conservation, à l'effacement et à la communication des données personnelles les concernant après leur décès.

Les personnes concernées peuvent également introduire une réclamation auprès de l'autorité de contrôle concernée, notamment lorsqu'elles considèrent leurs droits comme étant bafoués.

	Droit à la	**Droit d'effacement**	**Droit d'opposition**

	portabilité (art. 20)	(art. 17)	(art. 21)
Consentement (art. 6.1.a)	Oui	Oui	Pas directement, mais droit de retirer le consentement
Exécution d'un contrat (art. 6.1.b)	Oui	Oui	Non
Obligation légale (art. 6.1.c)	Non	Non, sauf si c'est l'objet de l'obligation légale	Non
Sauvegarde des intérêts vitaux (art. 6.1.d)	Non	Oui, en l'absence de motif impérieux	Non
Intérêt public (art. 6.1.e)	Non	Oui, en cas d'application de l'opposition et si ne compromet par la recherche	Oui, si l'objectif de la recherche n'est pas compromis et si intérêt public non affecté
Intérêts légitimes (art. 6.1.f)	Non	Oui, si ne compromet par la recherche	Oui, si l'objectif de la recherche n'est pas compromis et en l'absence de motif impérieux

Figure 2 : Tableau résumant l'applicabilité des droits de la personne en fonction du fondement juridique et du type de traitement.

4. Minimisation

Le cadre réglementaire européen prévoit la minimisation des données (article 5(1)(c) du RGPD), notamment lors de leur collecte, afin de ne traiter que les données strictement nécessaires, pertinentes et non excessives, sans préjuger des moyens pratiques mis en œuvre à cet effet.

Ceci signifie qu'en l'existence d'un lien permettant la ré-identification de la personne, la donnée est à considérer comme étant à caractère personnel. En recherche scientifique donc, le remplacement des données directement identifiantes de la personne, telles que le nom et prénom par un code d'identification est une méthode de minimisation des données largement utilisée. Toutefois, ceci ne permet pas de rendre les données "anonymes" comme il est souvent décrit, par abus de langage, mais plutôt "non nominatives", "codées" ou encore "pseudonymisées". Ces données restent des données répondant à la définition de données à caractère personnel.(4)

Par ailleurs, la suppression de ce code d'identification suffit rarement à rendre la donnée anonyme. En pratique, et considérant la puissance des moyens informatiques actuels, notamment avec les collectes massives de données donnant la possibilité de recouper et compléter ces données par croisement de bases de données, une donnée

concernant un échantillon réduit de personnes peut être considérée comme étant à caractère personnel avec seulement un nombre réduit d'informations.

A titre d'exemple, dans un autre domaine, citons l'initiative en 2014 des taxis new-yorkais de rendre publiques les informations de 173 millions de courses dans la métropole. Par croisement, notamment avec des photos de célébrités montant dans un taxi, et donc un lieu et un horaire de prise en charge, des paparazzi ont pu obtenir la destination voire l'adresse du domicile de ces célébrités, ainsi que le montant des pourboires laissés aux chauffeurs.(7)

En pratique, dans le domaine de la recherche en santé, il est donc difficile de concevoir la réalisation d'une recherche sur des données anonymes. Elles seront donc, la plupart du temps, indirectement identifiantes ou pseudonymisées, en vertu du principe de minimisation voulant que toute donnée non strictement nécessaire au chercheur ne soit pas traitée (article 5(c) RGPD). Les données publiées devront dans la plupart des cas l'être sous forme de statistiques agrégées que l'on pourra alors considérer comme anonymes.

Avec l'avènement du *Big data* et des algorithmes, cette notion, dont on comprend par ailleurs aisément l'utilité, pourra poser question dans le domaine de la santé. En effet, les conclusions tirées par les algorithmes de *machine learning* reposent souvent sur une interprétation des données ou des catégories de données non anticipées par l'humain. Dès lors, il est difficile de prévoir à l'avance quelles seront les données strictement pertinentes. Il convient donc d'avoir une interprétation souple de ce principe dans le cadre de la constitution des bases ou "entrepôts" de données constitués à des fins de recherche ultérieure ou d'utilisation des algorithmes.

5. Exactitude

Les données personnelles faisant l'objet d'un traitement doivent être exactes et, lorsque nécessaire, tenues à jour. Le responsable d'un traitement de données personnelles doit s'assurer que les données inexactes au regard des finalités pour lesquelles elles sont traitées sont effacées ou rectifiées.

De cet impératif découle le droit de rectification pour les personnes concernées, leur permettant de signaler au responsable de traitement les éventuelles données inexactes et de demander leur correction (cf. § I.A.1).

6. Durée de conservation

Le RGPD fixe le principe général que des données personnelles doivent être conservées pour une durée définie à l'avance, soit de manière fixe soit selon une logique prédéterminée.

Cette durée doit être cohérente avec les finalités poursuivies. Au sein du set de données concernant une personne, la durée de conservation de différentes catégories de données peut varier selon les finalités poursuivies où le fondement juridique des traitements. Pour cette raison, les catégories de données conservées peuvent également avoir des durées de conservation distinctes, selon des droits d'accès distincts, notamment lorsque des obligations légales de conservation s'appliquent.

Toutefois, pour les données traitées à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques, le RGPD permet de conserver ces données pour une durée plus longue, à ces seules fins (article 5(1)(e) du RGPD).

Cette disposition prend toute son importance pour permettre la création de bases ou entrepôts de données, ou collections d'échantillons biologiques, conservés tant qu'ils comportent un potentiel intérêt scientifique, y compris lorsqu'aucun projet de recherche spécifique n'est en cours.

7. Sécurité

Le responsable de traitement est tenu de mettre en place les mesures techniques et organisationnelles appropriées afin de permettre la protection des données contre tout traitement non autorisé ou illicite, leur destruction ou leur endommagement.

Il apparait dès lors indispensable pour la documentation interne du responsable de traitement d'évaluer pour tout traitement réalisé si les mesures techniques et organisationnelles mises en place correspondent bien à l'état de l'art au regard de la sensibilité des données traitées.

8. Documentation de la conformité

Le RGPD entérine le principe d'une documentation en interne et d'un autocontrôle en matière de protection des données. Il marque le passage d'une logique de formalités administratives préalables, qui prévalait sous la Directive 95/46/CE, à celui de protection des données dès la conception et par défaut, appelé en anglais *Privacy by design and by default*.

Ainsi, le système doit garantir la sécurité et l'intégrité des données par des protections physiques et logiques régies par une politique de protection des données. Cette démarche peut notamment se traduire par la prise en compte de la protection des données personnelles dans le système qualité de l'organisme.

Le cas échéant, ce travail de documentation peut notamment reposer sur le DPO de l'organisme.

a) Registre des activités de traitement

Le principal recueil pour la documentation de la conformité est le registre des activités de traitement, prévu par l'article 30 du RGPD. Ce registre contient la liste des traitements de données personnelles mis en œuvre par le responsable de traitement.

b) Analyse d'impact relative à la protection des données

Lorsqu'un traitement est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes concernées, le responsable de traitement est tenu d'effectuer une analyse de l'impact relative à la protection des données (DPIA pour *Data Privacy Impact Assessment*). C'est une analyse de risque centrée sur la perte, fuite, ou utilisation illégitime des données personnelles. Elle évalue donc les sources de risque, leur vraisemblance, leur gravité, et les compare aux mesures techniques, organisationnelles et juridiques mises en œuvre afin de parer à ces risques.

La nécessité de consulter l'autorité compétente en matière de protection des DCP n'intervient donc que si l'analyse d'impact relative à la protection des données effectuée au titre de l'article 35 indique que le traitement présenterait un risque élevé si le responsable du traitement ne prenait pas de mesures pour atténuer le risque (article 36 du RGPD).

La CNIL(8) et le G29(9) ont précisé que les traitements de données sensibles sur des personnes vulnérables doivent notamment faire l'objet d'une analyse d'impact. Par conséquent, le traitement de données de santé de patients doit systématiquement faire l'objet d'une analyse d'impact préalable. Cette obligation est également valable en cas de traitement de données génétiques.

c) Cartographie des traitements et système documentaire

Afin de maitriser un ensemble de traitements complexes, il se dessine donc la nécessité de mettre en place une cartographie des traitements et de l'infrastructure des systèmes d'information sur lesquels ils reposent. Cette cartographie permettra en effet au responsable de traitement de garantir la tenue d'un registre exhaustif et de

disposer de toutes les informations nécessaires pour réaliser l'analyse d'impact relative à la protection des données.

Un autre axe de maitrise consiste dans l'intégration de la protection des données dans le système documentaire, dans l'esprit notamment du *Privacy by design*. Ainsi, toute mise en œuvre, conduite, modification, ou fin d'un traitement de données personnelles entre dans un cadre de gouvernance prédéfini au sein de l'organisme.

III. La protection des données personnelles appliquée à la recherche en santé

A. Différents types de recherche et démarches réglementaires

1. Principes généraux

La mise en place d'une recherche en santé nécessitant la création ou l'utilisation de bases de données à caractère personnel nécessite d'être conforme

- à la réglementation qui encadre ce type de recherche (recherche impliquant la personne humaine, analyse des caractéristiques génétiques, etc.) pour les aspects propres à leur méthodologie et spécificités,
- ainsi qu'à la loi n°78-17 du 6 janvier 1978 dite loi Informatique et Libertés (LIL) pour les aspects relatifs à la protection des données personnelles.

La protection des données de santé est définie par le Chapitre III Section 3 de la LIL. L'article 65 de la LIL définit le champ d'application de cette section et en exclut notamment des traitements mis en place au sein des établissements pour leurs besoins et les traitements mis en œuvre par les institutions publiques pour leur fonctionnement, mais également les traitements fondés sur le consentement des personnes concernées. Ceci signifie en pratique que ces traitements ne sont pas soumis à l'autorisation de la CNIL ou à la conformité à des référentiels prévues par cette section.

1. Recherche impliquant la personne humaine (RIPH)

La loi n° 2012-300 du 5 mars 2012 a introduit la notion de **recherche impliquant la personne humaine (RIPH)** et le Décret n°2016-1537 du 16 novembre 2016 en précise la définition. Ce terme désigne les recherches organisées et pratiquées sur l'être humain en vue du développement des connaissances biologiques ou médicales (art. L. 1121-1 CSP).

Ceci désigne donc les recherches prospectives nécessitant le suivi d'un patient ou d'un volontaire sain. Elles ne comprennent donc pas les recherches sur les échantillons déjà collectés ou les bases de données préalablement constituées. Les recherches sur les registres existants ou les dossiers médicaux ne sont donc pas considérées comme des recherches impliquant la personne humaine si elles ne nécessitent pas de revenir au patient pour collecter des données supplémentaires.

Pour ces recherches, la Loi n° 2016-41 du 26 janvier 2016 (10) a donné compétence au CPP pour l'évaluation des aspects relevant de la protection des DCP (article 76 LIL). Cette évaluation entrait auparavant dans le champ du CCTIRS pour toutes les recherches dans le domaine de la Santé.

