

Les Cahiers d'Orphanet

Comptes-rendus



Colloque du 26 septembre 2016

DONNÉES DE SANTÉ DANS UN MONDE HYPERCONNECTÉ : ENJEUX, USAGES ET PERSPECTIVES

XVIIe Forum NTIC et les maladies rares

Lundi 26 septembre 2016

www.orpha.net





SOMMAIRE

| | |
|---|-----------|
| OUVERTURE DU FORUM | 4 |
| <i>Sophie DANCYGIER</i> | <i>4</i> |
| OUVERTURE DES SESSIONS DE LA MATINÉE | 4 |
| <i>Alain DONNART</i> | <i>5</i> |
| <i>Ana RATH.....</i> | <i>5</i> |
| DEVENIR NUMERIQUE DE NOS DONNEES DE SANTE | 6 |
| Les données de santé, entre définitions et | 6 |
| <i>Marc HANAUER</i> | <i>6</i> |
| Evolutions réglementaires concernant l'accès et l'utilisation des données de santé à des fins de recherche | 7 |
| <i>Frédérique LESAULNIER</i> | <i>8</i> |
| Données de santé dans un monde hyperconnecté : enjeux éthiques | 12 |
| <i>Emmanuelle Rial-Sebbag.....</i> | <i>13</i> |
| Les préoccupations des malades | 16 |
| <i>Viviane VIOLLET.....</i> | <i>17</i> |
| La vie insoupçonnée des données de santé | 20 |
| <i>Marc HANAUER</i> | <i>20</i> |
| Table ronde | 22 |
| OUVERTURE DES SESSIONS DE L'APRÈS-MIDI | 27 |
| LES DONNÉES DE SANTÉ, DU SOIN VERS LA RECHERCHE | 27 |
| Impact des filières et Réseaux européens de référence (ERN) | 27 |
| <i>Dr Guillaume JONDEAU</i> | <i>28</i> |
| Exploitation des données des dossiers patients : mythe ou réalité ? | 30 |
| <i>Rémy CHOQUET BNDMR</i> | <i>30</i> |
| Consentement et e-consentement..... | 33 |
| <i>Anne-Sophie LAPOINTE.....</i> | <i>34</i> |
| Table ronde et échanges avec la salle | 37 |
| <i>Rémy CHOQUET.....</i> | <i>41</i> |
| SPÉCIFICITÉ DES DONNÉES DE GÉNÉTIQUE | 42 |
| NGS : une nouvelle ère de collecte de données génomiques..... | 42 |
| <i>Ana Rath.....</i> | <i>42</i> |
| Données de génétique et réseaux sociaux..... | 45 |
| <i>Denis COSTELLO</i> | <i>45</i> |
| Découvertes fortuites lors du séquençage..... | 47 |
| <i>Anne-Sophie LAPOINTE VML</i> | <i>47</i> |
| Table ronde et clôture du Forum | 49 |



OUVERTURE DU FORUM

Sophie DANCYGIER

Déléguée générale de la Fondation Groupama pour la Santé

Bonjour à toutes et à tous. Bienvenue à ce XVIIe Forum !

Je tiens à remercier Ana Rath et son équipe pour la qualité de leur travail, mais aussi à saluer le projet stratégique de l'Alliance Maladies Rares, dont les objectifs sont très ambitieux. Nous sommes très heureux d'être vos partenaires depuis tant d'années.

Ce Forum s'inscrit pleinement dans notre axe de soutien à l'innovation et à la recherche. A cet égard, je suis heureuse de vous informer que nous reconduirons, cette année encore, notre formation des associations au financement participatif - outil très efficace de levée de fonds pour qui sait l'utiliser. Cette formation sera animée par la plateforme HelloAsso dont nous sommes partenaires. Désireux de renforcer encore notre soutien à la recherche, nous avons mis en place un nouveau prix grâce auquel nous soutiendrons le projet de recherche d'une équipe durant cinq ans, à hauteur de 100 000 euros par an. Les présidents du comité scientifique ont déjà auditionné neuf lauréats, qui seront auditionnés le 8 novembre prochain. Le conseil d'administration désignera ensuite le meilleur dossier, sur les recommandations du comité scientifique.

L'autre grand axe de notre mission est celui de la proximité et de la solidarité. Vous avez peut-être déjà participé à nos « balades solidaires » qui nous ont permis, l'année dernière, de collecter 220 000 euros de dons - qui ont été en totalité redistribués aux associations. Cette année, nous avons déjà collecté un peu plus de 212 000 euros.



OUVERTURE DES SESSIONS DE LA MATINÉE

Alain DONNART

Président d'Alliance Maladies Rares



Les associations sont également très fières d'être partenaires de la Fondation Groupama pour la santé, que nous remercions encore pour son accueil. L'Alliance Maladies rares et la Fondation Groupama sont nées quasiment au même moment et cheminent ensemble depuis le départ.

Je salue également Ana Rath, que je remercie d'associer l'Alliance Maladies Rares aux équipes d'Orphanet, ainsi que le président de Maladies Rares Info Services, Jacques Bernard.

Les associations, ce sont avant tout des personnes malades ou des aidants - la famille ou toute personne très impliquée dans les maladies rares. Je suis convaincu que les nouvelles technologies sont l'avenir et qu'elles deviendront le petit « remède miracle » qui soignera de nombreuses maladies et qui aidera les malades et leurs familles à être beaucoup plus libres dans leur vie.

Ana RATH

Directrice d'Orphanet



Je vous remercie pour votre présence à tous, en particulier Sophie Dancygier pour son accueil indéfectible tous les ans et Alain Donnart pour son implication qui permet à Orphanet de tenir tous les ans cet engagement de tenir un Forum annuel des associations.

Aujourd'hui, nous verrons comment sont générées et exploitées les données massives (Big Data) de santé, avec les risques et les opportunités que cela présente. Ce matin, nous traiterons des données de santé en général avant, cet après-midi, d'aborder le sujet particulier du partage des données de santé entre la clinique et la recherche. Nous terminerons cette journée en nous focalisant sur les données génétiques, qui sont d'une extrême importance pour le diagnostic et la prise en charge des maladies rares.



DEVENIR NUMERIQUE DE NOS DONNEES DE SANTE

Les données de santé, entre définitions et

Marc HANAUER
ORPHANET

Reprenons la définition du dictionnaire : la santé est l'état de bon fonctionnement de l'organisme, ou « le silence des organes », donc l'absence de maladie. En 1948, le préambule de l'OMS va plus loin en définissant la santé comme « un état complet de bien-être physique, mental et social qui ne consiste pas seulement en l'absence de maladie ou d'infirmité ». La santé est donc un concept relatif, à la fois biomédical et social, dépendant aussi des conditions de vie et de l'environnement.

Quelques notions de base

Les **données de santé** sont elles aussi un concept en évolution constante, ne serait-ce que parce que nous sommes de plus en plus producteurs de ce type de données. En l'absence de définition juridique, c'est la jurisprudence française et européenne qui s'applique.

Les données de santé en général comprennent toutes les données médicales ainsi que les données relatives aux déterminants généraux de santé, et à la santé d'une personne ou d'un groupe de personnes. Pour leur part, les données médicales sont les données sur l'état de santé physique ou mentale d'une personne, qui sont générées par un professionnel de santé dans un contexte médical (données de diagnostic et de traitement des données associées).

Par ailleurs, la récente apparition des données dites « de bien-être » accompagne le développement des objets connectés qui permettent de prendre des mesures de soi (nombre de pas, rythme cardiaque, etc.). Il s'agit, là encore, d'une notion juridiquement floue.

Qui plus est, dans un monde numérique et d'objets connectés, les données sont de plus en plus globalisées - et la santé n'y échappe pas. En 2014, la Commission « open data en santé » a remis au Ministre des affaires sociales et de la santé un rapport dans lequel elle tente de dresser une typologie des données de santé : données qui sont générées par les professionnels de santé, par les établissements de santé, par les autorités dans le cadre des services publics (systèmes de remboursement de soins), par les industriels et les entreprises du secteur de la santé (laboratoires, mutuelles, etc.), mais aussi données générées par la recherche et par les patients (témoignages, informations remontées sur des événements indésirables, informations générées à l'aide de capteurs et d'analyse de l'activité physique).

Enfin, les **données personnelles de santé** sont susceptibles de révéler une maladie ou une prédisposition à la maladie. Dans les systèmes d'information certaines d'entre elles sont « dés-identifiées » (pour conserver l'anonymat), mais peuvent être « ré-identifiantes » lorsqu'elles permettent d'identifier une personne *via* des croisements de données (parcours de soins, dates d'hospitalisation, etc.). C'est tout l'enjeu de l'open data (données ouvertes), tout particulièrement dans

le domaine des maladies rares où des données même très peu nombreuses peuvent être ré-identifiantes.

Définition européenne

Le règlement européen de mai 2016 relatif à la protection des données personnelles définit les données de santé comme les données à caractère personnel relatives à la santé physique ou mentale, indépendamment de leur source. Cette définition inclut également les données génétiques et biométriques (empreinte digitale pour allumer un téléphone, par exemple).

Quelles attentes ?

Il ressort d'une récente étude du cabinet Deloitte sur les objets connectés que plus de 70 % des personnes interrogées se déclarent prêtes à partager leurs données de santé.

Les attentes portent principalement sur :

- l'amélioration de la prise en charge des malades ;
- la réduction des dépenses de santé ;

- la création de bases de données permettant de croiser et de valoriser les informations ;
- la création de connaissances (en particulier dans le domaine de la recherche).

En outre, le marché des données de santé - estimé à plus de 300 milliards d'euros d'ici 2022 - représente un fort potentiel tant pour les industriels du médicament, l'économie de la santé, les assurances et les mutuelles que pour les entreprises technologiques et...les moteurs de recherche.

Vers une nouvelle médecine

Nous passerons rapidement d'une santé productrice de données à une santé déterminée par les données, et d'une médecine qui est l'art de guérir à une science de la mesure, ou une médecine dite « 4P » : prédictive, préventive, personnalisée et participative - afin que la prise en charge et les traitements soient adaptés aux caractéristiques des réactions individuelles dans des environnements différents.



Evolutions réglementaires concernant l'accès et l'utilisation des données de santé à des fins de recherche

Frédérique LESAULNIER

ITMO Santé publique, INSERM

Le droit de la protection des données de santé est en pleine mutation tant au niveau européen, que national.

Le cadre juridique de la protection des données de santé

La France dispose d'un cadre juridique très riche, qui définit les conditions d'accès et d'utilisation des données de santé et en assure la protection. Ce cadre traduit le caractère « sensible » des données de santé qui sont des données susceptibles de relever de l'intimité de la vie privée des individus, et qui doivent, à ce titre, bénéficier d'un statut protecteur particulier, largement baigné par le secret qui permet d'en garantir la confidentialité.

Chaque patient a ainsi le droit de bénéficier du secret médical, un droit qui s'enracine dans le droit au respect de la vie (C. civ., art. 9), qui trouve un prolongement dans le code de la santé publique et dont la violation est pénalement sanctionnée (CP, art. 226-13). De façon plus générale, les conditions d'accès et d'utilisation sont précisées par la loi « Informatique et libertés » du 6 janvier 1978, modifiée en 2004 pour transposer la Directive 95/46/CE du 24 octobre 1995.

Les données « sensibles » (parmi lesquelles figurent les données de santé) sont frappées d'une présomption de non-pertinence en raison du risque de discrimination qu'elles comportent, et sont

soumises à un principe d'interdiction de traitement, assorti d'exceptions et de garanties qui permettent l'encadrement consécutif de leur accessibilité et de leur usage.

