INTRODUCTION À LA PROTECTION DES DONNÉES PERSONNELLES DE SANTÉ

septembre 2021

Dr Julien Grosjean, PhD

Département d'Informatique et d'Information Médicales, CHU de Rouen & LIMICS
U1142 INSERM, Sorbonne Université & Université Sorbonne Paris Nord

Objectifs pédagogiques

- Assimiler et comprendre les principaux axes du RGPD et de la réglementation française en matière de protection et réutilisation des données produites dans le cadre du soin et de la recherche médicale
- Connaitre le vocabulaire dédié
- Savoir quelles données sont concernées et comment un établissement doit répondre aux exigences du RGPD
- Connaitre les droits des patients et comment les exercer
- Assimiler les principales recommandations/exigences en matière de protection des données personnelles en santé
- Avoir des notions en terme de réglementation vis-à-vis des Entrepôts de Données de Santé et de leurs spécificités au regard du RGPD

Quelles données ?

- Toutes les données personnelles sont dites « protégées » ou du moins doivent l'être
- On parlera ici uniquement des données informatisées (numériques) même si les données « papier » sont également protégées

- En santé, il peut s'agir de n'importe quelle donnée comme le poids, la taille, les pathologies, les symptômes, les interventions, les codages, les résultats de biologie ou encore les séquences biologiques comme le génome…

Quelles lois et cadres juridiques ?

- En France, deux niveaux se superposent :
  - En général d'abord, puisqu'au niveau Européen : le RGPD (Règlement Général sur la Protection des Données) depuis le 27 avril 2016
  - Spécifiquement, pour tout ce qui incombe à la France : la Loi Informatique et Libertés (LIL3) depuis le 6 janvier 1978 puis consolidée depuis le 21 juin 2018

- En France, le garant de l'application des lois et le cadré régulateur est la CNIL (Commission Nationale de l'Informatique et des Libertés), organisme indépendant. « *Elle a un rôle d'alerte, de conseil et d'information vers tous les publics mais dispose également d'un pouvoir de contrôle et de sanction.* »

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

Pourquoi protéger les données personnelles de santé ?

- Personne n'a envie de voir ses données divulguées sur le web surtout si on peut savoir de qui il s'agit…
- Beaucoup d'acteurs industriels rêveraient d'y avoir accès pour améliorer leur business : assurances, mutuelles, banques, sites de vente en ligne, etc.

- Toute information « sensible » est monnayable : vente sur le « dark net », etc. d'où l'intérêt grandissant des hackeurs
  - Exemples de plus en plus fréquents : CHU de Rouen en novembre 2019, CH de Dax, de Villefranche et une trentaine de laboratoires de biologie médicale en 2021

Vocabulaire

- Données à caractère personnel (= données personnelles)
- Traitement = tout opération informatique faite sur ces données (stockage, recherche, analyse, export…)
- Responsable de traitement = personne morale ou physique responsable en cas de contrôle ou d'infraction

- RIPH = Recherche Impliquant la Personne Humaine
- RNIPH = Recherche N'Impliquant pas la Personne Humaine

Points importants

- Base juridique de traitement : 6 bases existent mais 2 sont vraiment utilisées :
  - la 5 : Exécution d'une **mission d'intérêt public** (pour les organismes publics)
  - la 6 : Intérêts légitimes du Responsable du Traitement ou d'un tiers (pour les organismes privés)

- Consentement vs. Non-Opposition

consent ≠ ne s'oppose pas

Les Méthodologies de Référence (MR)

MR-001	**Recherches Interventionnelles = RIPH 1 & 2 => consentement**
MR-002	Études non interventionnelles de performances en matière de dispositifs médicaux de diagnostic in vitro
MR-003	**Recherches Non Interventionnelles = RIPH 3 => non-opposition**
MR-004	**RNIPH : la plupart du temps, études rétrospectives**
MR-005	PMSI + RPU (SNDS) par des acteurs publics (non commerciaux)
MR-006	PMSI + RPU (SNDS) par des acteurs privés (commerciaux)

MR-001

- « Traitement de données à caractère personnel présentant un caractère d'intérêt public »
- Recherches pour lesquelles l'inclusion d'une personne requiert le recueil de son consentement

➢ RIPH1 : Recherches « à risque » : évaluation de l'efficacité d'un traitement ou d'une procédure par exemples