Le Code de la santé publique distingue trois catégories de recherches impliquant la personne humaine.

a) Types de RIPH

(1) Les recherches interventionnelles qui comportent une intervention sur la personne non justifiée par sa prise en charge habituelle

Elles étaient auparavant désignées comme « recherches biomédicales ». Ces recherches, mentionnées au 1° de l'article L. 1121-1 du Code de la santé publique, comportent une **intervention sur les personnes non dénuée de risques pour celles-ci**. On compte parmi ces recherches, celles portant sur des médicaments, sur les autres produits de santé (mentionnés à l'article L. 5311-1 CSP), et aussi celles ne portant pas sur des produits de santé (par exemple les recherches interventionnelles en physiologie, en chirurgie, mais également les recherches portant sur des denrées alimentaire).

A noter que les dispositions du décret du 16 novembre sont applicables à toutes les recherches sur la personne humaine initiées après le 17 novembre 2016. Le décret présente donc cette particularité qu'il s'applique à la fois aux dispositions de la loi n°2012-300 du 5 mars 2012(11), dite loi Jardé dans leur rédaction initiale et dans leur rédaction modifiée par l'Ordonnance.

(2) Les recherches interventionnelles qui ne comportent que des risques et des contraintes minimes

Elles peuvent comporter l'utilisation de produits de santé, mais ceux-ci le sont alors dans les conditions habituelles d'utilisation. Elles **peuvent comporter des actes peu invasifs** (prélèvement veineux sanguins, imagerie non invasive…). Ainsi, une partie de ces recherches correspond à ce qui était antérieurement désigné comme « recherches visant à évaluer les soins courants ».(12)

La liste exhaustive des interventions ne comportant que des risques et des contraintes minimes est fixée par l'Arrêté du 12 avril 2018 fixant la liste des recherches mentionnées au 2° de l'article L. 1121-1 du Code de la santé publique.

(a) Mesures *transitoires*

En l'attente de l'entrée en vigueur du Règlement 536/2014 sur les essais cliniques portant sur le médicament, les recherches interventionnelles portant sur le médicament sont exclues de cette catégorie. En effet, ces dernières sont actuellement régies par la Directive 2001/20/EC qui ne prévoit pas de régime allégé pour leur autorisation. Ainsi, les recherches portant sur le médicament qui sont interventionnelles et qui ne comportent que des risques et des contraintes minimes seront automatiquement soumises au 1° de l'article L. 1121-1 du Code de la santé publique.

(3) Recherches non interventionnelles

Ce sont les recherches qui comportent un ou plusieurs actes ou procédures réalisés conformément à la pratique courante et mentionnées à l'arrêté du 12 avril 2018 fixant la liste des recherches mentionnées au 3° de l'article L. 1121-1 du Code de la santé publique.

Ces recherches peuvent porter sur des produits de santé ou des stratégies ou pratiques de soin et de prise en charge, sans randomisation des participants, ou sur des prélèvements d'échantillons biologiques selon des modalités jugées comme ne présentant pas de risque ou de contrainte.

b) Démarches réglementaires

Les RIPH nécessitent avant leur mise en œuvre un avis favorable d'un comité de protection des personnes (CPP) (article L. 1123-6 CSP).

Les RIPH de catégorie 1° nécessitent également l'autorisation préalable de l'Agence nationale de sécurité du médicament et des produits de santé (ANSM).

Sur les aspects relatifs à la protection des données personnelles, la CNIL peut établir des référentiels ou règlements types (art. 66.II LIL). Lorsque la recherche est conforme à l'un de ces référentiels applicables aux RIPH que sont les Méthodologies de référence (MR) 001 et 003, le traitement peut être mis en œuvre sans démarche supplémentaire auprès de la CNIL.

Lorsque la recherche n'est pas conforme à une MR, elle doit être autorisée par la CNIL, après avis du CPP.

(1) Méthodologie de référence 001 (MR-001)

La Délibération n° 2018-153 du 3 mai 2018, ou MR-001 est applicable aux recherches en santé avec recueil du consentement de la personne.

Elle concerne donc les RIPH de catégories 1° et 2° ainsi que les RIPH de catégorie 3° lorsqu'un examen des caractéristiques génétiques nécessite le recueil du consentement des participants.

Cette méthodologie, la première publiée par la CNIL dans le domaine de la recherche en santé, en 2006 et revue récemment, constitue un guide pour le traitement des données dans la recherche qui, si le responsable de traitement s'y conforme, permet d'éviter des démarches administratives. Cette méthodologie reprend les principes réglementaires tout en les appliquant au domaine visé.

Cette méthodologie, comme l'on montré ses évolutions successives depuis 2006, est un document évoluant avec l'avancée des techniques. A ce jour, notons par exemple qu'elle ne permet pas, dans son champ d'application, la signature électronique du consentement par les patients.

(2) Méthodologie de référence 003 (MR-003)

La Délibération n° 2018-154 du 3 mai 2018 vise, elle, les RIPH ne nécessitant qu'une information simple des participants à la RIPH. Ce sont donc les RIPH de catégorie 3° sans examen des caractéristiques génétiques.

Ses dispositions pratiques sont très similaires à celles de la MR-001.

(3) Exception en cas de recherche fondée sur le consentement

L'article 65 de la LIL exclut du champ du Chapitre III Section 3 de la LIL, et par la même dispense d'une autorisation préalable ou de la conformité à une Méthodologie de référence, les traitements fondés notamment sur le consentement explicite donné par les personnes pour des finalités déterminées (tel que défini à l'article 9(2)(a) du RGPD).

Il convient toutefois de rappeler que le fondement du consentement comporte d'importantes limitations méthodologiques, opérationnelles et éthiques (cf. §II.C.2.a).

2. Recherches sur des données collectées dans un autre cadre

Ces recherches n'impliquent pas directement la personne humaine dans la mesure où les données sont déjà collectées.

a) Principes généraux

La mise en place d'une recherche en santé sur les bases de données à caractère personnel nécessite d'être conforme à la loi n°78-17 du 6 janvier 1978 dite loi Informatique et Libertés (LIL).

Les traitements cités par l'article 65 sont exclus du Chapitre III Section 3. Ceci signifie qu'ils ne sont pas soumis à l'autorisation de la CNIL prévue par ce chapitre. Ces exceptions concernent notamment des traitements mis en place au sein des établissements pour leurs besoins et les traitements mis en œuvre par les institutions publiques pour leur fonctionnement, mais également les traitements fondés sur le consentement des personnes concernées.

Les recherches, études ou évaluations dans le domaine de la santé entrent pour la plupart dans le champ du Chapitre III Section 3. Elles sont donc soumises à l'avis du Comité d'expertise pour les recherches, les études et les évaluations dans le domaine de la santé (CEREES), et à l'autorisation de la CNIL. Le dossier de demande est soumis à l'INDS qui assure le secrétariat unique (articles L. 1462-1 CSP et 66-II LIL).

b) MR-004

La Délibération n° 2018-155 du 3 mai 2018, ou MR-004, fait entrer dans son champ les recherches en santé mises en œuvre à partir de données préalablement collectées mais sans appariement de données depuis plusieurs sources.

Cette MR-004 peut donc potentiellement être applicable à une multitude de recherches, évaluations ou études dans le domaine de la santé sur des données collectées dans un autre cadre, que ce soit lors du soin ou lors d'une autre recherche. Toutefois, plusieurs limitations sont à soulever.

(1) Absence d'appariement

La MR-004 exclut de son champ les recherches nécessitant un appariement "entre les données déjà existantes d'un même individu issues de plusieurs centres participants". Ceci peut potentiellement exclure les recherches qui nécessiteraient le suivi d'un patient dans différents établissements qui pourraient le prendre en charge, ou l'appariement de données médicales que détiendrait un médecin avec les échantillons biologiques stockés par ailleurs.

(2) Information des personnes

Il convient de saluer dans cette MR-004, l'apparition de nouvelles modalités d'information des personnes. Ainsi, la MR-004 permet au responsable de traitement de d'informer les personnes via un support unique sous réserve que les personnes aient été préalablement informées de son existence et de sa localisation (ex : adresse url du site internet où se trouve l'information).

Ceci permet notamment l'apparition de sites internet délivrant aux personnes une information dynamique auquel elles peuvent se référer lorsqu'elles souhaitent plus d'informations sur les recherches menées ultérieurement à partir des données les concernant.

La MR-004 clarifie également qu'il n'est pas nécessaire de délivrer aux personnes concernées une nouvelle information pour chaque nouveau traitement dès lors que la recherche ultérieure est compatible avec l'information délivrée initialement.

Toutefois, la MR-004 ne fait pas entrer dans son champ les exceptions prévues par les articles 13 et 14 du RGPD permettant au responsable de traitement de ne pas délivrer aux personnes toutes les informations prévues relatives au traitement des données lorsque cette information exigerait un effort disproportionné au regard des finalités.

c) Procédure d'autorisation

Les recherches qui n'entrent pas dans le champ de la MR-004 nécessitent une autorisation ponctuelle de la part de la CNIL, prise après avis du Comité d'expertise pour les recherches, les études et les évaluations dans le domaine de la santé (CEREES).

Le dossier de demande d'autorisation est déposé auprès de l'Institut national des données de santé (INDS) qui assure le secrétariat et peut également rendre un avis sur l'intérêt public que présente toute recherche en santé.

3. Recherches sur échantillons

En tant que support de données à caractère personnel, les recherches sur les collections d'échantillons biologiques entrent dans le champ de la réglementation relative à la protection des données personnelles.

En effet, ces échantillons contiennent pour la plupart du matériel qui est un support de données génétiques tout en étant un matériel considéré dans le Code de la santé publique comme possédant un statut particulier.

La donnée génétique est définie par le RGPD comme toute "donnée à caractère personnel relative aux caractéristiques génétiques héréditaire ou acquises d'un personne physique qui donnent des informations uniques sur la physiologie ou l'état de santé de cette personne physique et qui résultent, notamment, d'une analyse d'un échantillons biologique de la personne physique en question". Dès lors, elle englobe les données relatives à la génétique, à la génomique, et aux mutations. On comprend également qu'un échantillon biologique, même non identifié, peut livrer des données personnelles.

La recherche sur les échantillons se voit donc appliquer

- le cadre réglementaire relatif à la protection des données,
- le cadre réglementaire des RIPH, lorsqu'ils sont collectées et utilisées dans le contexte de ce type de recherche,
- et les dispositions relatives aux échantillons biologiques, découlant notamment de la loi ° 2004-800 du 6 août 2004 relative à la bioéthique.

d) Prélèvement d'échantillons biologiques réalisé pour les besoins de la recherche ou à l'occasion d'un geste réalisé dans le cadre du soin

Comme mentionné au paragraphe III.1.a), les arrêtés du 12 avril 2018 fixant la liste des recherches mentionnées au 2° et au 3° de l'article L. 1121-1 du CSP viennent encadrer les recherches sur les prélèvements effectués au cours du soin et dont le volume est adapté pour les besoins de la recherche ou qui sont directement utilisés à des fins de recherches. Elles constituent des recherches sur la personne mentionnées au 2° ou au 3° de l'article L. 1121-1 du CSP si le prélèvement est motivé par la recherche, adapté à la recherche, ou utilisé immédiatement après le prélèvement effectué.

e) Recherches sur les échantillons biologiques existants

Les prélèvements biologiques collectés au cours du soin ou d'une recherche antérieure peuvent être utilisés dans le cadre d'une nouvelle recherche.