Cette protection s'inscrit dans une dimension européenne. Notre « modèle européen » de protection des données, qui n'est pas universel, s'est construit par le droit et a placé la personne et sa protection au cœur du système. La protection des données personnelles est considérée comme un droit fondamental, inscrit dans la loi et se traduit par des règles de bon usage, des droits reconnus aux personnes concernées au respect desquels une autorité administrative indépendante (la CNIL) veille et peut prononcer des sanctions en cas de non-respect des règles.

De nouveaux défis pour la protection des données

Depuis 1995, il y a eu des transformations considérables des intérêts en jeu et un gigantesque changement d'échelle. Le droit sur la protection des données de santé doit être adapté, suite aux modifications du traitement des données depuis 1995 et le développement d'Internet, des communications, des objets connectés, qui permettent :

- une massification des données (Big data, Open data...) ;
- une banalisation de la captation automatique des données dans toutes les activités du quotidien, le plus souvent à l'insu des individus ;
- une évolution des usages et des comportements (stratégies

de dévoilement de soi sur le Web social, participation à des communautés) ;

- une capacité croissante à exploiter des volumes de nature et de sources diverses quasi instantanément, de plus en plus finement et pour des coûts de moins en moins élevés (grâce au cloud computing). ;
- une globalisation et une marchandisation des données personnelles qui sont au cœur du modèle économique du numérique

Le développement accéléré de la numérisation des données de chaque citoyen offrira des opportunités extraordinaires pour le progrès de la connaissance et la prise en charge médicale (individuelle et collective), mais aussi pour le développement économique et l'emploi. Le secteur de la santé n'échappe pas à ces évolutions. Dans le même temps, les enjeux d'emprise sur les données par aussi bien des acteurs tiers et de puissances étrangères (ce dont témoigne l'affaire Snowden) sont majeurs et il ne se passe pas de semaine sans que soient dévoilées de nouvelles failles de sécurité parfois massives.

Face à ces nouveaux défis, Les équilibres arbitrés en 1995 et en France en 2004 sont remis en cause. Les principes traditionnels de protection des données restent valables mais leur mise en œuvre effective est compromise ou ils sont difficilement applicables au phénomène du Big Data. Le Règlement européen du 27 avril 2016 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 propose de nouveaux leviers pour donner un nouveau souffle au régime de protection.

Le droit de la protection des données à la croisée des chemins

Le Règlement général relatif à la protection des données personnelles du 27 avril 2016 (adopté après quatre ans de négociations sera directement applicable à l'ensemble des Etats membres de l'Union européenne à compter du 25 mai 2018. Toutefois, les Etats restent libres de prévoir des règles plus protectrices, en particulier en ce qui concerne le traitement de « catégories particulières de données à caractère personnel » (art. 9).

Le règlement devra donc être articulés au niveau national, avec la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, et la loi n° 2016-41 de modernisation de notre système de santé du 26 janvier 2016 (art. 193) (spécifique à la santé), qui sera complétée par la loi n° 2012-300 du 5 mars 2012 relative aux recherches impliquant la personne humaine (dite loi «Jardé ») qui est entrée en application le 17 novembre 2016.

Le règlement européen du 27 avril 2016 : des catégories particulières de données

Le règlement évoque les « catégories particulières de données » (et plus les données sensibles) et ajoute à la liste initiale. Il étend le champ d'application du texte aux données génétiques et aux données biométriques aux fins d'identifier une personne physique de manière unique. Ces ajouts s'accompagnent de nouvelles définitions (art. 4) relatives aux données génétiques, biométriques et aux données concernant la santé. Celles-ci ne concernent plus seulement les

données qui permettent d'indiquer la pathologie dont peut être atteint un individu (les données médicales), mais sont étendues à toute donnée sur sa santé physique et mentale qu'elle qu'en soit la source et toutes les prestations de services de santé servies ainsi que l'environnement médico-social de la personne, auxquelles on va étendre le régime de protection.

Il élargit considérablement la notion de données de santé, qui recouvre désormais les données médicales, les prestations de soin ou encore le NIR, quelle que soit leur source.

Le règlement maintient le principe d'interdiction du traitement de données de santé ainsi que les dérogations classiques parmi lesquelles figurent les recherches scientifiques et statistiques.

Par ailleurs, il renforce significativement la responsabilité des acteurs et la prise en compte des risques, faisant passer d'une logique de formalités préalables à une logique de conformité dynamique et continue grâce au développement de politiques d'*accountability* (procédures internes décuplées) :

- obligation de documenter l'ensemble des mesures internes définies et prises afin de garantir la conformité au règlement ;
- obligation de tenir un registre de tous les traitements ;
- obligation de mener des analyses d'impact relatives à la protection des données pour les traitements qui présentent des risques élevés pour la vie privée ;
- obligation de prendre en compte la protection des

données dès la conception du projet « Privacy by design » (minimisation, anonymisation) ;

- obligation de notifier les failles de sécurité à l'autorité de contrôle dans les 72 heures qui suivent la prise de connaissance et aux personnes concernées si la violation présente un risque élevé pour elles.

De nouveaux outils ont été mis en place pour permettre aux acteurs du traitement de données sensibles de faire face à leurs responsabilités :

- désignation d'un délégué à la protection des données obligatoire pour tous les organismes publics traitant de données sensibles à grande échelle ; les autorités ou organismes publics, les organismes dont les activités de base consistent en des traitements de données de santé à grande échelle et les organismes dont les activités de base exigent un suivi régulier et systématique à grande échelle des personnes concernées promotion de codes de bonne conduite par secteur d'activité ;
- élaboration de procédures de certification et de labellisation.

Le règlement européen renforce également le droit des citoyens, qui attendent d'être de plus en plus acteurs de leur santé :

- obligation d'information renforcée et plus large ;
- consentement renforcé (indubitable) et modernisé ;
- nouveaux droits (droit à l'oubli, transparence sur les failles de confidentialité, etc.).

Enfin, il renforce le rôle des autorités de protection des données :

- augmentation des amendes également applicables aux sous-traitants (20 millions d'euros ou 4 % du chiffre d'affaires mondial d'une entreprise) ;
- possibilité de mener des actions conjointes ;
- création d'un nouveau *Board* en remplacement du G29, afin de favoriser la cohérence de l'interprétation des règles européennes.

La loi de modernisation du système de santé français

La France est l'un des rares pays européens disposer de bases de données administratives et médico-administratives nationales gérées par des organismes publics, concernant des domaines stratégiques (recours aux soins, médecine de ville, etc.) et couvrant la population de façon exhaustive, continue et permanente - en utilisant un identifiant individuel unique (le NIR ou un dérivé).

L'article 193 de la loi de modernisation de notre système de santé prévoit la création d'un système national des données de santé (SNDS) regroupant le Sniiram (Système national d'information inter-régimes de l'assurance maladie), le PMSI (Programme de médicalisation des systèmes d'information), les causes médicales de décès, les données relatives au handicap des MDPH (maisons départementales des personnes handicapées) et, à terme, un échantillon des données de remboursement de l'assurance maladie complémentaire. Toutes ces données seront chaînées en routine et leur accès et leur appariement

avec d'autres bases de données sont favorisés pour des utilisations qui répondent à un intérêt public moyennant la mise en œuvre de garanties très importantes visant à protéger la vie privée des individus. Un référentiel de sécurité est notamment en cours d'élaboration, qui accroît les obligations de traçabilité des accès et de confinement des données.

Deux finalités sont interdites par la loi : la promotion commerciale de produits de santé et la sélection du risque (exclusion d'un contrat d'assurance ou modification des cotisations et des primes d'assurance, par exemple).

La loi distingue différentes modalités d'accès en fonction du risque d'identification des personnes :

- les données anonymes sont librement accessibles et réutilisables gratuitement par tous ;
- toutes les autres données, qui portent en elles un risque d'identification des personnes sont accessibles de façon encadrée, contrôlée en distinguant des accès permanents (missions de service public) et des accès ponctuels (spécifiquement liés à un projet contribuant à une finalité d'intérêt public).

L'article 193 assouplit les conditions d'utilisation du NIR pour la recherche, ce qui sera de nature à favoriser les appariements entre le SNDS et les cohortes et les registres. Mais il fixe aussi de nouvelles contraintes pour la recherche, *via* de nouvelles procédures de demande d'avis et une articulation/coordination nécessaire avec d'autres textes, notamment les

textes d'application de la loi « Jardé ».

Parce qu'aucune équipe de recherche ne dispose, seule, de l'ensemble des compétences requises (juridiques, scientifiques, techniques, logistiques) pour exploiter ces données complexes, le besoin d'accompagnement des chercheurs et de mutualisation des savoir-faire et des compétence est réel, tant sur des aspects réglementaires, éthiques que sur des aspects scientifiques et techniques (SI). C'est pourquoi

l'Inserm, dans le prolongement des services d'accompagnement des chercheurs déjà mis en place, développe actuellement une plateforme de soutien aux chercheurs afin de promouvoir, au bénéfice de la santé des populations, l'utilisation des données de santé dans le respect du cadre éthique et réglementaire et de la confidentialité des données personnelles.





Données de santé dans un monde hyperconnecté : enjeux éthiques

Emmanuelle Rial-Sebbag
INSERM
Université Paul Sabatier-
Toulouse III

Il est primordial que les juristes puissent interagir avec les représentants d'associations.

Le droit et l'éthique entretiennent des rapports très étroits, sans pour autant être exactement la même chose. En effet, le droit est ce que sanctionne la puissance publique et que nous avons l'obligation de respecter. L'éthique est parfois sanctionnée par le droit, mais elle répond avant tout à une approche aspirationnelle : ce que l'on souhaite pour le bien commun - sans être nécessairement toujours en adéquation avec le droit. Elle pose la question du monde de demain, pour nous et pour les générations futures.

Données de santé et Big Data ?

Comme cela a été rappelé ce matin, il n'existe pas de définition juridique des données de santé en droit français. Certes, une définition est fournie par le règlement européen de 2016, mais il s'agit d'une définition large qui intègre toutes les données des pathologies, qu'elles soient ou non recueillies dans d'autres contextes que celui du colloque singulier.

Les guides de la Cnil apportent quelques indices sur les données qui pourraient être qualifiées de données de santé en France : données médicales, données médico-administratives, données de recherche, données de santé publique - toutes considérées comme des données sensibles. En droit,

cette sensibilité entraîne par des droits (pour les personnes auprès de qui l'on recueille les données) et des devoirs (pour les personnes qui recueillent et traitent les données).

Par ailleurs, le recueil de données de masse existe de longue date, dans le cadre des cohortes ou des études génétiques par exemple. Mais le Big Data est nouveau en ce qu'il interroge les sources et leur diversification :

- les progrès techniques, qui permettent l'interopérabilité des bases de données ;
- le *crowd sourcing*¹, qui démultiplie les données d'ordre déclaratif ;
- les objets connectés, qu'il s'agisse de dispositifs médicaux évalués ou contrôlés ou qu'il s'agisse de dispositifs non contrôlés car délivrant des données dites « de bien-être » mais dont certaines pourraient s'apparenter à des données de santé.

Cette provenance multiple des données, de même que leur facilité et leur rapidité de circulation donc d'échange posent aujourd'hui la question de la responsabilité. Partager ses données génétiques, c'est aussi partager celles de sa famille, par exemple. Et si chacun est libre de communiquer sur sa propre santé, il existe aussi une forme de

¹ Lorsque les personnes délivrent elles-mêmes leurs données, à travers des réseaux sociaux ou des groupes fermés.

responsabilité morale quant au respect d'un professionnel de santé, d'un hôpital ou d'autres personnes qui se trouvent dans la même situation de santé. Les individus ont de nouveaux devoirs du fait qu'ils produisent eux-mêmes et mettent à disposition leurs données de santé.