➢ RIPH2 : Recherches « à faible risque » : évaluation de l'efficacité d'une thérapeutique « douce » comme l'exercice physique ou la diététique par exemples

MR-003

- « Traitement de données à caractère personnel présentant un caractère d'intérêt public »
- Recherches pour lesquelles l'inclusion d'une personne requiert sa non-opposition

➢ RIPH3 : Recherches « sans risque » : peu ou pas interventionnelles ; les traitements sont administrés et les actes pratiqués dans le cadre habituel du soin

MR-004

- Recherches sur données déjà collectées (études rétrospectives)
- Recherches sur des données « non sensibles » (enquêtes de satisfaction, évaluation de la pratique, etc.)
- Tests de produits cosmétiques/alimentaires
- Expérimentations en sciences humaines et sociales

Niveaux de confidentialité des données

- Identifiantes
  - directement : Noms, numéros (téléphone, INSEE...), adresses...
  - indirectement : dates de naissance/décès ou tout élément à croiser (lieu, évènement, date)
- Pseudonymisées
  - Pas de noms ou d'identifiants « publics » : seulement des identifiants spécifiques à l'étude et éventuellement des données indirectement identifiantes...
- Anonymisées => **la loi de s'applique pas**
  - « Impossible » de retrouver un patient avec les données exposées : cadre réglementaire strict défini
- Agrégées
  - statistiques diverses : moyennes, nombres, graphiques, etc. ne comportant pas définition aucun donnée à caractère personnel

Le délicat problème de l'anonymisation

- Les 3 critères du G29 n° 05/2014 :
  - l'individualisation : est-il toujours possible d'isoler un individu ?
  - la corrélation : est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?
  - l'inférence : peut-on déduire de l'information sur un individu ?

- **Ainsi** :
  - un ensemble de données pour lequel il n'est possible ni d'individualiser ni de corréler ni d'inférer est a priori anonyme ;
  - un ensemble de données pour lequel au moins un des trois critères n'est pas respecté ne pourra être considéré comme anonyme qu'à la suite d'une analyse détaillée des risques de ré-identification.

https://www.cnil.fr/fr/le-g29-publie-un-avis-sur-les-techniques-danonymisation

Aspects organisationnels, conformités au RGPD

Responsables des traitements

- Responsable de Traitement (RT) : personne physique ou morale responsable de la recherche (=promoteur)
- Responsable scientifique : personne désignée par le RT pour conduire la recherche et s'assurer la bonne conformité aux réglementations
- Professionnel intervenant dans la recherche : collecte, surveille, etc.

- Sous-traitant : personne physique ou morale appartenant à un autre organisme que celui du RT et agissant sous-contrat ou convention en son nom

- Responsable Conjoint de Traitement : responsabilités partagées mais obligations et rôles définis à l'avance par contrat

Délégué(e) à la Protection des Données

- DPD ou DPO (*Data Privacy Officer*) : la personne clé d'une structure en terme de RGPD
- Missions :
  - encadrer tous les types de traitements et en s'assurer la conformité (suivi, documentation...)
  - informer et conseiller les RT
  - coopérer/communiquer avec l'autorité de contrôle (CNIL)

Documentations

- De façon générale, il est recommandé de documenter au maximum les procédures et les éléments connexes ; ces documents feront foi en cas de contrôle mais sont aussi un bon moyen de suivre l'évolution des traitements et de leurs conséquences

- Registre des traitements : consigner tous les types de traitements et tous les traitements (dates, responsables, spécificités, etc.)

Information/consentement patients

- En fonction des MR, il est tantôt indispensable de :
  - recueillir le consentement « éclairé » d'un patient pour participer à une étude
  - soit d'informer « individuellement » un patient de l'utilisation de ses données pour une étude

- Sans MR, il est nécessaire de justifier le choix de ne pas recueillir le consentement (et donc de ne faire que l'information individuelle)
- Dans de très rares cas, une information collective pourra être mise en place (cf. EDS)

Droits des patients

- Accès : copie de toutes les informations détenues
- *Effacement [peu applicable en l'état en matière de santé sauf études spécifiques]**
- *Opposition [applicable pour la réutilisation en recherche notamment mais sans objet dans le cadre du soin]**
- Portabilité : lié à l'accès, pour transférer ces données vers un autre établissement
- *Rectification [peu applicable en l'état en matière de santé]**
- Limitation du traitement *[peu applicable en l'état y compris en recherche]**