Ces recherches sont un type particulier de recherche sur des données collectées dans un autre cadre.

(1) Échantillons ayant permis d'établir un diagnostic

Il convient de souligner que, conformément à l'article R. 6211-44 CSP, les échantillons ayant permis d'établir un diagnostic sous soumis à une obligation de conservation durant dix ans. Ils ne peuvent donc pas être utilisés à des fins de recherche si celle-ci empêche la conservation par le médecin anatomopathologiste.

(2) Recherche à partir d'une collection d'échantillons prélevés au cours d'une recherche antérieure

Toute collection d'échantillons biologiques conservée en dehors du cadre d'une RIPH devra être préalablement déclarée au Ministère chargé de la Recherche, conformément aux articles L. 1243-3 et R. 1243-49 CSP.

4. Recherches sur le SNDS

La loi n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé (LMSS de 2016)(10) a prévu la création du Système national des données de santé (SNDS) afin de mettre à disposition du public et de la communauté scientifique les données recueillies à titre obligatoire et destinées aux services ou aux établissements publics de l'Etat ou des collectivités territoriales ou aux organismes de sécurité sociale (art. 193 de la LMSS de 2016).

Le SNDS rassemble et met à disposition les données du Programme de médicalisation des systèmes d'information (PMSI), du Système national d'information inter-régimes de l'assurance maladie (SNIIRAM), du système national d'information interrégimes de l'assurance maladie et des complémentaires de santé, les données sur les causes de décès (données du CépiDc), et un échantillon représentatif des données de remboursement (Echantillon généraliste de bénéficiaires, EGB).

Les modalités de mise en œuvre du SNDS et les conditions d'accès sont définies par le Décret n° 2016-1871 du 26 décembre 2016.(13)

Cet accès est organisé par deux institutions :

- Le Système national des données de santé (SNDS), dont le rôle est de rassembler et mettre à disposition les données de plusieurs bases mentionnées à l'article L. 1461-1 du CSP
- L'institut national des données de santé (INDS) qui veille notamment à la qualité des données et aux conditions de mise à disposition et s'assure de l'intérêt public des recherches qui sont conduites avec ces données, conformément à l'article L. 1462-1 du CSP.

Le principe général des accès est organisé selon des modalités distinctes (art. L.1461-3 CSP) :

- des organismes possédant un accès permanent pour des raisons de service public
- un accès sur autorisation
    - direct pour les laboratoires ou bureaux d'étude
    - par l'intermédiaire d'un laboratoire de recherche ou d'un bureau d'études pour les organismes produisant ou commercialisant des produits de santé lorsque ceux-ci ne peuvent démontrer que les modalités de mise en œuvre du traitement rendent impossible l'utilisation des données à des fins de promotion des produits de santé
    - par l'intermédiaire d'un laboratoire de recherche ou d'un bureau d'études pour les banques ou assurances lorsque ceux-ci ne peuvent démontrer que les modalités de mise en œuvre du traitement rendent impossible l'utilisation des données à des fins de modification des contrats d'assurance.

a. Mesures de simplification

Parmi les mesures destinées à simplifier l'accès aux données du SNDS, notons :

- Un accès simplifié à l'EGB, échantillon représentatif de la population française regroupant 300 000 personnes, en raison de son absence d'exhaustivité qui rend la réidentification des personnes plus difficile ;(14)
- Deux méthodologies de référence, permettant un accès au PMSI et aux données de l'Agence technique de l'information sur l'hospitalisation (ATIH) sous réserve notamment de l'absence d'appariement avec une base de données existante n'est réalisé :
    - La MR-005 pour l'accès par les établissements de santé et leurs fédérations,(15)

- La MR-006 pour l'accès par les organismes produisant ou commercialisant des produits de santé, via un laboratoire de recherche ou un bureau d'étude.(16)

<table>
<tr><td rowspan="2">Type de recherche</td><td colspan="3">RIPH</td><td rowspan="2">Recherche organisée et pratiquée sur la personne qui n'a pas les finalités d'une RIPH</td><td rowspan="2">Recherche sur des données ou échantillons collectés dans un autre cadre</td><td rowspan="2">Recherche sur le SNDS</td></tr>
<tr><td>Catégorie 1°</td><td>Catégorie 2°</td><td>Catégorie 3°</td></tr>
<tr><td>Champ d'application des MR</td><td colspan="2" rowspan="2">MR-001</td><td>MR-003
si pas d'examen des caractéristiques génétiques</td><td colspan="2" rowspan="2">MR-004</td><td rowspan="2">MR-005
/
MR-006</td></tr>
<tr><td>MR-001
si un examen des caractéristiques génétiques est effectué (car consentement requis)</td></tr>
<tr><td rowspan="2">En cas de non conformité avec les MR</td><td colspan="3">Avis CPP + autorisation CNIL (+ saisine INDS possible sur l'intérêt public)</td><td colspan="3">INDS (secrétariat unique + intérêt public) + avis CEREES + autorisation CNIL</td></tr>
<tr><td colspan="6">➔ La non-conformité peut notamment porter sur les aspects suivants : information de la personne, nature des données traitées, destinataires des données directement ou indirectement identifiantes, risque résiduel élevé suite à l'analyse d'impact, etc.</td></tr>
</table>

Figure 3 : Champ d'application des différentes Méthodologies de référence en fonction du type de recherche.

B. Transparence envers les personnes concernées

Préalablement à tout traitement des données les concernant, les personnes doivent être informées du traitement, de façon concise, transparente, compréhensible et aisément accessible, en des termes clairs et simples (article 12 RGPD).

1. Lorsque les données sont collectées auprès de la personne

L'article 13 du RGPD fixe le contenu de l'information délivrée lorsque les données sont collectées directement auprès des personnes.

La personne reçoit notamment les informations suivantes :

- l'identité et les coordonnées du responsable de traitement, et le cas échéant, de son représentant dans l'UE, afin de pouvoir les contacter facilement ;
- les coordonnées du délégué à la protection des données, le cas échéant ;
- la finalité du traitement et sa base juridique ;
- les destinataires ou catégories de destinataires des données ;
- l'existence éventuelle de transfert des données vers des pays tiers et les garanties offertes en matière de protection des données dans le cadre de ce transfert ;
- la période de conservation des données ou les critères pour la déterminer ;
- les droits dont dispose la personne ;
- si la fourniture des informations demandées est rendue obligatoire par la réglementation ou à la conclusion d'un contrat, ainsi que les conséquences éventuelles de leur non-fourniture ;
- l'existence d'une prise de décision automatisée, des informations utiles concernant la logique sous-jacente, et des conséquences envisagées d'un tel traitement pour la personne.

a) Granularité de la notion de finalité

En informant la personne concernée des finalités du traitement, le responsable de traitement doit être capable de fournir une information claire et simple tout en étant suffisamment précise. Sur cet aspect, le Groupe de l'article 29 (voir section II.A.4) indique que mentionner "des fins de recherche" n'est pas suffisamment clair car le type de recherche visé n'est pas précisé.(17)

La *Agencia Española de Protección de Datos* (AEPD), autorité de contrôle espagnole, dans une note juridique publiée en 2018, prenant en compte les considérants 33, 52, 157, 159 du RGPD, interprète que le consentement que pourrait donner une personne pour des recherches ultérieures ne saurait se résumer à un domaine restreint, comme

par exemple un sous-type de cancer. Ainsi, une granularité pertinente consisterait en des finalités définies par exemple comme la recherche en oncologie voire dans des domaines plus étendus.(18)

b) Recherches impliquant la personne humaine

La plupart des recherches en santé nécessitant une collecte de données auprès de la personne va répondre à la définition de RIPH.

Conformément à l'article L. 1122-1-1 du CSP, toute personne participant à une recherche sur la personne doit avoir préalablement reçu les informations décrites à l'article L. 1122-1 du CSP. Parmi ces informations, certaines correspondent aux informations prévues par le RGPD, d'autres s'y additionnent, notamment sur les sujets des risques et contraintes liés à la participation de la personne à la recherche.

Il est à noter que l'impératif de concision prévu par le RGPD sera parfois difficile à atteindre en raison de ce double impératif d'information applicable aux RIPH.

Dans le cas d'une recherche mentionnée au 1° de l'article L. 1121-1 du CSP, son consentement exprès doit être recueilli par écrit. Pour une recherche mentionnée au 2° du même article, le consentement exprès est nécessaire sans que la forme n'en soit précisée. Enfin, seule une non opposition du patient est nécessaire pour la participation à une recherche mentionnée au 3°.

2. Lorsque le traitement est réalisé à partir de données collectées dans un autre cadre

L'article 14 précise le contenu de l'information lorsque les données n'ont pas été collectées directement auprès des personnes, c'est à dire en cas d'utilisation secondaire de données collectées dans un autre cadre.

En plus du contenu déjà prévu par l'article 13 et cité précédemment, dans le cadre d'une utilisation ultérieure des données, l'article 14 prévoit que la personne est informée des catégories de données traitées et de leur provenance initiale. De manière logique, l'information du caractère obligatoire de la fourniture des données, prévu par l'article 13, n'est pas mentionnée dans l'article 14, dans la mesure où le responsable de traitement n'obtient pas ces données auprès de la personne concernée.

L'article 14 prévoit des exceptions à la délivrance de ces informations aux personnes concernées :

- lorsque la personne concernée dispose déjà de ces informations ;

- lorsque la fourniture de ces informations se révèle impossible, exigerait des efforts disproportionnés ou lorsque celle-ci est susceptible de compromettre gravement les finalités du traitement ;
- lorsque le droit communautaire ou national prévoit des modalités alternatives visant à protéger les intérêts des personnes ;
- lorsque les données doivent rester confidentielles en vertu notamment d'une obligation de secret professionnel.

a. Finalités compatibles

Conformément aux articles 5 et 14 du RGPD, en cas de traitement ultérieur, la finalité de ce nouveau traitement doit rester compatible avec la finalité initiale. Dans le cas contraire, la personne devra en toute logique être informée de la nouvelle finalité.

Dans le cadre du RGPD, les traitements ultérieurs dans le domaine de la recherche scientifique notamment jouissent du principe de finalité compatible (article 5). A ce titre, le considérant 33 du RGPD, explique que dans la mesure où il n'est pas toujours possible de cerner entièrement la finalité des recherches au moment de la collecte des données, la personne devrait pouvoir consentir à l'utilisation des données la concernant dans certains domaines de recherche. L'autorité de contrôle espagnole, l'AEPD, s'est positionnée sur ce sujet par le biais d'une note juridique indiquant que le consentement de la personne doit pouvoir être donné non pas pour chacune des recherches spécifiques, non plus pour des catégories trop limitées de recherches, tel que la recherche sur un type particulier de cancer, mais pour des champs larges, tel que par exemple la "recherche en oncologie".(18)

b. Efforts disproportionnés

Le G29 indique que lorsqu'un responsable de traitement souhaite invoquer l'effort disproportionné qu'exigerait l'information des personnes, celui-ci doit mettre en balance l'effort nécessaire avec l'incidence des effets sur la personne concernée dans le cas où elle ne recevrait pas ces informations.