Se pose aussi la question des utilisateurs « en chaîne » - avec un risque de perte de contrôle par l'individu lui-même. Aussi faudrait-il réfléchir à la protection de nouveaux droits, au-delà de ceux des personnes sources.

Enfin, la question de la gestion du risque informationnel (capacité ou incapacité de ré-identification des données) mérite elle aussi d'être posée. Dans un contexte de multiplication des sources et de croisement des données, quelles mesures mettre en œuvre et qui doit le faire : la loi ou les professionnels ? Il est encore difficile de savoir si le risque est réel ou potentiel - d'où la difficulté à proposer des mesures pour encadrer le risque informationnel ou y faire face

Questions éthiques

- Consentement et contrôle

L'objectif du consentement est de permettre aux personnes d'exprimer leur volonté. Le premier principe qu'il convient de garantir, dans le cadre de l'expression du consentement et de la maîtrise de l'utilisation des données de santé (identifiantes et sensibles) est celui de l'autonomie de la volonté. Ainsi, le consentement initial est juridiquement obligatoire (positif ou présumé) pour le recueil des données. Mais comment garantir que l'on

consent à des utilisations multiples, étalées dans le temps et déterritorialisées ? Le e-consentement est une voie possible, mais est-elle suffisante ?

- Accès et partage des données

L'utilisation des données de santé est largement encouragée. Elle est même un impératif pour le soin, mais aussi faire avancer la recherche et créer de la connaissance (*a fortiori* pour les maladies rares). Ce principe n'est donc pas à remettre en question, mais il doit s'accompagner de garde-fous et d'une gouvernance, pour assurer la confidentialité (en particulier dans le domaine des maladies rares où les effectifs sont très peu nombreux), la sécurité et l'intégralité des données. Il faut aussi continuer à rappeler - et là, les associations ont un rôle à jouer - le principe de protection de la vie privée : on a le droit de communiquer sur sa vie privée, à condition de ne pas porter atteinte à celle des autres. Toute cette approche ne peut pas se faire à l'échelle nationale, une gouvernance internationale est indispensable. En effet, tous les consortiums sont internationaux aujourd'hui.

- Justice et équité

Il existe des tensions entre le niveau individuel et le niveau collectif dès lors que les données numériques de santé (personnelles, donc individuelles) doivent être mises au service de la collectivité, pour le bien

commun. Du côté individuel, il s'agit de garantir la non-discrimination, la conformité des usages autorisés avec la finalité annoncée, mais aussi un bénéfice individuel (un médecin nous traite pour notre propre bénéfice). Du côté collectif, les valeurs à promouvoir sont celles de la solidarité, du bien commun et du bénéfice pour les générations futures. Deux principes fondamentaux se trouvent donc, d'emblée, en tension : le bénéfice individuel et le bénéfice collectif.

- Valeur et propriété

Les données de santé ont une valeur économique. Qui peut en récolter les fruits, qui en est le propriétaire ? Quel retour vers les donneurs ? Quels impératifs pour l'utilisateur ? Juridiquement, nous ne sommes pas propriétaires de nos données. Mais éthiquement, ce n'est pas certain. De tous ces doutes, découlent les droits économiques sur les données. Si je n'ai pas le droit de vendre mes données, pourquoi celui qui les a recueillies aurait-il le droit de le faire ? Par ailleurs, l'effort et la contribution des donneurs doivent être reconnus. Les données de santé sont données sur le modèle du don. Mais ne pourrait-on pas imaginer un système de reconnaissance pour les donneurs, sans qu'il s'agisse nécessairement d'une rétribution personnelle (clarté des informations, communication des résultats, etc.) ?

La nécessaire clarification du rôle des institutions

Qui fait quoi ? Comment s'assurer d'un usage intègre et pertinent de nos données (chartes, outils, etc.) ?

Quelle est la place du secteur privé dans cet environnement ? Le privé ne doit absolument pas être diabolisé, car il est essentiel qu'il contribue lui aussi à la production de connaissances. Pour autant, sa place doit-elle être la même que les autres ?

Les institutions publiques peuvent-elles constituer des tiers de confiance ?

Et quelle est la place des individus dans cette nouvelle gouvernance ?

En conclusion

Les géants du Web s'intéressent de très près à nos données de santé - et pas seulement dans une visée altruiste. Comment réguler une partie de ce modèle économique ?

Il est illusoire de continuer à garantir la protection et la confidentialité, car le risque d'identification est réel. Il faut donc trouver d'autres mécanismes :

- faire reposer le risque sur ceux qui traitent les données ;
- développer des cultures d'établissement ou d'entreprise fondées sur la loyauté et la transparence ;
- encourager le rôle de relais de l'information et de sensibilisation par les associations ;
- mettre en garde contre les risques d'atteinte de la vie privée

Enfin, la digitalisation et la numérisation de l'humain doivent

nous interroger sur notre futur. C'est un défi pour les générations futures, mais aussi pour notre responsabilité actuelle. Les relations humaines restent une base solide de production de connaissances. Restons donc prudents quant à la nature du monde que nous sommes en train de créer.

Ana Rath

Nous en sommes en train de construire l'idée d'acteurs responsables, de libertés pleinement exercées en toute responsabilité. Nous sommes tous des producteurs de données, des donneurs. A cet égard, les associations qui représentent des malades se situent au cœur de la tension entre l'intérêt privé et l'intérêt collectif.



Les préoccupations des malades

 Viviane VIOLLET

Comité Ethique et Société de
l'Alliance Maladies Rares

Depuis un an, le Comité « éthique et société » de l'Alliance Maladies Rares travaille sur la question du traitement des bases de données et j'en profite pour remercier tous les membres de ce comité pour le travail accompli.

La spécificité des maladies rares dans le traitement des données de santé

La numérisation des données de santé est une avancée incontournable mais qui présente toutefois des problèmes éthiques pour toutes les maladies, y compris les maladies rares dont les spécificités sont légion : 80 % de nos maladies ont une signature génétique et le profil génomique est parfois indispensable pour comprendre et soigner la maladie. La composante génétique y est plus importante que dans les autres maladies, engageant de fait la parentèle. Certes, il importe de faciliter la recherche, et les patients sont les premiers à le souhaiter. Mais il faut aussi informer ces derniers qu'il peut exister des données imprévues, et les tenir au courant de tous les résultats, pas uniquement ceux pour lesquels ils font une analyse - tout en se posant, malgré tout, la question de la pertinence de l'annonce de certains résultats en l'absence de traitement et de prise en charge.

Le faible nombre de malades rend la collecte plus difficile et plus longue. Chaque donneur est précieux, en particulier dans le cas de maladies très rares, et la pression morale peut

être forte. Sans compter que l'errance thérapeutique implique aussi la collecte de données intrusives. Un membre du Comité « éthique et société » a d'ailleurs fait le témoignage suivant :

« Ma maladie est tellement rare qu'en donnant mes données, j'ai tout donné - pour moi et pour aider les autres. Mais aujourd'hui, je ne me sens pas si bien que cela. Je crois que j'ai trop donné. »

J'ai entendu, ce matin, que l'intérêt pour le collectif ne doit pas se substituer à l'intérêt pour sa vie privée. Il faut, en effet, apprendre à la protéger. Outre la formation, un véritable accompagnement est nécessaire.

Les motivations des malades à fournir leurs données de santé

Les motivations à livrer ses données sont différentes suivant qu'il s'agit de données de soin ou de données de recherche. Dans le premier cas, le malade n'a pas besoin de participer à la recherche pour être soigné. Il n'a pas de consentement à signer et sans hésitation, même sous le coup de la douleur, il fait ce qu'on lui demande et il « donne » tous les renseignements demandés d'ordre cliniques, analytiques etc.. Dans le deuxième cas, il a plus de recul mais il est généralement très motivé car il peut attendre des bénéfices individuels et souhaiter apporter sa contribution à une meilleure connaissance de sa maladie en général.

Dans tous les cas, la motivation essentielle est l'attente d'un bénéfice pour soi, pour ses proches et pour

les autres malades, avec une volonté de ne pas subir de conséquences négatives. D'emblée, il existe chez les malades un esprit de solidarité qui peut amener une confusion entre soin et recherche, vie privée et intérêt collectif. D'où, encore, une nécessité absolue d'information. Il faut prendre le temps d'informer et d'expliquer ce qu'est le consentement. C'est essentiel, pour nous.

Les malades redoutent particulièrement l'utilisation de leurs données par des utilisateurs peu ou non bienveillants (assurances, banques, etc.).

Les questions des malades face à l'utilisation de leurs données de santé

Avec l'immédiateté permise par le numérique, les malades ont parfois le sentiment que les choses peuvent leur échapper. Dans ce contexte, les questions sont nombreuses.

- *Quid* de la possibilité de refuser de fournir des données ou de revenir sur son consentement ?
Théoriquement et juridiquement c'est possible, mais psychologiquement il est difficile de ne pas céder à la pression notamment dans le cadre des soins. Il est important que le professionnel responsable de la recherche prenne du temps pour expliquer et qu'il trouve le moyen de garder le lien avec le malade qui fait intégralement partie de la recherche.
- *Quid* de la finalité du don et de la transparence : où vont les examens, et pour quelle(s) recherche(s) - d'autant qu'il

semblerait que la loi ouvrirait la possibilité d'un changement de finalité, pour permettre des études comparatives notamment ? Là encore, on revient au consentement du patient et à la transparence sur le devenir des échantillons prélevés. Les malades veulent savoir à quoi vont servir leurs données cliniques ou analytiques ou leurs prélèvements.

- *Quid* du circuit des données (lieux de stockage, transferts, conservation), de l'accès aux données (par quels professionnels, avec quelle fiabilité du codage et de l'anonymisation) et du secret médical (surtout lorsque le dossier médical est accessible par des ordinateurs interconnectés, par exemple quand les mutuelles sont au sein des banques) ?
- *Quid* du rendu des résultats ?
- *Quid* des promesses de la génétique ? Cette dernière génère un espoir énorme. Mais tiendra-t-elle toutes les promesses que les malades mettent en elle ? De nouveaux dispositifs sont mis en place, sans qu'on en connaisse encore les bienfaits. Le séquençage à haut débit est intéressant, mais que se passera-t-il lorsque chacun découvrira tout son génome humain ?
- *Quid* de l'éventualité d'imprévus ?

Faut-il accepter de donner sans retour ? Travailler sur les maladies rares aide aussi pour les autres maladies, plus communes. Dès lors, nous donnons. Mais nous devons

réfléchir à ce que nous pouvons donner, et à qui - donc à faire preuve de plus de vigilance pour donner avec clairvoyance et, le plus souvent, sans retour.

La plus grande banque de données est Google, qui dispose de plusieurs millions de génomes. En outre, l'hyperconnexion des données permet à ces dernières de voyager, y compris dans des secteurs commerciaux ou des pays étrangers où leur protection n'est pas garantie. Nous avons d'ailleurs peur des pirates informatiques, qui peuvent décrypter des codages et des systèmes d'anonymisation. Dans nos maladies rares, ce peut être lourd de conséquences.

Les malades sont souvent en détresse. Ils acceptent donc parfois un peu trop rapidement d'être recruté pour des essais ou d'être directement mis en lien avec des laboratoires. N'oublions pas d'être bienveillants avec nous-mêmes et de respecter notre vie privée.

Le rôle des associations

Dans un monde hyperconnecté, les associations jouent un rôle qui se situe entre vigilance et conseil, en participant aux actions d'information sur les données de santé, mais aussi un rôle de précurseur pour la mise

en place de garde-fous. Pour nos maladies, les données sont rares. Mais elles sont traitées par une multiplicité d'acteurs : ceux-ci savent-ils suffisamment qu'ils sont investis de la confiance des malades, et que cela exige à la fois une pleine conscience de la responsabilité de chacun et de la bienveillance ?