***La base juridique de traitement va avoir une influence directe sur l'applicabilité de ces droits : en cas « d'Exécution d'une mission d'intérêt public », le promoteur peut s'y soustraire**

https://www.cnil.fr/fr/preparer-lexercice-des-droits-des-personnes

Exercice des droits des patients

- Que ce soit via consentement ou information, le patient DOIT être sensibilisé à l'exercice de ses droits concernant ses données (pourquoi, comment)
- La personne référente dans un établissement est toujours, par défaut, le ou la DPD. En fonction des études, cela peut être aussi les acteurs de l'étude qui sont directement en contact avec le patient (médecins...). Ceux-ci devront faire remonter la volonté du patient au (à la) DPD

https://www.cnil.fr/fr/preparer-lexercice-des-droits-des-personnes

Exigences/recommandations générales

- Toujours partir du principe que la donnée peut être « mal utilisée »
- Attention aux outils « faciles » : **la plupart des boites mail, serveurs de partage de documents, etc. ne sont pas compatibles avec les exigences de sécurité du RGPD**
- Mettre en œuvre un maximum de moyens de **façon appropriée** afin d'assurer la sécurité de ces données
- **Documenter** les processus afin de bien suivre leur évolution et de faciliter les contrôles

- **Minimiser les données** : ne collecter que les informations strictement nécessaires à l'étude ; sécurité mais aussi gain en terme méthodologique...
- **Pseudonymiser voire anonymiser les données au maximum**
- **Ne jamais transmettre les données** sauf ce qui est prévu dans le cadre d'une étude ; dans ce cas, **chiffrer les données**

- S'assurer que le patient est **informé et éclairé** et est en mesure d'exercer ses droits
- **Être transparent** sur son activité y compris auprès des autres professionnels...
- En cas de doute, **s'informer ou contacter d'autres professionnels** susceptibles d'avoir une réponse, y compris la CNIL elle-même

Le cas particulier des Entrepôts de Données de Santé

Encadrer les Entrepôts de Données de Santé

- Les Entrepôts de Données de Santé (EDS) sont des outils informatiques permettant la collection, l'intégration puis le traitement des données de santé provenant d'un grand nombre de sources d'information clinique (dossier patient informatisé, système d'information des laboratoires et d'imagerie, prescription informatisée, dossier infirmier…)

- Outils récents et **méthodologie inverse** à celles encadrées habituellement
- En France, il faut une autorisation spécifique délivrée par la CNIL : un dossier complet doit être déposé puis évalué
- Pour uniformiser les pratiques et la méthodologie des EDS, la CNIL est en train de formaliser un « référentiel » à la manière des MR

Objectifs des EDS

- recherche, étude ou évaluation dans le domaine de la santé ;
- production d’indicateurs et pilotage stratégique de l’activité du ou des établissements ou centre où s’exercent des activités de prévention, de diagnostic et de soins concernés

Gouvernance des EDS

- Comité de pilotage : orientations stratégiques et scientifiques (chef(fe) d'établissement, direction de la recherche, etc.)
- Comité scientifique et éthique : rend un avis préalable et motivé sur les propositions de projets exploitant l'EDS

- De façon optionnelle : comité opérationnel qui s'occupe de la mise en œuvre et de la maintenance de l'EDS

Quelles données dans les EDS ?

- Données à caractère personnel « adéquates, pertinentes et limitées » à ce qui est nécessaire au regard des finalités du traitements : issues du dossier médical et/ou de projets de recherche précédemment réalisés (durée de conservation non expiré)

- Plus précisément : poids, taille, comptes-rendus médicaux, résultats d'examens/d'analyses, imagerie, antécédents, maladies, actes et données médico-administratives (PMSI)
- En fonction des EDS, on peut aussi ajouter : photographies/vidéos, enregistrements audio, données génétiques, données relatives à la vie professionnelle, habitudes de vie, mode de vie, vie sexuelle...