Le responsable de traitement met alors en place des mesures appropriées telles que l'affichage de ces informations sur un support accessible au public, ainsi que d'autres mesures telles que la réalisation d'une analyse d'impact, la minimisation des données, etc.

Il est à noter que si le responsable de traitement n'informe pas les personnes en raison de l'effort disproportionné exigé par l'information exclut à ce jour la recherche

en santé de la MR-004 et nécessite par conséquent une autorisation préalable de la CNIL.

c. Cas des échantillons biologiques

En addition des dispositions du RGPD, les échantillons biologiques doivent répondre aux dispositions du Code de la santé publique et du Code civil.

Par conséquent, la personne sur laquelle a été opéré ce prélèvement ou cette collecte, doit être informée au préalable de leur finalité d'utilisation.

Deux situations sont à distinguer :

- Pour une recherche ne portant pas sur l'examen des caractéristiques génétiques, il peut être dérogé à l'obligation d'information (article L. 1211-2 CSP)
    - lorsque celle-ci se heurte à l'impossibilité de retrouver la personne concernée,
    - ou lorsqu'un des comités consultatifs de protection des personnes mentionnés à l'article L. 1123-1, consulté par le responsable de la recherche, n'estime pas cette information nécessaire.
    - Si le patient est décédé, dans la mesure où il n'a pas exprimé d'opposition de son vivant, l'utilisation des échantillons et des données à de nouvelles fins est possible, sauf s'il s'agit de tissus ou cellules germinaux.
- Pour une recherche portant sur l'examen des caractéristiques génétiques, il peut être dérogé à l'obligation d'information lorsque celle-ci se heurte à l'impossibilité de retrouver la personne concernée et qu'un des comités consultatifs de protection des personnes mentionnés à l'article L. 1123-1, consulté par le responsable de la recherche, n'estime pas cette information nécessaire (article L. 1131-1-1 CSP).

Toutefois, ces dérogations ne sont pas admises lorsque les éléments initialement prélevés consistent en des tissus ou cellules germinaux (article L. 1211-2 CSP). Dans ce dernier cas, toute utilisation pour une fin autre que celle du prélèvement initial est interdite en cas de décès de l'intéressé.

On pourra s'interroger sur l'articulation de ces dispositions avec l'information prévue par le RGPD. Le Code de la santé publique, dans ces articles, ne mentionne que l'information portant sur la finalité d'utilisation. Dès lors, si la personne a été informée de la finalité d'utilisation de l'échantillon mais que l'information reste

lacunaire au regard du RGPD, il est possible que seuls le CEREES et la CNIL doivent être consultés.

3. Retrait de consentement

On entend par « retrait » de consentement, l'acte du patient d'exprimer la fin de son consentement à participer à la recherche. C'est donc en pratique un arrêt du consentement plutôt qu'un retrait. Le consentement qui avait été précédemment donné n'est plus valide. En revanche, sa validité entre le moment de la signature et le retrait n'est pas remise en cause.

Le retrait du consentement peut se faire à l'oral comme à l'écrit, quelle que soit la modalité de recueil du consentement initial. Si l'écrit est la solution à favoriser en termes de traçabilité, on ne peut pas l'imposer au patient.

Dans le cas d'un retrait de consentement oral, il est recommandé de mettre en place une procédure garantissant la traçabilité du retrait de consentement.

L'article 2 de l'ordonnance n°2016-800 du 16 juin 2016 a clarifié le devenir des données en cas de retrait de consentement par un ajout au dernier alinéa de l'article L. 1122-1-1 du CSP :

Dans le cas où la personne se prêtant à une recherche a retiré son consentement, ce retrait n'a pas d'incidence sur les activités menées et sur l'utilisation des données obtenues sur la base du consentement éclairé exprimé avant que celui-ci n'ait été retiré.

Dans le cas où le traitement de données ne repose pas sur le fondement du consentement, un retrait de consentement à la participation à la recherche n'a pas lieu, selon les dispositions du RGPD, d'avoir un impact sur l'utilisation des données.

Toutefois, le RGPD prévoit notamment que les personnes concernées ont un droit d'opposition au traitement, qui pourrait entrainer l'effacement des données. On pourra s'interroger sur l'articulation entre cette disposition et l'exercice des droits des personnes prévus par le RGPD, et s'il faut considérer les deux textes comme applicables ou si les dispositions du Code de la santé publique relatives aux RIPH sont des règles spéciales qui dérogent aux règles générales (*Specialia generalibus derogant*).

Pour les essais cliniques portant sur le médicament, l'article 28(3) du Règlement (UE) 536/2014 (qui n'est pas encore entré en application) prévoit des dispositions

similaires, qui distingue explicitement le retrait de consentement de l'exercice des droits prévus par la Directive 95/46/CE, désormais remplacée par le RGPD :

[...] Sans préjudice de la directive 95/46/CE, le retrait du consentement éclairé n'a pas d'incidence sur les activités déjà menées et sur l'utilisation des données obtenues sur la base du consentement éclairé avant que celui-ci ne soit retiré.

C. Transfert des données en dehors de l'Union européenne

Conformément au Chapitre V du RGPD, le transfert de données à caractère personnel en dehors de l'Union européenne ne peut avoir lieu que dans les conditions suivantes :

- S'il s'effectue vers un Etat reconnu par la Commission européenne comme assurant un niveau de protection adéquat à l'égard du traitement des données (art. 45 RGPD)
- S'il s'effectue selon des garanties appropriées (art. 46 RGPD), et notamment si des Clauses contractuelles types (CCT) publiées par la Commission européennes sont établies entre les parties procédant au transfert (art. 46.2(c) RGPD).
- S'il s'effectue au sein d'une entreprise dont les filiales ont établi des règles d'entreprise contraignantes validées par la Commission européenne (art. 47 RGPD)
- S'il s'effectue, à titre dérogatoire, dans des situations particulières (art. 49 RGPD).

1. Décisions d'adéquation

La Commission publie, après évaluation, la liste des pays pour lesquels la protection garantie aux données personnelles est jugée équivalente à celle prévue par le RGPD.

Dans certains cas, cette décision pourra être partielle et ne s'appliquer qu'un champ restreint de traitements.

Notons par exemple que le *EU - U.S. Privacy Shield* (Bouclier de protection des données UE - Etats-Unis)(19) formule de façon ambiguë son applicabilité aux données pseudonymisées ("key-coded", c'est à dire lorsque chaque patient est identifié par un code remplaçant les données nominatives) utilisées dans le contexte de la recherche en santé (Annexe II, III.14.g.i du *Privacy Shield*). Ceci a conduit le G29, dans son interprétation, à exclure ces données et traitements du champ d'application de ce *Privacy Shield*.(20)

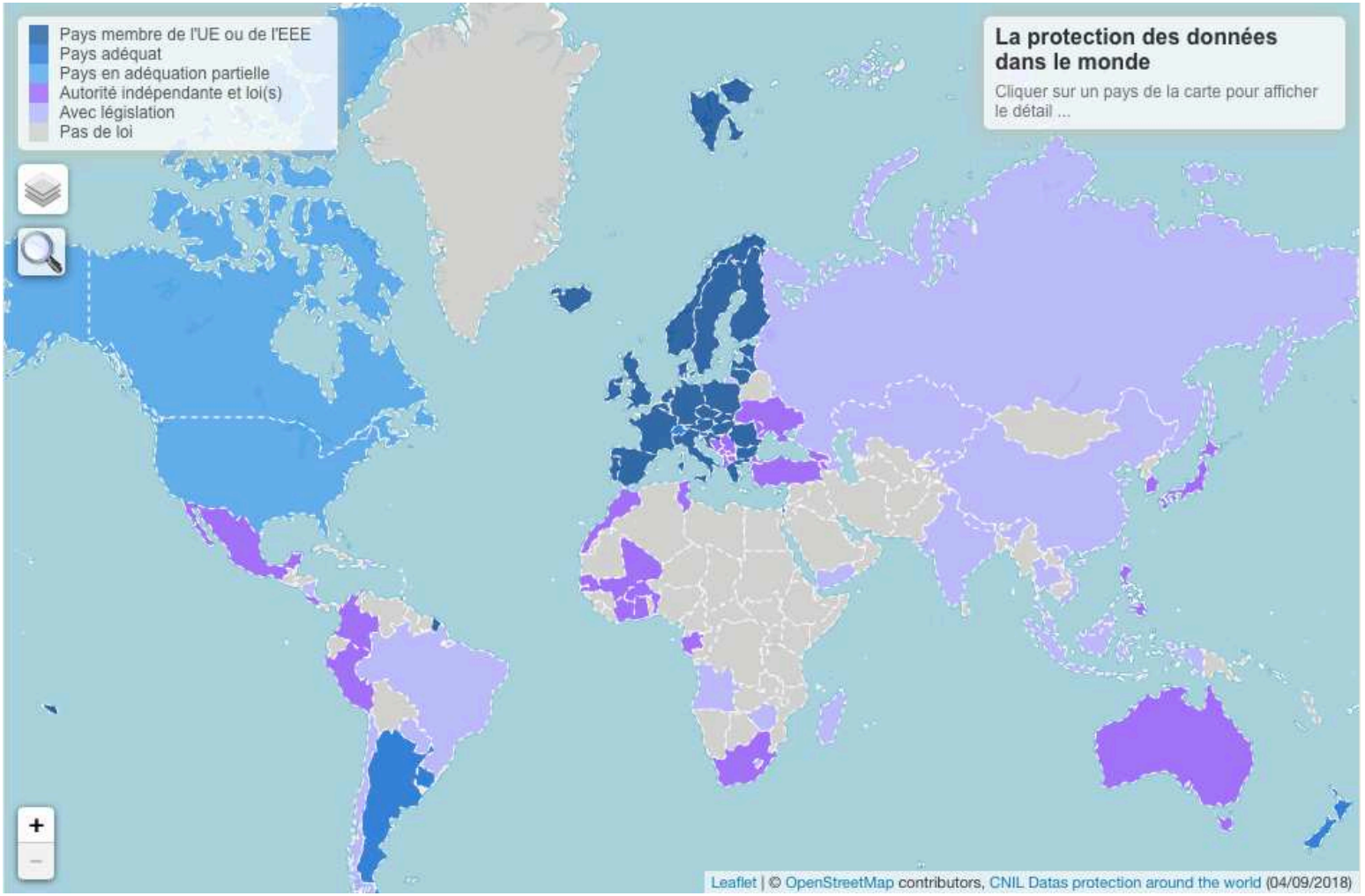

Figure 4 : Carte de la réglementation relative à la protection des données personnelles dans le monde (CNIL, 24/1/2019)

2. Dérogations pour des situations particulières

Dans le cadre de la Directive 95/46/CE, le G29 a indiqué que ces exceptions doivent s'interpréter de manière restrictive et ne s'appliquent pas dans le cas de transferts massifs ou répétés.(21) En pratique, dans la recherche clinique, le consentement du patient au transfert n'est donc pas un fondement suffisant pour permettre le transfert de données.