Il est important de rappeler les droits des patients, mais aussi de parler du renforcement des droits du citoyen. C'est parce que nous avons un esprit citoyen que nous avons obtenu beaucoup d'avancées. Il nous faut continuer !

Alain DONNART Alliance Maladies Rares

Je suis ravi que la commission « Ethique et société » ait été prolongée dans le cadre du plan stratégique 2015-2020 de l'Alliance Maladies Rares.





La vie insoupçonnée des données de santé

Marc HANAUER

ORPHANET

Aujourd'hui, nous sommes tous producteurs de données de santé - parfois sans le savoir ou sans nécessairement en être conscients. Sur les réseaux sociaux, par exemple, les gens conversent de manière naturelle sur leur pathologie, produisant ainsi des données de santé qui peuvent être exploitées (traitement suivi, taux de glycémie, lieu de prise en charge, nom d'un praticien et avis sur celui-ci, etc.).

La question de l'exploitation des données postées sur les réseaux sociaux pour faire de la détection d'effets secondaires ou pour recruter dans le cadre d'essais cliniques s'est d'ailleurs très vite posée. De très nombreux articles citent d'ailleurs Twitter dans le cadre de la pharmacovigilance, et le projet européen Web-RADR a pour objectif de mettre en place d'un outil de surveillance des réseaux sociaux dans cette optique.

Les ambitions des acteurs du Web et des nouvelles technologies sont réelles, tant l'enjeu de la capacité d'analyse et de traitement de données est gigantesque au regard du marché de la santé. Google, par exemple, a signé un accord avec des hôpitaux londoniens lui permettant de récupérer 1,6 million de dossiers patients afin d'alimenter son intelligence artificielle « Deep Mind ». Les patients en question n'avaient sans doute pas conscience que leurs données seraient transmises dans ce cadre. De manière plus générale, le rachat de sociétés peut aussi se

traduire par du transfert de données patients.

Le Web du génome commence lui aussi à émerger. Produire de l'information de données génétiques n'est pas nouveau, mais cette production s'est accélérée d'autant que le coût de séquençage a significativement baissé. Un consortium américain, dont l'un des membres est Google (qui a lancé Google Genomics il y a quelques années) a ainsi créé un moteur de recherche sur les mutations génétiques.

Aujourd'hui, le nouveau terrain de bataille pour des acteurs comme Google et Amazon est le stockage d'ADN dans le *cloud*. En effet, le marché principal de ces acteurs n'est pas la vente de produits, mais la mise à disposition de plateformes de stockage et d'hébergement de données, et pas uniquement dans le domaine de la santé. Ils proposent ainsi d'héberger gratuitement des bases de données génomiques réputées comme « 1000 Genomes Project ». Pour citer un autre exemple, Amazon facture entre trois et cinq dollars seulement par mois pour stocker un génome humain complet. Ce service présente un réel intérêt pour les universités, tant en termes de délai de traitement de l'information qu'en termes de sécurité qu'elles ne maîtrisent pas nécessairement et en général elles ne disposent pas de la capacité suffisante en terme d'architecture technique pour évoluer aussi rapidement que les besoins en stockage et traitement des données.

Enfin, les données de santé ont une valeur sur le *darkweb*. La cote d'un

dossier médical, par exemple, s'établit à 17 dollars aujourd'hui. En 2015, un laboratoire français d'analyse sanguine s'est vu pirater 15 000 dossiers : les pirates ont menacé de diffuser sur le Web les informations contenues dans ces dossiers si le laboratoire ne lui versait pas de rançon. De la même façon, le système d'information d'un hôpital californien a été totalement

immobilisé suite à l'introduction de hackers ayant crypté ses données, les rendant inutilisables, et menaçant de ne les décrypter que contre le versement d'une rançon. Ces exemples montrent que les tentatives d'intrusion n'ont pas uniquement vocation à voler des données de santé.





Table ronde

- *Viviane VIOLLET, Alliance Maladies Rares*
- *Frédérique LESAULNIER, INSERM*
- *Emmanuelle RIAL-SEBBAG, INSERM*
- *Marc HANAUER, Orphanet*
- *Ana RATH, Orphanet*

Mado GILANTON, association APAISER

Les associations sont potentiellement en capacité de disposer d'importantes bases de données, notamment *via* la procédure d'adhésion en ligne. Ces bases sont-elles soumises à la réglementation de la Cnil ?

Par ailleurs, nos fichiers sont parfois envoyés par mail. Faut-il cesser de le faire ? Nous avons compris qu'il ne fallait rien stocker sur le *cloud*, mais de quels autres moyens disposons-nous pour agir à la fois en sécurité et en toute responsabilité ?

Frédérique LESAULNIER

Un fichier de patients étant un fichier de données à caractère personnel et d'une très grande sensibilité (l'intitulé de l'association renseignant directement sur la pathologie du malade, même si le fichier d'adhérents ne comporte que des prénoms et des noms), il est naturellement soumis à la réglementation de la CNIL. En tant que responsables d'association, vous êtes responsables du traitement des données que vous collectez. Toutes les règles que nous avons énoncées ce matin s'imposent à vous.

Si les fichiers sont hébergés chez un tiers prestataire, il est indispensable que ces derniers soient certifiés et agréés. S'ils sont hébergés par

l'association elle-même, des règles de sécurité doivent être mises en place et documentées. En outre, le fichier doit être déclaré à la CNIL - par le biais d'une déclaration simple, sans demande d'autorisation préalable dans la mesure où la constitution du fichier repose sur le consentement des personnes.

Marc Hanauer

L'hébergement du site Web chez un tiers prestataire d'une association coûte moins cher que l'hébergement d'une base de données de santé, qui nécessite de s'adresser à un hébergeur agréé pour cela. La question doit donc se poser au cas par cas, en fonction des données recueillies. La plus grande vigilance quant à la nature des données que l'on sera en capacité de stocker est indispensable, d'autant que toutes les associations n'ont pas les moyens adéquats.

De la salle

Chaque hôpital est relié à la Caisse d'assurance maladie, qui nous identifie clairement comme patient d'une maladie rare. Toute personne mal intentionnée peut avoir accès à nos données.

Emmanuelle RIAL-SEBBAG

Les personnels des institutions que vous venez de citer ont, par leur

statut, une obligation de confidentialité, y compris dans le cadre du secret partagé (lorsqu'une donnée a le droit de circuler entre plusieurs professionnels de santé, tous soumis à la même obligation de confidentialité). Certes, personne n'est jamais à l'abri d'un acte malveillant comme la divulgation d'une donnée à une personne non autorisée. Mais cette divulgation est illégale, qu'elle soit le fait d'un individu ou d'une entreprise. Il s'agit d'une fraude pénalement réprimée.

Frédérique LESAULNIER

Les obligations de prise de mesures effectives de sécurité (référentiel, authentification, traçabilité, etc.) renforcent l'obligation de confidentialité. Tous les niveaux de sécurité seront progressivement relevés. Le piratage d'un établissement hospitalier est possible, mais reste marginal en France.

De la salle

Le règlement européen sur le consentement éclairé dans le cadre de la recherche modifiera-t-il ce dispositif, dans son objet, dans sa durée et dans son usage ?

Emmanuelle RIAL-SEBBAG

Je n'ai pas l'impression qu'il change significativement les choses par rapport à la directive de 1995 sur la recherche biomédicale. Il vise à concilier un consentement spécifique pour une finalité donnée avec les nouvelles façons de faire de la recherche et d'utiliser des données personnelles, avec des usages en cascade - qui ne peuvent donc pas se satisfaire d'un objectif unique. Le nouveau règlement élargit davantage les finalités que le consentement en tant que tel. Je précise, toutefois,

qu'il ne s'agit là que de mon interprétation.

Frédérique LESAULNIER

En effet, il s'agit de permettre de conduire une autre recherche portant sur une même maladie rare. Il est important de ne pas déposséder les patients de leur consentement, ce qui passe par un consentement spécifique. Les droits dits de la personnalité sont indisponibles, ce qui signifie que l'on ne peut pas en disposer sans consentement spécifique.

Une participante

Nous distinguons soin et recherche. Lorsque nous nous rendons à l'hôpital pour recevoir un soin, nous donnons des images (radiographies) sans consentement, ou du moins avec un consentement implicite. Ces données sont, dans notre esprit, à visée diagnostique. Mais nous ne savons pas ce qu'elles deviennent ensuite. Qu'en est-il des fonds de tube et des lames de biopsie, par exemple, pour lequel on n'a pas demandé son consentement au patient ?

Emmanuelle RIAL-SEBBAG

Dans le contexte du soin, le consentement est certes implicite, mais l'information du patient est obligatoire - dans la relation du colloque singulier. Dans le contexte de la recherche, le cadre est celui de l'exploitation d'une ressource. C'est une tout autre échelle.

Concernant l'utilisation des fonds de tube ou des plaques de biopsie, la réglementation est stricte depuis 2004. En effet, le code de la santé publique prévoit qu'à l'exception des cheveux, des ongles et du lait maternel, aucun élément du corps humain ne peut être utilisé sans

consentement. Le patient doit *a minima* être informé que, de manière résiduelle et sans porter atteinte à la prise en charge médicale, une partie de sa biopsie sera utilisée en recherche, stockée dans la tumorothèque X dont Monsieur Y est responsable - et qu'il peut s'opposer à cette utilisation. De la même façon, les déchets opératoires ne peuvent plus être utilisés à l'insu du patient.

Frédérique LESAULNIER

Il en va de même pour les données médicales. Si le médecin qui a recueilli vos données souhaite les transmettre à un tiers (une personne externe à l'équipe de soin), il est obligé de respecter la réglementation de la CNIL sur la levée du secret et le tiers doit être autorisé à recevoir ces données.

Une meilleure information passe nécessairement par une meilleure implication du patient.

Caroline ROATTA, association Bernard Pepin pour la maladie de Wilson

Le rôle des associations change significativement, avec l'utilisation des réseaux sociaux. Le partage des données se structure en plusieurs niveaux : le partage pour la recherche, mais aussi le partage sur les réseaux sociaux. Ce dernier est-il toujours naïf et irresponsable ? N'est-il pas nécessaire aussi, pour certains patients ? En Amérique du Sud, par exemple, les réseaux sociaux sont quasiment l'unique source d'information sur les maladies rares. Comment pouvons-nous nous adapter à ce nouveau type de médiation ? Nous avons parfois le sentiment qu'il existe un décalage entre nos pratiques quotidiennes et la loi - ce qui nourrit la peur.

Emmanuelle RIAL-SEBBAG

Nous partageons cette interrogation, qui s'avère fondamentale. La grande difficulté vient du fait que la plupart des réseaux sont ouverts : même en communiquant à titre individuel sur Facebook, nous laissons des traces. Peut-être faudrait-il constituer des groupes fermés, mais aussi réinventer la fonction de modérateur ?

Marc Hanauer

Je précise que même au sein de groupes fermés sur Facebook, l'information appartient à Facebook.

Par ailleurs, le partage d'informations correspond souvent à une volonté d'échange et de partage d'expérience. Cette recherche de relations humaines est démultipliée par les réseaux sociaux - dont il ne faut pas nier les effets positifs.

Enfin, il y a quelques années, nous avons élaboré un guide pratique de l'animation et de la modération des forums. Mieux vaut être acteur et promoteur de bonnes pratiques sur les réseaux sociaux qu'en être totalement absent.

De la salle

Les MDPH - ou plus largement le secteur médico-social - détiennent de nombreuses informations médicales. La numérisation de leurs archives est-elle soumise à la protection des données ?