- **Toutes ces données doivent être pseudonymisées** (voire anonymisées)

Durée de conservation / portée

- En fonction des cas :
  - 20 ans maximum **à compter de la collecte** (« fenêtre glissante »)
  - 20 ans maximum pour le dossier complet : si le patient ne revient pas 20 ans après sa dernière venue, les données sont supprimées, sinon le dossier est toujours « actif »

- Ces critères posent au moins deux problèmes :
  - (non) exhaustivité d'un dossier (données manquantes potentielles et impossibilité de savoir si c'est le cas)
  - cas particulier de la pédiatrie/néonat : suivi à la long terme

Qui peut accéder aux EDS ?

- Les professionnels de l'établissement concerné (médecins, pharmaciens, chercheurs, administratifs, étudiants)
- En fonction du niveau de confidentialité des données exploitées, certains professionnels n'auront pas les même droits : par exemple, un personnel administratif n'aura, la plupart du temps, accès seulement aux données agrégées ; ceci fait partie des tâches du comité scientifique et éthique

Procédure d’accès à un EDS

- Requiert la soumission d’un dossier au CSE de l’établissement : avis éthique (est-ce légitime de faire cette recherche ?), scientifique (méthodologie, données à collecter ?) et technique (est-ce faisable et quels sont les biais potentiels ?)
- Une étude de faisabilité est faite parallèlement
- En cas de feu vert du CSE : constitution/mise à disposition du jeu de données
- Le cas échéant, export(s)

- Tout accès est nominatif et tracé (journal)

Information aux patients

- Trois niveaux :
  - Après la constitution initiale de l'EDS : information individuelle aux patients lors de leur venue dans l'établissement (divers moyens possibles : paragraphe dédié sur un courrier de rdv, compte-rendu, SMS de rappel de rdv, liasse de soins, etc.)
  - Pendant la constitution initiale de l'EDS : la plupart du temps, le nombre de patients à contacter rétrospectivement est trop important (exemple : près de 2 millions de patients au CHU de Rouen entre 2000 et 2020) ; dans ce cas, une dérogation motivée peut être accordée par la CNIL. Dans tous les cas, (au moins) une information publique est exigée afin de prévenir la population « autour de l'établissement » (communication « grand public »)
  - Un affichage dédié dans chaque service (salles d'attentes par exemple) doit être mis en place pour prévenir que la réutilisation des données de soins est possible dans le cadre d'études exploitant un EDS

Information aux professionnels

- En fonction des EDS, certaines informations concernant les professionnels de santé de l'établissement peuvent être collectées
- Ceux-ci doivent être informées de quelles données exactes il s'agit

- Une note individuelle est remise à chaque professionnel
- Une communication interne plus large est en outre recommandée (magazine, intranet...)

EDS et éthique professionnelle

- L'exploitation des pose un problème plus complexe qui est la « priorité » des données issues des soins :
  - les chefs de services voire les médecins doivent-ils donner leur accord ou simplement être informés que les données qu'ils produisent sont exploités par d'autres professionnels de l'établissement ?
  - les publications scientifiques et notamment « l'authorship » sont également un enjeu épineux

- Au niveau légal, rien n'encadre ni définit ces problématiques ; le CSE doit pouvoir s'assurer de bon déroulé des travaux entre collègues mais la réalité est complexe

Spécificités en terme de sécurité

- Par définition, le but d'un EDS est de collecter TOUTES les informations (ou du moins un très grand nombre)
- Les mesures de sécurité doivent être adaptées (liste non exhaustive) :
  - données pseudonymisées via tables de correspondances
  - données non structurées anonymisées (« floutées »)
  - disques/partitions chiffrées
  - double authentification sur les postes informatiques
  - mots de passe forts
  - personnels habilités identifiés (administrateurs notamment)
  - locaux contrôlés (serveurs, postes d'accès)
  - accès/requêtes tracés
  - ...

Recueil de données et exports

- Seules les données strictement nécessaires à l'étude sont collectées... mais en pratique pas toujours simple. Exemple : les données non structurées (comptes-rendus, etc.) qui contiennent souvent plus d'informations et qui sont difficiles à isoler

- En pratique, deux options sont possibles :
  - créer un sous-jeu de données limité à l'étude (N dossiers de patients, voire filtrer les types de données) (=*datamart*)
  - exporter des données anonymisées/pseudonymisées/identifiantes selon les cas. Exemples d'identification obligatoire : prévention, pharmacovigilance