Il semble donc que les dérogations prévues par l'article 49 du RGPD doivent s'interpréter selon un champ similaire, en l'absence de nouvelle indication.

3. Règles d'entreprise contraignantes

Ces règles d'entreprise contraignantes, ou BCR (*Binding Corporate Rules*), peuvent être élaborées pour une entreprise multinationale afin de permettre les transferts de données transfrontaliers au sein de ses filiales. Elles sont validées par la Commission européenne au sein d'un long processus (deux ans environ). Ces règles peuvent également ne prévoir dans leur champ d'application que les transferts liés à certains types de traitements.

4. Clauses contractuelles types

En synthèse, les clauses contractuelles types (CCT), malgré l'absence de réécriture par la Commission européenne depuis l'entrée en vigueur du RGPD restent le principal fondement permettant le transfert de données personnelles dans le cadre de coopérations internationales pour la recherche en santé.

Il en existe plusieurs modèles :(22)

- Des clauses pour le transfert d'un responsable de traitement vers un sous-traitant situé en dehors de l'EEE
    - Un modèle selon la Décision de la Commission 2001/497/CE du 15 juin 2001, prévoyant notamment une responsabilité solidaire des parties en cas de brèche de données,
    - Un modèle selon la Décision de la Commission 2004/915/CE du 27 décembre 2004, prévoyant notamment une responsabilité limitée pour chaque partie au dommage effectif subi ;
- Des clauses pour le transfert de données d'un responsable de traitement vers un autre responsable de traitement, selon la Décision de la Commission 2010/87/UE du 5 février 2010.

Ces CCT s'utilisent telles quelles, sans modification possible par les parties, au-delà des parties à compléter pour préciser le champ de l'accord.

Notons que depuis 2016, ces clauses font l'objet de critiques et d'une remise en cause(23) dans la mesure où elles ne garantissent pas notamment que la loi du pays destinataire ne puisse prévaloir sur cet accord et entraver la sécurité des données transférées, comme ce peut être le cas avec le *Patriot Act* aux Etats-Unis.

IV. Difficultés à surmonter

A. Consentement et fondements juridiques

1. Articulation en articles 6 et 9 du RGPD

Le mécanisme du fondement juridique sur lequel peut reposer un traitement fait l'objet de plusieurs interprétations.

L'interprétation la plus largement acceptée semble être que tout traitement doit répondre à un fondement cité à l'article 6(1) du RGPD et, lorsque le traitement porte sur des données sensibles, remplir une des conditions de l'article 9(2).

Toutefois, certaines interprétations, en Allemagne notamment, appliquant le principe du droit civil hérité de la loi romaine, indiquent que l'article 9 est plus spécifique que l'article 6 et donc applicable de manière exclusive lorsque le traitement porte sur des données sensibles.(24) L'article 6 s'appliquerait alors aux traitements ne portant pas sur des données sensibles et l'article 9 aux traitements de données sensibles.

Le RGPD ne semble pas indiquer la nécessité de s'en tenir à un fondement unique (la formulation utilisée à l'article 17(1)(b) semble indiquer le contraire). Il convient de souligner que le choix de plusieurs fondements juridiques, ou conditions de l'article 9, lorsqu'ils sont valables, pourrait permettre de renforcer la légitimité du traitement et l'information donnée aux personnes si l'un des traitements venait à être remis en cause.

Toutefois, certaines interprétations veulent que ne soit choisi qu'un seul fondement juridique par finalité (ou sous-finalité) du traitement.

2. Applicabilité des fondements juridiques de l'article 6 et conditions de l'article 9

Par ailleurs, la portée de chacun des fondements et conditions et leur champ d'application font eux aussi l'objet d'interprétations variées.

Ceci semble dû à la fois à des interprétations nationales différentes et à des divergences de traduction du Règlement dans les différents langages de l'Union.

Dans le domaine de la recherche en santé, notamment, les fondements et conditions touchant à l'intérêt public font l'objet de nombreux débats.

L'alinéa 1(e) de l'article 6 indique :

le traitement est nécessaire à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement;

La Commission européenne a publié au printemps 2019 un document de questions et réponses sur l'articulation entre les dispositions du RGPD et du Règlement (UE) n°536/2014 relatif aux essais cliniques de médicament.(25)

Préalablement à la publication de ce document, le CEPD avait publié son avis sur le projet de la Commission européenne.(26)

Ces deux documents indiquent que les essais cliniques peuvent reposer sur les combinaisons suivantes de fondement juridique de l'article 6 du RGPD et exceptions de l'article 9 :

- Fondement et exception liés au consentement de la personne concernée (art. 6(1)(a) et art. 9(2)(a)), tout en clarifiant que ce n'est pas parce que la méthodologie de la recherche nécessite de recueillir le consentement de la personne que le consentement doit nécessairement servir de fondement juridique au traitement des données ;
- ou fondement de mission menée dans l'intérêt public (art. 6(1)(e)) ou de l'intérêt légitime du responsable de traitement (art. 6(1)(f)) en combinaison avec l'exception relative aux raisons d'intérêt public dans la santé publique (art. 9(2)(i)) ou l'exception relative à la recherche scientifique (art. 9(2)(j)).

Le CEPD et la Commission européenne considèrent que ces fondements s'appliquent différemment selon le type de finalités poursuivi par les traitements menés dans le cadre de l'essai clinique ou à sa suite.

- Lorsque le traitement de données vise à protéger la santé en assurant des normes rigoureuses de fiabilité et sécurité pour les médicaments, le fondement 6(1)(c) est considéré comme applicable, en combinaison avec l'exception 9(2)i).
- Lorsque le traitement de données est purement mené dans un but de recherche, alors il est possible de se reposer sur
    - le fondement 6(1)(a) en combinaison avec l'exception 9(2)(a),
    - ou le fondement 6(1)(e) ou 6(1)(f) en combinaison avec l'exception 9(2)(i) ou 9(2)(j).

Si l'on prend l'exemple des essais cliniques portant sur le médicament, la Directive n° 2001/20/CE, actuellement en vigueur, et le Règlement (UE) n° 536/2014, qui entrera

en vigueur en 2020 et remplacera la Directive, tous deux relatifs aux essais cliniques de médicaments à usage humain, que ces recherches ne peuvent être mises en œuvre qu'au regard d'un bénéfice escompté pour les participants ou la santé publique (article 28).

Par conséquent, il semble que ces traitements, lorsque légitimement mis en œuvre, pourraient être éligibles au fondement 6(1)(e) et à l'exception 9(2)(i) du RGPD.

Toutefois, le CEPD et la Commission indiquent que le traitement peut se fonder sur l'intérêt public (6(1)(e)) lorsque celui-ci est défini par la loi européenne ou nationale et qu'il entre directement dans les missions dont est investi le responsable de traitement. Selon cette interprétation, le fondement de l'intérêt public ne serait donc applicable, pour un même type de traitement, qu'aux seuls organismes directement investis de telles missions. Par exemple, un organisme privé menant des essais cliniques sans que sa mission n'ait été définie en tant que telle par les pouvoirs publics ne saurait donc se prévaloir du fondement de l'intérêt public.

Dans la formulation de l'article 6(1)(e), on pourra s'interroger sur cette interprétation et sur la portée de "dont est investi le responsable de traitement" :

- s'applique-t-il uniquement à "l'exercice de l'autorité publique" dans "relevant de l'exercice de l'autorité publique dont est investi le responsable du traitement"
- ou à l'intégralité de l'expression "à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique" restreignant ainsi son champ d'application à non pas toutes les "missions d'intérêt public" mais uniquement aux "missions d'intérêt public dont est investi le responsable de traitement", sous-entendu, par une autorité publique.

La lecture de la version du RGPD en langue anglaise renforce ce questionnement :

« processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller; »

En traduction littérale et non officielle : "le traitement est nécessaire à l'exécution d'une tâche menée dans l'intérêt public ou dans l'exercice d'une autorité officielle dont est investi le responsable de traitement".

On pourrait alors penser que la "mission d'intérêt public" consiste en tout traitement ayant pour finalité l'intérêt public. Il est également moins clair dans la version anglophone que l'expression "dont est investi le responsable de traitement"

s'applique à la fois aux "missions" et à "l'autorité publique" du responsable de traitement.

L'autorité de contrôle britannique considère que ce fondement 6.1(e) est applicable lorsque le responsable de traitement :

- mène une tâche dans l'intérêt public, tel que défini par la loi,
- ou exerce une autorité officielle définie par la loi.(27)

Si elle peut paraitre bénigne au premier abord, l'interprétation du CEPD et de la Commission européenne viennent poser de réelles questions sur le fondement applicable à certains traitements, et surtout à l'applicabilité du droit d'opposition qui en découle. En effet, la réalité est loin d'être aussi schématique que le laissent entendre les documents de la Commission européenne et du CEPD, et les frontières entre les finalités de traitement sont poreuses.

Quel fondement viendra-t-on appliquer à l'utilisation des données pour un dépôt d'autorisation de mise sur le marché, ou à la transmission des données de sécurité d'un promoteur vers un exploitant du médicament qui ne serait pas impliqué directement dans l'essai mais pourrait tirer de ces données des informations de pharmacovigilance utiles ? Peut-on imaginer que soient alors utilisées des bases de données tronquées, dans lesquelles manqueraient des données lorsque des personnes auraient exercé leur droit d'opposition, ce qui est rendu possible lorsque le fondement n'est pas l'intérêt public ?

Pour les réutilisations ultérieures de bases de données existantes, la robustesse des résultats doit-elle dépendre des droits qu'auront pu exercer entretemps les personnes ?

En France, l'article 66.I de la LIL pose comme condition au traitement de données de santé qu'elles le soient dans l'intérêt public, et de rajouter, selon les termes de l'article 9(2)(i) du RGPD, que "[l]a garantie de normes élevées de qualité et de sécurité des soins de santé et des médicaments ou des dispositifs médicaux constitue une finalité d'intérêt public". Cette mention a été ajouté lors des débats parlementaires portant sur le projet de loi relative à la protection des données personnelles (qui fut ensuite publiée en tant que loi n°2018-493 du 20 juin 2018) avec l'objectif de rappeler que les organismes privés peuvent répondre à cet impératif d'intérêt public.(28) Toutefois, en prenant en compte les interprétations précédentes, bien que la loi pose cet impératif de répondre à l'intérêt public, il ne permet pas à tout traitement licite de

données de santé de reposer sur le fondement 6(1)(e) car ceci ne présupose pas une « mission d'intérêt public ».

B. Partage des données au sein de la communauté scientifique

1. Contexte

La réutilisation et le partage des données et des échantillons permettent à la fois

- de baisser les coûts de la recherche lorsque les données nécessaires ont déjà été collectées et peuvent être rendues disponibles,
- de valoriser et tirer tout le bénéfice du travail de collecte, de sécurisation et de gestion des données et échantillons collectés,
- de respecter la confiance accordée, l'effort de participation et de don d'échantillon fourni par les personnes concernées en tirant de ces données tout le parti qu'elles pourraient souhaiter pour la recherche en santé.

Le travail de recherche dans le domaine de la santé sur des bases de données collectées dans un autre cadre est souvent rendu difficile ou retardé en raison du manque de préparation des conditions qui vont permettre ce partage. C'est un élément essentiel de la gouvernance des données.