Frédérique LESAULNIER

Bien sûr, d'autant que le cadre réglementaire a élargi le secret professionnel de la sphère médicale à la sphère médico-sociale. La notion

d'équipe de soins couvre donc aussi le personnel médico-social.

Plus le temps passe, plus je suis persuadée qu'il n'est plus possible de travailler sur l'élaboration de guides de référence ou de formulaires types (recueil de consentement, par exemple) sans vous, sans les associations de patients.

Ana Rath

Le patient doit être acteur à tous les niveaux : le soin et la recherche.

Alain Donnart

Les patients et les associations sont naïfs : nous racontons tout et n'importe quoi sur les réseaux sociaux ! C'est souvent parce que nous recherchons tous azimuts des conseils et des informations.

Marc Hanauer

Il ne faut pas avoir peur des réseaux sociaux, mais avoir bien conscience des problématiques qu'ils posent.

D'où l'intérêt d'y être proactif et de promouvoir les bonnes pratiques.

Nous devons apprendre à vivre avec un monde de données globalisées. En avoir conscience doit nous inciter à envisager des modes de régulation et d'action les mieux adaptés.

Frédérique LESAULNIER

N'hésitez pas non plus à utiliser des pseudonymes - même si, dans votre situation, il est indispensable de se reconnaître pour échanger. Par ailleurs, le droit évolue, notamment en termes de droit à l'oubli.

Ana Rath

Le risque numérique existe, mais il faut faire avec - ce qui implique de le connaître, de se responsabiliser et de responsabiliser les autres. Dans cette démarche, les malades doivent être de véritables acteurs, du soin à la recherche en passant par la communication.



OUVERTURE DES SESSIONS DE L'APRÈS-MIDI

LES DONNÉES DE SANTÉ, DU SOIN VERS LA RECHERCHE



Impact des filières et Réseaux européens de référence (ERN)

Dr Guillaume JONDEAU
Cardiologue, Coordinateur Centre
Expert, AP-HP

Au sein du centre national de référence sur les syndromes de Marfan et apparentés, différents professionnels (cardiologue, ophtalmologues, pédiatres, radiologues, biologistes moléculaires, assistantes sociales, psychologues) recueillent de très nombreuses données médicales (imagerie, prélèvements tissulaires, etc.) sur les patients atteints de pathologies aortiques. Ces données représentent une véritable richesse pour la recherche. Elles sont entrées dans une base de données au fur et à mesure de la consultation - dans le centre national de référence mais aussi dans les différents centres de compétences sur le territoire.

L'intérêt des données médicales des patients pour les études

Une récente étude observationnelle à partir des données que nous recueillons auprès des patients a permis d'évaluer le risque de dissection aortique en fonction du périmètre de l'aorte, donc la pertinence du risque opératoire. Suite à ces résultats, les recommandations européennes et américaines ont été modifiées.

Nous avons également mené une étude interventionnelle, avec une participation plus active des patients, pour tester un médicament après des résultats de recherche prometteurs sur des souris. Nous avons pu inclure 300 personnes souffrant d'un syndrome de Marfan, après leur avoir

fait signer un consentement. Au terme de trois ans d'étude, nous avons montré que la vitesse de dilatation de l'aorte restait strictement la même dans le groupe placebo et dans le groupe qui avait reçu la molécule active - ce qui signifie que ce qui fonctionne très bien chez la souris est sans effet chez l'homme.

Les données des patients permettent aussi de conduire des études translationnelles (du patient à la recherche, puis de la recherche au patient). Nous suivons ainsi une famille dont certains membres présentent une dilatation de l'aorte mais chez lesquels aucun des gènes connus n'est en cause. Nous avons donc screené tout le génome des personnes sélectionnées, pour tenter de trouver une anomalie dans un gène nouveau puis rechercher d'autres mutations dans le même gène chez d'autres patients souffrant de la même pathologie en dehors de cette famille. En l'occurrence, nous avons trouvé une mutation dans un TGFB2 causant une forme familiale d'anévrisme de l'aorte thoracique. Aujourd'hui, ce gène est systématiquement inclus dans le panel des gènes recherchés en cas de forme familiale d'anévrisme de l'aorte.

La filière santé maladies rares FAVA-Multi

Le deuxième Plan national maladies rares (2011-2014) a créé 23 filières maladies rares. Parmi elles, la filière FAVA-Multi est constituée depuis le 1^{er} septembre 2014, pour une durée de cinq ans. Elle regroupe :

- trois centres de référence en charge de pathologies vasculaires génétiques (Syndromes de Marfan et apparentés, Maladies vasculaires rares, maladie de Rendu Osler) ;
- avec trente centres de compétences rattachés aux centres de référence ;
- trois associations de patients et de familles ;
- trois équipes de recherche affiliées ;
- sept laboratoires de diagnostic approfondi.

Nous avons un projet de recherche relatif à la paroi des vaisseaux, mais ce dernier n'est pas encore complètement défini. Quoi qu'il en soit, ce changement d'échelle s'avérera à coup sûr bénéfique.

Le réseau européen de référence VASCern

VASCern postule pour être reconnu comme le réseau de référence européen pour les maladies vasculaires rares multi-systémiques. 31 centres nationaux de référence souhaitent y participer, dont les trois centres français de la filière FAVA-Multi.

Les maladies vasculaires rares multi-systémiques ont été réparties en cinq groupes. Un groupe de patients a également été constitué, avec des représentants des cinq groupes de pathologies, ainsi que cinq groupes transversaux.

Il ressort des recensements des différents centres nationaux de référence européens que 6 784 adultes auraient une maladie

aortique en Europe (dont 1 200 en France). Et chaque année, les nouveaux patients représentent 2 040 adultes et 600 enfants. La participation à un réseau européen permettra donc un véritable changement d'échelle - ce qui est fondamental dans les maladies rares. Dans une précédente étude, par exemple, un travail à grande échelle avait permis de construire une courbe de survie faisant passer une pathologie de « gravissime » à « grave » (avec un taux de 75 % de survie à 75 ans). Cette approche multi-centrique et multi-pays a permis d'annuler les biais des recherches initiales et d'adapter le message sur cette pathologie.

Le réseau européen de référence VASCern

La recherche, ce sont avant tout de très nombreuses heures de travail !

Par ailleurs, je tiens à faire passer le message que la participation des patients à la recherche n'est pas dangereuse : les chercheurs sont plutôt bien intentionnés. Ils ne cherchent pas à dérober des données, mais à aider les patients.

Enfin, la confiance dans l'utilisation de ces dernières implique le retour sur investissement des patients, mais aussi des cliniciens et des chercheurs.



Exploitation des données des dossiers patients : mythe ou réalité ?

Rémy CHOQUET
BNDMR

L'un des enjeux de la Banque nationale de données maladies rares (BNDMR) est la mise en place d'un jeu de données commun au niveau national, afin de faciliter l'identification des patients atteints de maladies rares dans le système généraliste et de permettre la conduite d'études nationales visant à éclairer la décision publique tout en garantissant le caractère privé des données utilisées. L'une des stratégies est de s'appuyer sur les dossiers patients hospitaliers du réseau expert.

Le dossier patient

L'article R 1112.2 du code de la santé publique définit le dossier patient comme « *un document écrit, individualité, actualisé, qui regroupe l'ensemble des informations nécessaires et suffisantes pour déterminer et coordonner l'action du personnel soignant* ». Le contexte dans lequel les données sont collectées a donc toute son importance. Les données collectées dans le cadre du soin ont *a priori* pour seul but l'amélioration de la prise en charge du patient. Elles n'ont *a priori* pas d'objectif de recherche ou d'exploitation statistique.

Par ailleurs, « *le dossier patient prend en compte l'aspect préventif, curatif, éducatif et relationnel du soin. Il est le support de la démarche de soins et l'élément fondamental de la cohérence et de la continuité des soins. Il est nécessaire pour assurer le suivi du soin au sein d'une équipe*

de prise en charge ». La définition légale de l'équipe de prise en charge est en train d'évoluer, notamment avec la mise en place des territoires de santé.

Enfin, « *le dossier patient peut avoir une forme électronique ou papier. Il peut aussi prendre la forme d'un cahier d'observations personnelles à destination du professionnel de santé (et, par extension, à son équipe de soin, mais sans utilisation à l'extérieur)* ». Aujourd'hui, la majeure partie des archives est encore au format papier. D'autant que la numérisation ne va pas encore de soi pour tous les professionnels de santé, certains considérant que l'existence d'un écran entre le médecin et son patient n'est pas optimale dans la relation.

L'informatisation des données médicales : état des lieux

L'atlas des systèmes d'information hospitaliers indique que :

- 92 % des établissements ont engagé l'informatisation du dossier patient (85 % pour le dossier de soin) ;
- 55 % ont achevé leur projet d'informatisation (contre 51 % en 2014) ;
- 52 % ont achevé l'intégration des résultats d'examens de laboratoire (seulement 31 % pour l'imagerie).

Qui plus est, le niveau d'informatisation est peu avancé pour les dossiers médicaux structurés de spécialité.

En France, l'informatisation concerne en priorité et par ordre d'importance :

- le pilotage économique ;
- la prescription électronique ;
- les résultats d'imagerie ;
- l'agenda et les ressources de l'hôpital ;
- le parcours de soins au sein de l'hôpital ;
- le dossier patient.

Aujourd'hui, 80 % des dossiers patients contiennent à des données en texte libre (compte rendu de sortie, notes manuscrites de suivi de prise en charge). Tous contiennent également des données structurées pour le diagnostic de la prise en charge et des actes médicaux effectués donnant lieu à un remboursement par l'assurance maladie de la PEC tarifé. Mais seuls certains d'entre eux comportent des résultats d'examens (plus ou moins standardisés) ou des données sur la prescription médicamenteuse.

Quels enjeux ?

Pour les professionnels de santé, les principaux enjeux de l'informatisation sont :

- la facilitation de la collecte d'informations ;
- la capacité à retrouver l'information pertinente au bon moment ;
- l'aide à la prise de décision ;
- la garantie de sécurité, de confidentialité et de transparence des traitements.

En termes de traitement de l'information, les enjeux sont nombreux également :

- retrouver des informations pertinentes dans le cadre de la prise en charge ;
- identifier des patients correspondants à des critères ;
- identifier des patients comparables ;
- générer des alertes pour la sécurité de la prise en charge ;
- caractériser le phénotype des patients (pour permettre une interprétation pertinente des données) ;
- évaluer la lourdeur de la prise en charge médicale (et identifier s'il y a suffisamment de centres de référence, par exemple).

Reste à savoir si ces données peuvent être réutilisées par la santé publique et la recherche. La question est toujours posée, notamment pour le caractère informationnel qu'elles portent et les questions relatives à la qualité de celles-ci.

Quels outils pour faciliter la réutilisation des données de santé ?

Il existe des « entrepôts de données de santé », mais l'interprétabilité des données du soin reste encore trop souvent limitée à la personne qui les a entrées.

D'autres méthodes permettent d'interroger les données : des algorithmes (pour reconstruire un phénotype), de la fouille (ou datamining, pour trouver des motifs ou analyser des textes libres) ou des statistiques (générer des connaissances). Mais le plus souvent, les résultats restent encore assez restreints.

Pour augmenter la puissance analytique des informations qui

figurent dans les dossiers patients, il convient de produire davantage de connaissances. Les systèmes organisés de connaissances, notamment, permettent à la machine de mieux connaître les termes et les concepts comme « terminologie », « nomenclature », « modèle » ou encore « ontologie ». Ils sont construits par des sociétés savantes et des organisations internationales comme Orphanet.