Ceci passe par

- des conditions de sécurité adéquates pour le partage (canaux de transfert sécurisés, travail dans une "bulle", un environnement informatique sécurisé, etc.),
- la définition claire des rôles des différents acteurs, formalisée par contrat,
- la définition d'un cadre réglementaire approprié,
- la délivrance d'une information appropriée aux personnes concernées,
- et dans l'idéal, le formatage des données et leur étiquetage selon des normes communes aux acteurs afin de faciliter leur appréhension par les chercheurs qui n'auraient pas participé à leur collecte initiale (travail de *data stewardship*).

2. Difficultés liées à l'encadrement réglementaire

En France, la Loi informatique et libertés s'ajoute en surcouche au RGPD. Elle définit les conditions précises de l'encadrement des traitements de données de santé.

La CNIL a su prendre des mesures facilitatrices pour les traitements jugés les plus courants, ce sont les Méthodologies de référence (MR).

Lorsqu'un traitement ne peut être conforme à une MR, le cadre réglementaire est toutefois assez restrictif. En effet, le responsable de traitement d'une recherche en santé n'a que deux options :

- l'autorisation du traitement par la CNIL ;
- ou l'autorisation d'un groupe de traitements similaires, réalisés par un même responsable de traitement, par le biais d'une décision unique de la CNIL.

Si cet encadrement peut paraitre suffisant pour une recherche ponctuelle menée par un responsable de traitement ou un même groupe de responsables de traitement conjoints, il peut être très limitant dans des situations de collaboration entre organismes.

En effet, il semble que ce cadre réglementaire ait été créé dans l'optique de recherches ponctuelles, faisant principalement l'objet d'une collecte de données directement auprès des personnes concernées, ou avec une exploitation ultérieure des données limitée. Ce cadre apparait rapidement limitant dans une perspective de réutilisation des données préalablement collectées, comme par exemple les recherches à partir d'une même base de données. En effet, dès que l'organisme réalisant la recherche (responsable de traitement) va changer, une nouvelle autorisation est à obtenir, y compris lorsque le nouvel opérateur adhère aux conditions de l'autorisation initiale et lorsque les recherches sont menées selon des modalités similaires et poursuivant des finalités proches

Il est alors difficile d'organiser des collaborations internationales et d'exploiter les bases de données selon leur plein potentiel, y compris lorsque l'information des personnes aura été délivrée de manière exhaustive et les modalités des collaborations ultérieures auront été définies à l'avance.

On regrettera, donc dans la réglementation française actuelle, l'absence de possibilité pour un responsable de traitement soumis à la loi française de réaliser une demande d'autorisation à la fois pour un large projet de constitution de base de données (soit centralisée en entrepôt, soit décentralisée et accessible via un outil unique) et pour les recherches ultérieures que pourrait permettre cette infrastructure, y compris lorsque des limites et garde-fous sont préalablement définis.

3. Difficultés liées à la transparence envers les personnes

Comme précédemment exposé, en l'absence d'une information conforme aux articles 13 ou 14 du RGPD, le responsable de traitement ne peut se conformer à une Méthodologie de référence et doit effectuer une demande d'avis et une demande

d'autorisation additionnelle de son projet de recherche. Cette difficulté est particulièrement exacerbée lorsque le responsable de traitement de données concernant des personnes situées en France est lui situé à l'étranger et a une connaissance imparfaite des spécificités de la loi française, bien qu'il applique les dispositions du RGPD.

Le considérant 33 du RGPD concède qu'il est parfois difficile de connaitre entièrement et précisément dès la collecte des données toutes les finalités de recherche scientifique qui pourront leur être appliquées. Il semble dès lors, qu'il puisse être développé un système selon lequel la personne pourrait consentir à des utilisations ultérieures, "pour ce qui est de certains domaines de la recherche ou de certaines parties de projets de recherche, dans la mesure où la finalité visée le permet".

Le considérant 50 nous éclaire sur le concept de finalité compatible, qui devrait pouvoir notamment s'appliquer à la recherche scientifique. Ce concept est repris par l'article 5(1)(b) du même Règlement. Le considérant 50 indique par ailleurs que la base juridique du traitement initial devrait alors pouvoir s'appliquer au traitement ultérieur.

Le RGPD prévoit, dans ses articles 13 et 14, une transparence envers les personnes concernées par le traitement de leurs données basée sur un modèle formel avec notamment l'identification précise du responsable de traitement, de son délégué à la protection des données, de la finalité du traitement, des mesures encadrant l'éventuel transfert hors UE des données, de la durée de conservation, de l'origine des données lorsqu'elles ne sont pas collectées auprès de la personne, etc.

Si ces informations sont complètes, le maintien de leur exactitude et exhaustivité en cas de recherche ultérieure ou de partage de données nécessite une fréquence de ré-information et un degré de détail qui ne parait pas toujours pertinent pour les personnes concernées.

Une étude menée aux États-Unis auprès de 771 participants à des essais cliniques s'est penchée sur l'opinion des participants aux essais cliniques sur le partage des données les concernant.(29) La majorité des personnes interrogées (82%) se sont déclarées en faveur du partage de leurs données en raison des bénéfices perçus pour la science, bien que certaines se soient déclarées inquiètes des risques liés au partage des données. Une proportion de participants a indiqué des inquiétudes relatives au fait que le partage de données puisse décourager les participants aux essais cliniques

(37%), que ces informations soient utilisées à des fins marketing (34%) et que les informations soient volées (31%). Parmi les principaux bénéfices perçus par les participants, ceux-ci ont indiqué l'assurance que la conduite de l'essai conduira au maximum de bénéfices scientifiques possible et l'accélération de la recherche.

Dans son rapport sur le thème du numérique en santé, le Comité consultatif national d'éthique (CCNE) est arrivé à une conclusion similaire et préconise une information des personnes portant sur les modalités du partage plutôt que sur les finalités précises de chacun des traitements ultérieurs(30) :

"[...] *De surcroît, dans le domaine de la recherche en santé, l'accès aux données issues du soin, des systèmes de santé, ou d'autres bases de données personnelles représente un progrès majeur en évitant un double recueil des données, lequel peut représenter un coût considérable pour les études interventionnelles (essais cliniques, essais d'intervention) ou observationnelles (grandes enquêtes transversales, cohortes). Il serait donc de bonne pratique, dans toute la mesure du possible, de recueillir un consentement qui permette aux personnes d'autoriser le partage de leurs données en sachant comment elles vont être partagées (plan de partage), plutôt que pourquoi (par qui et pour quelle recherche).*"

Cette information serait alors complétée par une information disponible sur le site internet de chacun des responsables de traitement, conformément à la recommandation du CEPD.(17)

Dans le contexte réglementaire actuel, la recherche ultérieure ou le partage des données peuvent être rendus difficiles ou retardés en raison de l'absence, dans l'information initialement délivrée aux personnes concernées, soit de certaines mentions prévues par le RGPD, soit de la référence à un support unique d'information dynamique (cf. III.1.b)(2)).

Pour cette raison, des projets de recherche, par ailleurs conformes à la réglementation et aux référentiels, tels que la MR-004, viennent nécessiter un avis du CEREES et une autorisation de la CNIL lorsque cette information n'a pas été délivrée.

Toutefois, la délivrance de ces informations se heurte à plusieurs obstacles une fois la collecte des données terminées. En effet, une certaine proportion des personnes est perdue de vue, par exemple suite à un déménagement, à l'arrêt du suivi, ou suite à un décès dont le responsable de traitement n'a pas connaissance. La délivrance de l'information peut également concerner un large nombre de patients et représenter un défi logistique conséquent.

Dans certains cas, notamment chez des patients touchés par des pathologies graves, la délivrance de l'information pourra être au mieux maladroite, créer un stress chez un patient atteint d'une pathologie grave en rémission ou en phase palliative, voire être délivrée à une personne non concernée et réveiller alors un souvenir douloureux ou encore révéler par ce biais une information de santé que n'aurait pas communiquée la personne concernée à ses proches.

4. Opportunités et prochains rendez-vous législatifs

Le cadre juridique et réglementaire n'est toutefois pas figé et plusieurs possibilités d'évolutions vont prochainement se présenter.

a) Ma Santé 2022

En septembre 2018, la Ministre des Solidarités et de la Santé, Agnès Buzyn, a annoncé le projet de loi dit "Ma Santé 2022" visant à "propose[r] une vision d'ensemble et des réponses globales aux défis auxquels est confronté le système de santé français".(31)

Ce projet de loi, dont le vote est prévu au mois de juin 2019, en procédure accélérée, comporte notamment plusieurs mesures portant la création d'une Plateforme des données de santé, qui remplacerait l'actuel Institut national des données de santé.

Cette loi ou ses décrets d'application pourraient intégrer des mesures visant à adapter le cadre réglementaire de la protection des données de santé à ses enjeux et ses contraintes.

b) Projet de révision de la loi bioéthique

En janvier 2018, le processus de révision de la loi bioéthique a été lancé avec l'ouverture par le Comité consultatif national d'éthique (CCNE) des Etats généraux de la bioéthique. Les données de santé et la génétique ont fait partie des thèmes débattus lors de ces Etats généraux.

Le projet de loi de loi sur la révision de la loi bioéthique devrait être présenté par le gouvernement avant l'été 2019.

A cette occasion, des définitions du droit national telles que de l'"examen des caractéristiques génétiques" pourraient être revu pour s'articuler plus aisément avec celles du RGPD.

A ce titre, les dispositions actuelles du Code de la santé publique, provenant des versions antérieures de la loi bioéthique, concentrent les garde-fous sur l'"examen des caractéristiques génétiques". En effet, actuellement, il est plus compliqué du

point de vue réglementaire d'analyser un échantillon pour connaitre le variant d'un gène exprimé par celui-ci que de faire une nouvelle recherche sur de larges portions de génome provenant d'un échantillon qui aurait été préalablement analysé. En effet, dans le premier cas, s'appliquent à la fois le cadre réglementaire de la protection des données et de la bioéthique, alors que dans le second, seule la protection des données est à considérer.

Alors que le prix des analyses génomiques a chuté ces dernières années et que ces analyses se banalisent dans la recherche, on pourra s'interroger sur le maintien de cette quasi-sacralisation de l'acte de l'analyse génomique. Il semble en effet que c'est par leur nature et les informations qu'elles contiennent que ces informations sont sensibles, et non pas en fonction de la manière dont elles sont obtenues.

c) Méthodologies de référence

Les Méthodologies de référence de la CNIL sont conçues comme des documents que l'autorité peut faire évoluer en fonction du besoin.

Dès lors, avec le retour d'expérience dont commence à disposer la CNIL, peut-être qu'une révision prochaine des MR sera envisagée, sur un élargissement de leur champ d'application, ou l'intégration de nouvelles possibilités qui commencent à se répandre en pratique, telle que le recueil de consentements électroniques dans les essais cliniques, ou des modalités d'information des personnes plus souples.

d) Code de conduite

Si le législateur et les autorités peuvent faire évoluer le cadre réglementaire, il est également possible pour les acteurs de la recherche, d'écrire et de demande l'autorisation par la CNIL ou par la Commission européenne d'un "code de conduite", en application de l'article 40 du RGPD.