Pour la fouille de données, par exemple, l'outil I2b2 permet de trouver plus rapidement des patients qui pourraient être éligibles à certaines études cliniques. A l'institut Imagine, le système Dr Warehouse permet pour sa part de trouver des patients similaires (un système équivalent a été développé à Rennes).

Pour l'aide à la décision, IBM Watson a été le premier ordinateur capable de trouver les questions à une réponse donnée en langage naturel dans le cadre d'un jeu télévisé, grâce à son immense base de connaissances. Aujourd'hui, il s'agit de voir s'il pourrait faire la même chose dans le domaine médical : pourrait-il indiquer quelle est la maladie en présence à partir de la description des symptômes présentés par un patient ?

D'autres exemples semblent prometteurs concernant la caractérisation du phénotype.

En fin de compte, il n'existe pas de solution unique, et aucune solution n'est pas applicable à tous les hôpitaux, toutes les questions, toutes les maladies et tous les malades. Les articles actuels soulignent d'ailleurs les limites de ces outils. L'une d'entre elles est que l'on retrouve surtout des patients très malades lorsque

l'on fouille pour trouver des patients pour entrer dans une étude, ce qui a un effet sur la qualité de la recherche.

Concernant la santé publique et les études, les méthodes classiques (échantillonnage, cohorte, registres, etc.) sont toujours d'actualité. En effet, on observe globalement peu d'intérêt à l'utilisation des données des dossiers patients pour les études car elles présentent plusieurs problèmes : qualité des données, manque de définition et de contexte pour les interpréter, contexte de collecte peut avoir un biais sur l'information collectée, le caractère très compliqué de l'accessibilité à grande échelle de ces données.

Une information cependant, une méthodologie de référence Cnil (MR-003) est parue en août 2016, pour les études sur données (non interventionnelles) ne nécessitant pas de consentement éclairé spécifique, mais une simple information active. Elle a pour ambition de pallier en partie la problématique de la définition *a priori* la finalité du traitement dans ces entrepôts de données hospitaliers.

En conclusion

La BNDMR a défini des objectifs d'exploitation, car il existe une relation étroite entre ces objectifs et la donnée qui sera collectée. Elle intègre la collecte spécifique dans les dossiers patients des hôpitaux qui y sont prêts, ou nous mettons à disposition une application de saisie spécifique. Ses sources sont au nombre de trois : le dossier patient informatisé, les BaMaRa et les registres/cohortes. Mais le cadre juridique de l'interconnexion de ces

derniers dispositifs avec la BNDMR est encore à l'étude.

Par ailleurs, il existe plus de 350 bases de données dans les centres de référence et de compétences maladies rares en France. Une réflexion est en cours pour savoir s'il est pertinent de les intégrer à la BNDMR.

Enfin, la base CEMARA mis en place par le Pr Paul Landais est actuellement déployé dans 40 % des centres de référence maladies rares. Il reste l'un des projets les plus emblématiques en termes de collecte

nationale de données pour les maladies rares. 350 maladies et 300 000 malades sont référencées. La file active est d'environ 40 000 nouveaux cas par an. *A priori*, un seul cas observé sur deux est confirmé, ce qui souligne la forte proportion des cas non encore confirmés. Même s'il existe un risque de biais en raison d'anomalies de saisie, ces chiffres sont intéressants dans la perspective de la mise en place du troisième Plan national maladies rares, en lien avec le plan « France Médecine génomique 2025 ».



Consentement et e-consentement

Anne-Sophie LAPOINTE
VML

La notion d'éthique de la recherche est en constante évolution. Du paternalisme du chercheur propre aux années 1940, nous sommes progressivement passés à une régulation protectrice, puis un accès accru à la recherche (parallèlement à l'affirmation du patient autonome). Et depuis quelque temps, nous nous orientons de façon de plus en plus marquée vers un partenariat avec les associations et les communautés de patients.

Particularités des maladies rares pour la recherche et le soin

Les questions autour du consentement et du e-consentement se posent avec une force particulière concernant les maladies rares, qui se caractérisent par

- une errance de diagnostic ;
- des difficultés à recruter des patients pour les essais cliniques ;
- une information limitée et complexe ;
- un manque de connaissances lié à une histoire naturelle de la maladie peu renseignée et des données réduites ;
- une dispersion géographique ;
- un manque de financement (seuls 5 % de traitements spécifiques maladies rares sont financés, pour 6 à 10 000 maladies rares dont 80 % d'origine génétique).

Accompagner les parcours de vie et soutenir le développement de la recherche

Le mouvement associatif a accompagné les deux premiers Plans nationaux maladies rares, la création des centres de référence et des filières santé maladies rares et la mise en œuvre du règlement européen sur le médicament orphelin. Il est le fruit de la volonté commune de toutes les associations de travailler dans le même sens. Pour aller plus loin encore, l'Alliance Maladies Rares plaide d'ailleurs en faveur d'une co-construction du troisième Plan national maladies rares.

Le consentement : comment faire confiance ?

Certains chercheurs envisagent l'éthique comme une contrainte dans le bon déroulement de leur travail. Je considère, pour ma part, qu'elle est éminemment concrète et qu'elle s'adapte à chacun. Ce n'est pas une philosophie, mais une affaire de tous les jours. C'est la raison pour laquelle le consentement dépend de l'histoire de chacun.

Sincérité, loyauté, fiabilité doivent être les maîtres mots du passage du soin à la recherche.

Comment l'information peut-elle participer à un consentement libre et éclairé et contribuer à une recherche éthique ?

Il est important que les protocoles d'essais cliniques soient bien construits et adaptés aux malades. Par ailleurs, les bases de données constituent des outils indispensables

pour améliorer les connaissances. Aussi est-il essentiel que les patients puissent participer à leur développement, pour nourrir une recherche participative.

Les enjeux d'une recherche participative sont nombreux :

- saisie individuelle par le patient pour des données plus fournies que celles du dossier patient ;
- mise en place et suivi des bases de données ;
- partage des données ;
- accès aux données ;
- collection des échantillons biologiques ;
- recrutement ;
- information sur les essais cliniques ;
- construction et élaboration des essais cliniques ;
- retour des résultats issus de la recherche ;
- information sur les avancées de la donnée.

Une association de patients n'a pas pour unique rôle de faire du recrutement de patients. Elle constitue un véritable atout pour la recherche.

Dans le cadre de ma thèse, j'ai conduit deux études afin de recueillir le ressenti des différents acteurs. J'ai notamment interrogé des patients et leurs proches quant à leur souhait de participer à une recherche avec un recueil de données personnelles cliniques et biologiques visant à créer une base sur les leucodystrophies : 80 % des patients et 86 % leurs proches ont répondu par l'affirmative. J'ai également demandé aux associations de patients quelle

pouvait être leur utilité dans la gestion des registres. Les résultats sont les suivants :

- aider au recrutement (71 %) ;
- aider à définir les données pertinentes à entrer dans le registre (58 %) ;
- élaborer les outils d'information et de consentement (53 %).

De la désignation orpheline au post-AMM, les associations sont présentes dans de très nombreux comités. Cela participe de la confiance du patient, qui constate que son point de vue du patient est pris en compte dès le départ.

Quels enjeux ?

Il présente de nombreux points positifs :

- un partage facilité des données ;
- un retour des résultats par newsletter ;
- une autonomie de choix et, le cas échéant, un re-consentement ou un retrait facilité du candidat ;
- la possibilité de ré-orienter la recherche ;
- l'accès aux données et aux informations ;
- la mise en place d'un comité de suivi avec les associations.

Pour autant, des points négatifs sont également identifiés :

- le risque de fracture sociale liée au numérique ;
- le risque d'un e-consentement standardisé ;
- la difficulté à créer des partenariats associations/chercheurs ;

- la question de la validité du consentement (est-il réellement éclairé ?) ;
- la mise à mal de la singularité de la relation médecin/patient ;
- la difficulté d'écoute ;
- la question de la garantie de la stricte anonymisation des données, *a fortiori* pour les maladies rares.
- l'écoute et l'empathie ;
- la confiance en soi ;
- la non-malfaisance (singularité de la relation de soin, accompagnement de tous les acteurs) ;
- la bienfaisance (qualité des outils, acquisition des compétences, diffusion de l'information à l'entourage).

Les enjeux éthiques du e-consentement sont donc les suivants :

- l'équité et la justice (une juste place pour tous) ;
- l'autonomie ;
- le partage d'expérience ;

Dans ce contexte, il est indispensable de renforcer la formation des associations, pour permettre l'acquisition de connaissances et de compétences et afin que nous soyons tous acteurs de la recherche clinique et du développement du médicament.



Table ronde et échanges avec la salle

- *Dr Guillaume JONDEAU, AP-HP*
- *Rémy CHOQUET, BNDMR*
- *Paulette MORIN, Syndrome de Marfan*
- *Anne-Sophie LAPOINTE, VML*
- *Ana RATH, Orphanet*

Emmanuelle Rial-Sebbag

*Quid des méthodes dites **Adapting Licensing** (thérapies innovantes mises sur le marché par anticipation en vue de les adapter à la « vraie vie ») ? Elles peuvent être vues comme un progrès, mais qu'en est-il de l'évaluation des risques ?*

Anne-Sophie LAPOINTE

L'évaluation des risques par un comité juridique et d'éthique est une question d'importance. C'est l'un de nos champs de réflexion.

Jacques BERNARD, Maladies Rares Info Service

En tant que citoyen, j'ai le sentiment que les malades ont tout intérêt à faire part de leurs données et à accepter qu'elles soient rassemblées : cela peut favoriser la recherche et la clinique, mais aussi la participation des malades aux décisions de santé qui les concernent. Je dirais donc oui au partage des données médicales, mais attention aux utilisations abusives, nocives et pernicieuses.

Guillemette PADOU, Inserm

Comment parvenez-vous à évaluer la qualité des résultats des travaux de recherche au regard des difficultés rencontrées dans l'exploitation des données ? La qualité de l'information

peut varier selon la qualité de son recueil.

Rémy CHOQUET

Le débat porte souvent sur les risques de dérive de la réutilisation du Big Data - sous réserve que cela soit un jour juridiquement permis. Mais la réutilisation effective n'est pas encore une réalité. Nous ne pouvons pas encore appuyer sur ce type de données pour générer de la connaissance et alimenter la recherche. La question est celle du bon calibrage du bon outil pour le bon usage. La bonne pratique tend à mettre en place des outils spécifiques, pour des usages spécifiques.

Dr Guillaume JONDEAU

Le plus souvent, lorsque nous menons une étude clinique, nous ne nous appuyons pas sur notre base de données mais nous effectuons une lecture centralisée, homogène et reproductible de tous les examens relatifs au paramètre étudié. Pour le dire autrement, nous sortons de la base de données pour en créer une autre, spécifique.

Une participante

Où en est l'interopérabilité des bases de données des hôpitaux de l'AP-HP ?

Rémy CHOQUET

L'interopérabilité des données médicales sera rendue possible par la mise en place d'un système unique, en l'occurrence Orbis, qui est en cours de déploiement dans l'AP-HP.

Je tiens à préciser que l'on peut tenter de créer un système référentiel des données de santé, pour standardiser au mieux l'information collectée et colligée, mais ce n'est pas pour autant que ces données seront interprétables dans le cadre de la recherche. Dans le soin, le professionnel de santé peut interpréter l'information. Aujourd'hui, c'est encore très difficile de sortir ces données de leur contexte pour appliquer des méthodes statistiques afin de générer des connaissances. Nous testons plus des hypothèses sur ces entrepôts de données, qu'il faudra valider ensuite plus classiquement.

Dr Guillaume JONDEAU

L'interprétation d'une base de données est très complexe, en effet. D'autant que le sens des données dépend de la finalité de la création d'une base de données.