Un tel code de conduite pourrait permettre de définir les modalités de délivrance d'une information claire, concise et pertinente ainsi que faciliter l’exercice des droits des patients.

C. Vers une véritable gouvernance des données

1. De nouvelles menaces

Les hôpitaux, en France, dans l'Union européenne et à l'international, font l'objet de cyber-attaques de plus en plus fréquentes. Celles-ci exposent notamment les établissements au risque, d'exposer des informations confidentielles sur leur état de santé, ou d'utiliser ces données pour usurper leur identité, ou tout simplement de perdre les informations essentielles au suivi de leurs patients ou à l'archivage légal de leurs données.(32–37)

Du côté des organismes de recherches (qui peuvent être ces mêmes établissements de soin), s'ils pourraient réparer les dommages suite à un cambriolage matériel, qu'en est-il de leurs données ? Quel organisme de recherche saura se remettre parfaitement de l'effacement de données, fruit de plusieurs années de recherche et trésor potentiel à valoriser pour des recherches ultérieures ?

2. De nouvelles opportunités

En parallèle, la numérisation du soin et de la recherche, le développement de nouvelles technologies, algorithmes, télémédecine reposant sur les données de santé nécessitent une organisation de l'infrastructure des systèmes d'information.

L'opposition entre donnée de soin et donnée de recherche qui a pu prévaloir par le passé tend à se dissiper. Il devient de plus en plus clair que la donnée issue du soin va pouvoir constituer une base ou un complément pour la recherche.

La possibilité de dupliquer et interconnecter les données ouvre par ailleurs de nombreuses perspectives de collaboration entre chercheurs. En France, le projet de "*Health Data Hub*" ou "Plateforme des données de santé", conduit par le Ministère des Solidarités et de la Santé, vise par exemple à structurer et garantir l'interfaçage des données du soin au niveau national afin de permettre la recherche à l'échelle nationale.(38)

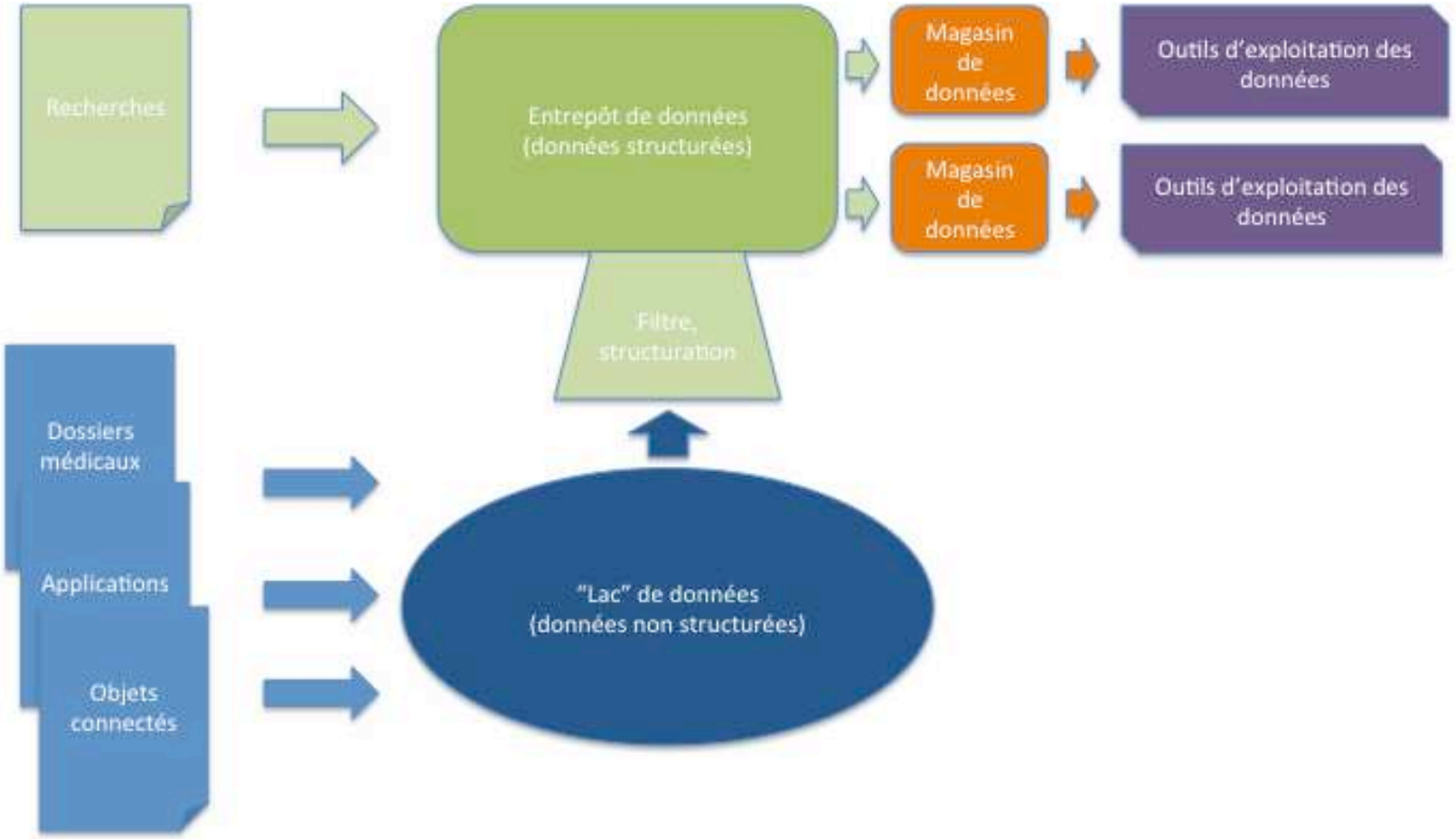

Figure 5 : Exemple d'organisation de l'infrastructure des données

3. Une convergence de la démarche

Ainsi, la protection des données personnelles requise par le RGPD contribue directement à protéger les données des menaces et organiser l'activité afin d'ouvrir de nouvelles opportunités. Garantir la protection des données nécessite pour le responsable de traitement de développer une vraie politique interne de protection, traçabilité et anticipation des risques relatifs aux données personnelles.

Toutefois le respect de la réglementation ne suffit pas lorsqu'il s'agit de préparer l'avenir de la recherche. C'est une composante à intégrer dans une démarche globale et convergente de gouvernance de la donnée.

Cette gouvernance vise à organiser l'infrastructure de traitement de la donnée, afin d'en définir le format, les modalités de collecte, les circuits, les accès, les utilisations, les outils, la valorisation, la conservation, etc. En bref, c'est une réflexion globale sur l'organisation de ce qu'il convient de considérer comme une ressource à la fois sensible et précieuse.

Les organismes bénéficient de l'émergence de nouvelles techniques pour la protection et l'exploitation des données. L'organisation de la gouvernance des données permettra d'adopter rapidement des outils d'intérêt tout en limitant le gâchis financier ou humain qui peut parfois être déploré par l'adoption de techniques inadéquates.

Parmi ces thématiques, citons par exemple la réflexion actuelle, non sans difficultés, autour du *blockchain* pour protéger et tracer les données(39), ou des exemples d'applications pratiques tels que la plateforme Substra qui vise à favoriser le partage et l'exploitation des données de santé(40).

V. Conclusion

Si le secteur de la recherche en santé était auparavant déjà très réglementé, le RGPD est venu placer la protection des données personnelles sur le devant de la scène médiatique et dans la conscience des acteurs et personnes concernées.

La réglementation relative à la protection des données personnelles vient renforcer pour les organismes de recherche les impératifs de documentation interne et de transparence envers les personnes concernées.

De ces obligations réglementaires découle la nécessité d'organiser une réelle gouvernance des données qui vise à organiser les traitements de données dans leur ensemble, leur infrastructure, traçabilité et formalisme. Toutefois, ces démarches ne sont pas réalisées en vain ou pour le seul respect de la réglementation. Elles sont l'occasion pour les organismes de se prémunir contre les risques liés aux fuites ou pertes de données et d'enclencher dès aujourd'hui la transition numérique du soin et de la recherche.

Du point de vue réglementaire, certains progrès restent à faire, afin de clarifier la mise en pratique de certaines dispositions du RGPD ou leur harmonisation au sein de l'Union européenne et l'interface avec les pays tiers.

Dans ces démarches, les professionnels de santé sont à l'interface, avec pour rôle d'utilisateur des nouvelles technologies, de curateurs dans la sélection des nouveaux outils, de participants actifs à leur développement, mais également de point de contact des patients et participants aux recherches.

La santé numérique et les algorithmes vont modifier profondément la pratique du soin et de la recherche. Dans cette évolution la compréhension des enjeux liés à la protection des données personnelles est essentielle pour accompagner et rationaliser ce changement.

Bibliographie

1. Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

2. BRAIBANT G. Données personnelles et société de l'information : rapport au Premier ministre sur la transposition en droit français de la directive numéro 95-46 [Internet]. La Documentation française. [cité 23 avr 2016]. Disponible sur: http://www.ladocumentationfrancaise.fr/rapports-publics/984000836/index.shtml

3. Ordonnance n° 2018-1125 du 12 décembre 2018 prise en application de l'article 32 de la loi n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles et portant modification de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés et diverses dispositions concernant la protection des données à caractère personnel.

4. Groupe de travail Article 29 sur la protection des données. Avis 05/2014 sur les Techniques d'anonymisation [Internet]. Commission européenne; 2014. Disponible sur: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_fr.pdf

5. Groupe de travail Article 29 sur la protection des données. Lignes directrices sur le consentement au sens du règlement 2016/679 [Internet]. Commission européenne; 2018. Disponible sur: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051

6. Legitimate interests [Internet]. Information Commissioner's Office (UK). 2018 [cité 13 oct 2018]. Disponible sur: https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/legitimate-interests/

7. Hern A. New York taxi details can be extracted from anonymised data, researchers say. The Guardian. 27 juin 2014;Technology.

8. Ce qu'il faut savoir sur l'analyse d'impact relative à la protection des données (AIPD) | CNIL [Internet]. [cité 2 déc 2018]. Disponible sur: https://www.cnil.fr/fr/ce-quil-faut-savoir-sur-lanalyse-dimpact-relative-la-protection-des-donnees-aipd

9. Groupe de travail Article 29 sur la protection des données. Lignes directrices concernant l'analyse d'impact relative à la protection des données [Internet]. Commission européenne; 2017. Disponible sur: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236

10. Loi n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé. 2016-41 janv 26, 2016.

11. Loi n° 2012-300 du 5 mars 2012 relative aux recherches impliquant la personne humaine. 2012-300 mars 5, 2012.

12. Ministère des Affaires sociales et de la santé. Recherches impliquant la personne humaine, Questions-réponses [Internet]. 2016 [cité 16 déc 2018]. Disponible sur: https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/recherche-et-innovation/recherches-impliquant-la-personne-humaine/

13. Décret n° 2016-1871 du 26 décembre 2016 relatif au traitement de données à caractère personnel dénommé « système national des données de santé ». 2016-1871 décembre, 2016.