Rémy CHOQUET

L'exemple de l'ordinateur Watson d'IBM est très intéressant. Il arrive à trouver de l'information à propos de l'état de santé d'un patient dans la bibliographie médicale, c'est un très bon indexeur. Rappelons-nous qu'aujourd'hui l'intelligence artificielle n'existe que dans un cadre prédéfini par l'homme. Les méthodes de traitement de l'information, les statistiques, le datamining sont plus efficaces qu'auparavant, mais elles sont aujourd'hui limitées pour traiter des données interprétées par l'homme dans un contexte non décrit comme les données issues de la

clinique et de dossiers patients qui ne renseignent que des éléments de prise en charge et non le phénotype des patients.

Emmanuelle RIAL-SEBBAG

Avez-vous eu des discussions avec votre Service d'information avant l'adoption de la MR-003 de la Cnil ?

Rémy CHOQUET

La MR-003 permet de faire des recherches sur données à partir d'hypothèses sur des entrepôts de données sans avoir nécessairement identifié en amont la totalité des usages possibles. L'AP-HP a placé les garde-fous nécessaires en termes organisationnels : toute mise en place d'étude ou de pré-étude sur l'entrepôt de données doit se faire avec l'accompagnement des personnes habituées à réutiliser les données du soin, et de façon très encadrée. En outre, cette méthodologie de référence présente l'intérêt d'alléger la procédure juridique de demande d'autorisation avant toute recherche.

La question est avant tout terminologique : qu'est-ce qu'une « recherche » ? Il est difficile de demander à un patient son consentement pour entrer dans une étude de prévalence ou d'incidence, par exemple. D'autant que l'information recherchée servira à tout le monde. La question du bien commun avait d'ailleurs été posée lors de la préparation de la loi de santé, mais elle a finalement été laissée de côté en raison de la difficulté à définir ce concept.

Anne-Sophie LAPOINTE

Comment la BNDMR et Radico pourraient-elles s'articuler entre, et quelle serait la place du

consentement dans cette articulation ?

Rémy CHOQUET

La BNDMR et Radico n'ont pas la même finalité. Radico vise à répondre à des questions très précises, qui nécessitent des données spécifiques, donc un consentement spécifique. Pour sa part, la BNDMR est un système de screening visant à aider les professionnels à constituer des cohortes de sous-populations de patients et à répondre à des questions nationales visant à adapter l'offre de soins à la demande, comme par exemple mieux appréhender la question du temps au diagnostic maladie rare.

Ana RATH

Je coordonne actuellement l'action commune menée par les Etats membres de l'Union européenne pour les maladies rares, dont l'une des missions consiste à aider les futurs réseaux européens de référence à mettre en place une interopérabilité dans le cadre du soin et de la recherche. Dans ce cadre, une *Interoperability Task Force* a été constituée, en partenariat avec l'action commune pour la e-santé. L'un des premiers sujets de travail sera l'élaboration de standards de collecte de données dans le cadre du soin mais aussi des registres. Un autre sujet sera le e-suivi du malade avec une spécificité maladie rare pour la prise en charge en situation d'urgence dans un autre pays que le sien. Ce n'est qu'un pilote pour l'instant, mais ce pourrait être un atout majeur dans le cadre de la directive sur les soins transfrontaliers.

Alain DONNART

Paulette, quelles sont les relations que votre association entretient avec les centres de référence concernant le traitement des informations contenues dans le dossier patient ?

Paulette MORIN

Nous avons participé au financement de la mise en œuvre d'une base de données internationale sur les syndromes de Marfan et apparentés. Notre volonté est que les patients puissent verser directement des informations complémentaires dans leur propre dossier. Cela permettrait de mieux connaître notre maladie.

De manière générale, j'observe la grande qualité des relations entre les chercheurs et les malades. Je ne peux qu'encourager les associations à se rapprocher des équipes de soin et de recherche, notamment lors de l'élaboration des notes d'information et de consentement. Lorsque les associations sont persuadées du bien-fondé de la recherche, elles sont mieux à même de motiver leurs adhérents à entrer dans un protocole.

Caroline ROATTA

Une association de patients qui délivre des données dans le cadre d'un projet de recherche peut-elle accéder aux données de santé des hôpitaux ?

Dr Guillaume JONDEAU

Tout un chacun a accès à ses données médicales (données de soin). Il suffit de les demander. Quant aux données de recherche, vous pouvez demander qu'elles ne soient pas utilisées. Mais je ne suis pas certain que vous y ayez accès.

Anne-Sophie LAPOINTE

Nous suggérons de plus en plus souvent de faire de l'information globale, dans le cadre d'une newsletter annuelle par exemple, pour indiquer à quelle recherche ont servi les données des patients.

Dr Guillaume JONDEAU

Les résultats des recherches que nous avons conduites avec des données des patients sont disponibles dans notre centre de référence.

Rémy CHOQUET

La CNIL a demandé à l'AP-HP de mettre en place une page Web présentant toute demande de recherche sur les données (par qui, comment, pour quelle finalité). L'objectif est que les patients soient informés d'un projet de recherche et puissent s'opposer à l'utilisation de leurs données. C'est une bonne pratique qui doit se généraliser.

De la salle

Nous recevons régulièrement des rapports sur les recherches que nous finançons.

Dr Guillaume JONDEAU

Lorsque nous bénéficions d'un financement, il est de notre devoir de présenter un retour sur investissement.

De la salle

Dans le cadre de l'Open Data, les associations peuvent-elles faire des recherches sur les données auxquelles elles ont contribué ?

Dr Guillaume JONDEAU

Non. Ce n'est pas parce qu'une association nous aide que nous lui transmettons les data de patients.

Rémy CHOQUET

Si les données sont véritablement ouvertes, vous avez la capacité d'y accéder. Mais dans le domaine de la santé, les données intéressantes n'existent pas ou peu en réel Open Data.

De la salle, association Kourir

Nous voudrions connaître quel est le nombre exact d'enfants atteints par notre pathologie. La littérature américaine fait état de 20 000 enfants, mais les professionnels de santé indiquent qu'ils ne seraient en fait que 3 000. Où est la vérité ?

Rémy CHOQUET

Nous essayons de construire la vérité, que nous ne connaissons pas encore. Il ne s'agit pas d'Open Data, mais de résultat sur les données. A l'horizon 2017-2018, nous devrions disposer d'un processus permettant aux associations de patients de poser des questions, après validation de notre comité scientifique, mais également d'accéder à des rapports publics sur les maladies rares en France.

De la salle

Quelles sont les personnes en charge de l'informatisation des données, dans les hôpitaux ? J'imagine qu'en cette période de forts déficits budgétaires, vous devez manquer de moyens.

Dr Guillaume JONDEAU

Dans notre centre de référence, nous avons mis en place un système qui permet d'entrer les données au fur et à mesure de la consultation. Il est très difficile d'avoir des bases de données complètes. Sans compter que la saisie est très chronophage, nécessite une organisation assez lourde et impose un écran entre le médecin et son patient. Nous pouvons aussi confier cette tâche à des attachés de recherche clinique, sous réserve de pouvoir en financer.

Rémy CHOQUET

La charge de travail est colossale. Le simple référencement de 350 000 cas de CEMARA a nécessité dix ans de travail pour plus de 1500 professionnels de santé. La re-labelisation des centres maladies rares (2016-2017) prévoit un mécanisme de participation au recueil de la BNDMR.





SPÉCIFICITÉ DES DONNÉES DE GÉNÉTIQUE

NGS : une nouvelle ère de collecte de données génomiques

Ana Rath

Orphanet

Le génome humain (totalité de l'ADN contenu dans les cellules) compte 20 000 gènes environ. Lire chacun de ces gènes nécessiterait plusieurs centaines d'années. Toutefois, seuls 1,5 % des gènes codent pour des protéines et déterminent des fonctions dans les différentes cellules de notre corps. Ce codage intervient dans des portions appelées « exomes ». Les autres cellules sont silencieuses ou servent à réguler les exomes.

Nous connaissons une véritable accélération en termes de séquençage, car nous sommes désormais en mesure de lire de très nombreuses séquences codantes simultanément et ce, à moindre coût - étant entendu que « lire » ne signifie pas nécessairement « comprendre ».

Gènes et maladies

La base Orphanet compte aujourd'hui 6 000 maladies, dont 2 740 (72 %) sont des maladies génétiques (ce chiffre s'actualisant très rapidement), pour lesquelles la causalité de 3 179 gènes (dont certains locus et séquences non-codantes) a été prouvée. Cela signifie que les causes de près de 40 % des maladies génétiques connues n'ont pas encore

été découvertes. Le manque de connaissances est plus grand encore qu'il n'y paraît, dans la mesure où certains cas de maladies dont la causalité génétique a été découverte restent inexplicables. Il reste donc une grande proportion de malades pour lesquels un diagnostic génétique ne peut pas être affirmé, comme nous l'avons vu tout à l'heure concernant les syndromes Marfan et apparentés.

La plupart des gènes causaux d'une maladie rare sont des gènes codants. Cela signifie qu'un *whole exome sequencing* (WES- séquençage de l'exome) permettrait dans la plupart des cas de les trouver. Et pour les gènes non-codants, un *whole genome sequencing* (WGS- séquençage du génome) permettrait sans doute de poser le diagnostic. Ce séquençage complet requiert plus de temps et coûte encore très cher. En tout état de cause, certains généticiens pensent qu'il serait intéressant d'en faire le premier examen.

Un accroissement sensible des découvertes de gènes causaux a été rendu possible à compter de 2012, avec l'utilisation du NGS (next-generation sequencing) en plus des techniques classiques. En parallèle, les coûts du séquençage ont significativement chuté : il coûte désormais moins cher de commencer par du *whole genome sequencing* ou du *whole exome sequencing* que de

passer par toute une batterie d'investigations (imagerie, etc.). La littérature considère que nous aurions gagné six ans dans l'histoire de la maladie si nous avions commencé par du séquençage du génome ou de l'exome, ce qui aurait permis de poser un diagnostic et de cibler les autres investigations à effectuer - sous réserve que le gène soit connu, bien sûr.

Séquençage : de l'échantillon à l'interprétation

Le séquençage du génome est très complexe. Ce n'est pas non plus une simple question de technique. Il ne suffit pas d'entrer un échantillon sanguin dans une machine pour connaître le diagnostic.

Aucune machine ne peut directement lire les brins d'ADN. Ces derniers doivent être préalablement découpés, réassemblés en fonction des séquences de référence et annotés. Ce marquage des gènes (pour indiquer où ils commencent et où ils se terminent) peut se faire par routine automatique, mais une annotation manuelle finale reste indispensable.

Enfin, les résultats doivent être interprétés, ce qui impose de comparer la séquence du malade avec une séquence de référence ou des séquences de parents ou apparentés, afin d'identifier les écarts. Il peut s'agir d'une mutation pathogène, mais aussi de nombreuses variations par rapport à la séquence de référence (si elle existe), qui ne peuvent pas toujours être interprétées en fonction des connaissances. Il faut donc du temps pour comprendre la signification clinique de l'intégralité du génome humain : ce n'est pas parce qu'on

peut lire chaque gène que l'on peut tout comprendre.

Par ailleurs, il est possible de faire du séquençage dans le cadre du diagnostic comme dans celui de la recherche. Dans le premier cas, l'évolution des connaissances permettra peut-être d'interpréter ultérieurement certains résultats - ce qui pose le problème de la génération gigantesque de données génomiques : où les conserver (Google, Amazon et autres proposent déjà des capacités de stockage sur lesquelles les pouvoirs publics peineront à s'aligner), lesquelles conserver, pour combien de temps et dans quel but ? Les résultats d'une recherche pourraient permettre d'améliorer le diagnostic, si ce n'est pour le malade initial, au moins pour les générations futures. Dès lors, comment dés-identifier les données tout en permettant de les ré-identifier ultérieurement pour interpréter un cas particulier ?