14. Délibération n° 2018-134 du 12 avril 2018 portant homologation de conditions de mise à disposition de l'échantillon généraliste des bénéficiaires (EGB) et des bases de données thématiques appelées « datamarts » du Système National d'Information Inter Régimes de l'Assurance Maladie (SNIIRAM). 2018-134 avril, 2018.

15. Délibération n° 2018-256 du 7 juin 2018 portant homologation d'une méthodologie de référence relative aux traitements de données nécessitant l'accès par des établissements de santé et des fédérations aux données du PMSI et des résumés de passage aux urgences (RPU) centralisées et mises à disposition sur la plateforme sécurisée de l'ATIH (MR 005).

16. Délibération n° 2018-257 du 7 juin 2018 portant homologation d'une méthodologie de référence relative aux traitements de données nécessitant l'accès pour le compte des personnes produisant ou commercialisant des produits mentionnés au II de l'article L. 5311-1 du code de la santé publique aux données du PMSI centralisées et mises à disposition par l'ATIH par l'intermédiaire d'une solution sécurisée (MR 006).

17. Groupe de travail Article 29 sur la protection des données. Lignes directrices sur la transparence au sens du règlement (UE) 2016/679 [Internet]. Commission européenne; 2018. Disponible sur: http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227

18. Agencia Española de Protección de Datos. Informe acerca de la incidencia que en el ámbito de la investigación biomédica pudiera producir la plena aplicación a partir del 25 de mayo de 2018 del Reglamento General de Protección de Datos [Internet]. 2018 [cité 4 nov 2018]. Disponible sur: https://www.aepd.es/media/informes/2018-0046-investigacion-biomedica.pdf

19. Commission Implementing Decision (EU) 2016/1250 of 12 July 2016 pursuant to Directive 95/46/EC of the European Parliament and of the Council on the adequacy of the protection provided by the EU-U.S. Privacy Shield (notified under document C(2016) 4176) (Text with EEA relevance) [Internet]. OJ L, 32016D1250 août 1, 2016. Disponible sur: http://data.europa.eu/eli/dec_impl/2016/1250/oj/eng

20. Groupe de travail Article 29 sur la protection des données. Avis 01/2016 sur le projet de décision concernant le caractère adéquat du bouclier de protection des données UE-Etats-Unis [Internet]. Commission européenne; 2017. Disponible sur: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236

21. Document de travail relatif à une interprétation commune des dispositions de l'article 26, paragraphe 1, de la directive 95/46/CE du 24 octobre 1995 [Internet]. Commission européenne; 2005 [cité 29 mai 2016]. Disponible sur: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2005/wp114_fr.pdf

22. Model contracts for the transfer of personal data to third countries [Internet]. European Commission - European Commission. [cité 15 déc 2018]. Disponible sur: https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/model-contracts-transfer-personal-data-third-countries_en

23. Irish High Court. Data Protection Commissioner v Facebook Ireland Limited [2016 No. 4809 P.]. mars 10, 2017.

24. Molnár-Gábor F. Germany: a fair balance between scientific freedom and data subjects' rights? Hum Genet. 2018;137(8):619□26.

25. European Commission, Directorate General for Health and Food Safety. Question and Answers on the interplay between the Clinical Trials Regulation and the General Data Protection Regulation [Internet]. 2019 [cité 14 avr 2019]. Disponible sur: https://ec.europa.eu/health/sites/health/files/files/documents/qa_clinicaltrials_gdpr_en.pdf

26. European Data Protection Board. Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) [Internet]. 2019 [cité 30 mars 2019]. Disponible sur: https://edpb.europa.eu/our-work-tools/our-documents/opinion-art-70/opinion-32019-concerning-questions-and-answers-interplay_en

27. Information Commissioner's Office (ICO). Public task [Internet]. ICO website. 2018 [cité 16 déc 2018]. Disponible sur: https://icoumbraco.azurewebsites.net/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/public-task/

28. Assemblée nationale. Première séance du mercredi 07 février 2018, Session ordinaire, Compte rendu intégral [Internet]. 2018 [cité 13 mai 2019]. Disponible sur: http://www.assemblee-nationale.fr/15/cri/2017-2018/20180136.asp#P1182191

29. Mello MM, Lieou V, Goodman SN. Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing. N Engl J Med. 7 juin 2018;378(23):2202□11.

30. Comité consultatif national d'éthique pour les sciences de la vie et de la santé (CCNE). Numérique & santé : quels enjeux éthiques pour quelles régulations ? [Internet]. 2018. Disponible sur: https://www.ccne-ethique.fr/fr/publications/numerique-sante-quels-enjeux-ethiques-pour-quelles-regulations

31. DICOM_Lisa.C, DICOM_Lisa.C. Dossier de presse : Ma santé 2022 : un engagement collectif [Internet]. Ministère des Solidarités et de la Santé. 2019 [cité 15 mai 2019]. Disponible sur: https://solidarites-sante.gouv.fr/actualites/presse/dossiers-de-presse/article/dossier-de-presse-ma-sante-2022-un-engagement-collectif

32. Reynaud F. Dans les hôpitaux français, « le vrai, gros piratage n'a pas encore eu lieu ». Le Monde.fr [Internet]. 9 juill 2017 [cité 22 sept 2018]; Disponible sur: https://www.lemonde.fr/pixels/article/2017/07/09/dans-les-hopitaux-francais-le-vrai-piratage-le-gros-il-n-a-pas-encore-eu-lieu_5158152_4408996.html

33. Cyberattaques : les hôpitaux britanniques, principales cibles atteintes [Internet]. lesechos.fr. [cité 16 déc 2018]. Disponible sur: https://www.lesechos.fr/13/05/2017/lesechos.fr/0212077825043_cyberattaques---les-hopitaux-britanniques--principales-cibles-atteintes.htm

34. Edwards E. Tullamore hospital suffers ransomware attack [Internet]. The Irish Times. [cité 16 déc 2018]. Disponible sur: https://www.irishtimes.com/news/ireland/irish-news/tullamore-hospital-suffers-ransomware-attack-1.3699397

35. The biggest healthcare data breaches of 2018 (so far) [Internet]. Healthcare IT News. 2018 [cité 16 déc 2018]. Disponible sur: https://www.healthcareitnews.com/projects/biggest-healthcare-data-breaches-2018-so-far

36. Jalali MS. Defending hospitals against life-threatening cyberattacks [Internet]. The Conversation. [cité 16 déc 2018]. Disponible sur: http://theconversation.com/defending-hospitals-against-life-threatening-cyberattacks-93052

37. Une attaque informatique paralyse 649 ordinateurs du CHU de Montpellier [Internet]. egora.fr. 2019 [cité 26 mai 2019]. Disponible sur: https://www.egora.fr/actus-pro/hopitaux-cliniques/47745-une-attaque-informatique-paralyse-649-ordinateurs-du-chu-de

38. DICOM_Lisa.C. Rapport Health Data Hub, mission de préfiguration [Internet]. Ministère des Solidarités et de la Santé. 2018 [cité 16 déc 2018]. Disponible sur: https://solidarites-sante.gouv.fr/ministere/documentation-et-publications-officielles/rapports/sante/article/rapport-health-data-hub-mission-de-prefiguration

39. Blockchain et RGPD : quelles solutions pour un usage responsable en présence de données personnelles ? | CNIL [Internet]. [cité 16 déc 2018]. Disponible sur: https://www.cnil.fr/fr/blockchain-et-rgpd-quelles-solutions-pour-un-usage-responsable-en-presence-de-donnees-personnelles

40. OWKIN lance le projet collaboratif Substra pour libérer le plein potentiel de l'IA dans la santé en assurant la protection des données [Internet]. [cité 16 déc 2018]. Disponible sur: /contenu/owkin-lance-le-projet-collaboratif-substra-pour-liberer-le-plein-potentiel-de-lia-dans-la

41. Data for Good [Internet]. Data for Good. [cité 2 déc 2018]. Disponible sur: https://dataforgood.fr/

42. Serment d'Hippocrate pour data scientist [Internet]. [cité 2 déc 2018]. Disponible sur: https://www.hippocrate.tech

Serment de Galien

Je jure, en présence des maîtres de la faculté et de mes condisciples :

D'honorer ceux qui m'ont instruit dans les préceptes de mon art et de leur témoigner ma reconnaissance en restant fidèle à leur enseignement.

D'exercer, dans l'intérêt de la santé publique, ma profession avec conscience et de respecter non seulement la législation en vigueur, mais aussi les règles de l'honneur, de la probité et du désintéressement.

De ne jamais oublier ma responsabilité, mes devoirs envers le malade et sa dignité humaine, de respecter le secret professionnel.

En aucun cas, je ne consentirai à utiliser mes connaissances et mon état pour corrompre les mœurs et favoriser des actes criminels.

Que les Hommes m'accordent leur estime si je suis fidèle à mes promesses.

Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.

Serment d'Hippocrate pour *Data Scientist*

Le 26 juin 2018, la communauté Data for Good(41) a proposé un Serment d'Hippocrate pour Data Scientist ou pour toute autre personne travaillant avec la donnée.(42)

Celui-ci apparaissant comme un complément pertinent au Serment de Galien au vu du sujet de cette thèse, je me permets donc de l'insérer ici.

En tant que professionnel(le) amené(e) à

collecter, stocker, traiter, modéliser, analyser des données et/ou à concevoir des algorithmes, des produits informatiques ou des interfaces,

je suis conscient(e) de l'impact que peut avoir mon travail sur des individus et sur la société dans son ensemble.

C'est pourquoi je m'engage à respecter les 5 principes suivants :

1. Intégrité scientifique et rigueur

J'exploiterai les données avec toute la rigueur requise et en conformité avec les meilleurs standards de ma profession.

2. Transparence

J'informerai de façon compréhensible et précise toutes les parties prenantes sur les finalités, les modalités et les implications potentielles de mon utilisation des données.

3. Equité

Je veillerai à toujours m'assurer que des individus ou des groupes ne soient pas discriminés par rapport à des critères illégaux ou illégitimes, de façon directe ou indirecte, sur la base de mes travaux sur les données.

4. Respect

J'exercerai mon activité professionnelle en respectant la vie privée et la dignité des personnes dans toutes leurs dimensions.

5. Responsabilité et indépendance

J'assumerai mes responsabilités en cas manquements ou de conflits d'intérêt et je donnerai l'alerte si des actes illégaux liés à des données sont constatés.

Résumé et mots clés

Résumé

Le RGPD, et plus généralement l'encadrement réglementaire relatif à la protection des données personnelles, a renforcé la prise de conscience de ces enjeux dans la recherche en santé. Toutefois, c'est un texte difficile à appréhender sans une bonne connaissance des notions qu'il aborde.

Cette thèse vise à clarifier les principales notions et principes de la réglementation, leur application à la recherche en santé, et fait état des difficultés qui restent à surmonter par les acteurs de la recherche pour le plein essor de l'exploitation sécurisée des données dans le domaine de la santé.

Mots clés

Santé, Recherche, Données à caractère personnel, Données personnelles, Données de santé, Gouvernance, RGPD, GDPR, Informatique et libertés, Protection, Vie privée, Transparence, Consentement, Droits, Partage, Réglementaire, Union européenne.