Enfin, pour accroître les connaissances sur les maladies génétiques, il est indispensable d'interopérer à l'international. L'échelle d'un centre et même l'échelle nationale sont insuffisantes. Seule une échelle internationale permet d'avancer infiniment plus vite, sous réserve de parler le même langage.

Interopérabilité

Des collaborations internationales se mettent en place à l'échelle internationale, grâce notamment au Matchmaker Exchange. L'objectif est de disposer d'une base de données collectant des données ou des méta-données, pour permettre à toutes les autres bases d'interopérer ou, *a minima*, d'être découvrables.

En France, nous sommes un peu en retard dans le domaine du séquençage massif du génome. Plusieurs instances se sont saisies de ce sujet et ont rédigé un rapport intitulé « Plan France Médecine génomique 2025 ». Ce plan vise notamment la problématique du séquençage du côté du soin. En l'occurrence, il suggère d'inscrire le séquençage dans le parcours de soins, pour lier les données de génétique et les données cliniques - même si nous ne savons pas encore vraiment comment cela se mettra en place.

Trois projets pilotes ont été proposés, donc l'un porte sur les maladies rares. Il conviendra de l'intégrer au troisième Plan national maladies rares. L'objectif est d'avoir 60 000 génomes (20 000 trios, le patient et ses parents) à horizon 2020 pour amorcer la génomique à grande échelle, puis de connecter cette base de données à l'international (ce qui ne va pas sans poser de très importants problèmes en termes de partage de données) afin d'accélérer le développement des connaissances.

Lever les verrous technologiques est indispensable pour faire évoluer les cadres réglementaire et éthique.





Données de génétique et réseaux sociaux

Denis COSTELLO
RareConnect/EURORDIS

J'ai déjà eu l'occasion de vous présenter RareConnect, le réseau social pour les associations de patients atteints d'une maladie rare. Nous dénombrons 800 associations partenaires, 90 communautés, un réseau de 300 modérateurs formés et 100 000 visiteurs par mois. L'animation est assurée par une équipe de cinq personnes, chargées de garantir la qualité des informations publiées et de modérer les échanges.

Vers une communauté mondiale

La problématique que prétend résoudre le numérique est la « queue longue » (*long tail*), c'est-à-dire le « très rare » (<1/2000 en Union européenne). Je considère que les réseaux sociaux ont un véritable rôle à jouer, ne serait-ce que pour identifier les patients, un peu à l'instar d'Amazon qui réalise la grande partie de son bénéfice en vendant des livres assez rares, tandis que les librairies physiques le font grâce à la vente de best-sellers.

Il est primordial de trouver des échantillons au niveau international (surtout lorsqu'une seule famille est concernée par une maladie rare dans un pays) pour pouvoir identifier les gènes causaux. Si un médecin a rencontré un profil de phénotype ou même s'il dispose du génome séquencé, ce n'est pas suffisant - surtout s'il n'a jamais rencontré un

tel profil avant. D'où l'importance du *Datamining* (fouille de données) et du *Datasharing* (partage de données), donc du rôle des réseaux sociaux.

L'objectif est de faire émerger une communauté mondiale pour une maladie donnée. D'où l'importance du phénotypage pour :

- décrire les caractéristiques de la maladie ;
- distinguer différents syndromes (parfois similaires en apparence) ;
- établir des corrélations génotype/phénotype.

Données génétiques et diagnostic?

Une véritable révolution a été permise par le HPO (Human Phenotype Ontology), l'ontologie développée à Berlin par Peter Robinson qui permet une description à la fois très fine et très uniforme des caractéristiques de la maladie. Cela permet aux ordinateurs de travailler sur l'ontologie. Si deux patients présentent un profil similaire, les algorithmes peuvent calculer ce qu'ils ont en commun ainsi que la distance entre leurs deux cas (en pourcentage).

L'algorithme « Exomizer », également développé par Peter Robinson, repose sur l'analyse des corrélations entre les phénotypes de la souris et les phénotypes humains. Il permet d'identifier les similitudes

entre deux cas particuliers. Les premiers tests ont permis d'augmenter par trois les possibilités de diagnostic, même s'il reste encore un grand nombre de faux positifs.

Par ailleurs, chaque hôpital peut déployer le réseau PhenomeCentral, outil informatique en open source permettant de décrire le phénotype de leurs patients de manière anonyme, de calculer la distance entre deux cas et de notifier les médecins dont les patients ont le même profil. Il est déjà déployé très largement au Canada et il commence à l'être en Europe.

Il convient en outre de noter que RareConnect a signé un partenariat avec l'équipe de PhenomeCentral, pour développer une version qui permettrait aux patients eux-mêmes de renseigner leur propre phénotype et de décrire leur maladie, puis de le faire valider par leur généticien. De nombreux patients ont été inclus dans ce projet, de façon à ce que l'outil réponde aux besoins de leur vie quotidienne. Nous ne sommes pas les seuls à nous lancer dans cette démarche, même si la plupart des initiatives viennent des États-Unis.

Nous pourrions aussi recourir aux moteurs de recherche, pour mieux identifier les associations de patients dont l'expertise est réelle dans telle ou telle maladie, ou pour le recrutement pour les essais cliniques.

Enfin, il serait également intéressant de nourrir des discussions transversales, selon le modèle des réseaux sociaux.





Découvertes fortuites lors du séquençage

Anne-Sophie LAPOINTE
VML

Les découvertes fortuites sont aussi appelées « données non sollicitées ». Elles posent, elles aussi, la question du consentement.

Deux paramètres sont clés dans l'articulation des intérêts collectifs et individuels, pour protéger les personnes.

- L'information de la personne
Le patient doit être systématiquement prévenu de toute anomalie (autonomie absolue), mais aussi de la découverte d'une information qui pourrait s'avérer pertinente (autonomie régulée). En revanche, il ne doit être prévenu qu'à condition que la pertinence de l'information soit fortement validée et qu'une prise en charge médicale soit très souhaitable (afin d'éviter une perte de chance significative). Enfin, il peut ne jamais être prévenu, ce qui pose la question délicate de la responsabilité, mais aussi celle de l'information à la parentèle.
- La pertinence médicale de données
La pertinence des informations doit être validée par l'investigateur principal, un comité de pilotage, un comité de protection des personnes, un expert désigné *a priori* et

une structure collégiale *ad hoc*, avec une liste préétablie des différents membres. Toutefois, la pertinence d'une donnée pouvant évoluer de façon très forte et rapide, le comité éthique de l'Inserm n'est pas favorable à une liste préétablie. Par ailleurs, la pertinence médicale des données, c'est une validité analytique, une validité clinique, une utilité clinique et une ELSI. La problématique des faux positifs est réelle, dans ce domaine.

Etat des lieux de la réflexion du groupe de travail « tests génétiques : découvertes annexes et inattendues de la recherche » (2015-2016)

Les études dans lesquelles les incidentalomes (découvertes fortuites, données non sollicitées) sont attendus doivent-elles avoir des consentements différents des autres ? Il est très difficile de consentir en amont sans savoir sur quoi, d'autant qu'avec le séquençage à haut débit, il est possible de trouver une énorme masse d'informations. D'où l'importance de prévoir un consentement expliquant clairement le processus des tests génétiques et le risque de découvertes annexes et inattendues.

Un groupe de travail a été constitué pour étudier les points de tension, les conseils à donner et les outils à mettre en place. Il travaille en lien avec des équipes de recherche, des

associations de patients et des institutions comme l'Inserm. L'objectif est de développer une filière médecine génomique, mais aussi une filière industrielle.

Quelles perspectives ?

La solution résiderait peut-être dans la création d'un portail, pour faire le lien entre le génotype et le phénotype, c'est-à-dire rassembler toutes les données non sollicitées que l'on ne connaît pas. Le rapport France Médecine génomique 2025 évoque la piste d'un centre de référence technologique, d'innovation, d'expertise et de transfert.

Par ailleurs, dans le domaine des maladies rares, nous ne pourrions pas nous affranchir de l'idée de travailler à l'international. Les futurs réseaux européens de référence seront de véritables accélérateurs pour le partage des données et l'amélioration des connaissances.

Les questions restent nombreuses. Concernant le séquençage, par exemple, le consentement est demandé à un individu, mais nous savons que toute la famille est concernée lorsqu'il s'agit d'une maladie génétique. Un véritable accompagnement est nécessaire. Il passera par une plus grande communication, *via* des flyers ou des webinars par exemple. Dans cette optique, la formation des associations est indispensable, pour comprendre les enjeux du séquençage à haut débit et savoir expliquer la pertinence de telle ou telle découverte. Par ailleurs, l'article 16-10 du Code civil indique que « *le consentement mentionne la finalité de l'examen* ». Or la finalité de l'examen n'est pas connue, pour le type de données dont nous parlons.

Il convient de distinguer les découvertes secondaires (issues de l'hypothèse de départ) des découvertes non sollicitées (auxquelles l'on ne s'attend pas du tout). C'est le dilemme entre soin et recherche, avec d'un côté le juridique et, de l'autre, l'éthique l'accompagnement et la bienfaisance.





Table ronde et clôture du Forum

- *Ana Rath, Orphanet*
- *Denis Costello, Eurordis*
- *Anne-Sophie LAPOINTE, VML*

De la salle

Certaines découvertes fortuites interviennent dans le cadre du diagnostic prénatal également, lors de la recherche d'une trisomie 21 par exemple. Il peut s'agir d'une découverte concernant une anomalie chromosomique rare et pathogène. Or lorsqu'ils signent leur consentement pour un diagnostic prénatal, les parents ne sont pas vraiment informés que le médecin pourra trouver autre chose que la trisomie 21.

Anne-Sophie LAPOINTE

Le consentement adapté aux découvertes non sollicitées reste encore à construire. Cette question sera traitée par la Filière.

Ana RATH

Dans le cadre d'une amniocentèse, on recherche toutes les anomalies chromosomiques - dont la trisomie 21 est la plus fréquente. L'information des parents est peut-être insuffisante, dans certains cas.

Les découvertes non sollicitées sont un peu différentes : elles portent sur des anomalies qui n'étaient pas recherchées.

Paulette MORIN

L'agence de Biomédecine a demandé au comité « Ethique et société » de l'Alliance Maladies Rares de participer à la rédaction de textes à sortir sur le diagnostic prénatal (DPN) et le diagnostic pré-implantatoire (DPI). Il

existera différents consentements (je ne veux rien savoir donc on ne fait pas l'analyse, je veux savoir sous telles conditions, etc.), qui seront très clairement expliqués.

De la salle

Nous avons aussi le droit de ne pas savoir, en effet.

De la salle

On ne pense pas à demander le consentement des parents en fonction de ce qui sera découvert à l'occasion d'un test génétique sur un de leurs enfants.

Ana RATH

Vous posez là le problème de l'annonce du diagnostic, indépendamment du séquençage.

Anne-Sophie LAPOINTE

Je rappelle que toute recherche génétique requiert obligatoirement un consentement spécifique.

Alain DONNART

L'annonce à la fratrie est très complexe. L'Alliance Maladies Rares souhaiterait, à cet égard, qu'un psychologue soit systématiquement présent au moment de l'annonce du diagnostic.

De la salle

D'où l'importance du rôle des associations en termes d'accompagnement et de proximité avec les malades.

Alain DONNART

En effet. Les associations font de l'écoute permanente.

Paulette MORIN

Le généticien qui fait le prélèvement doit fournir la liste des associations de patients agréées, c'est une obligation légale. Apporter de l'aide,

de la connaissance et du soutien est pleinement notre rôle

Ana RATH

Je vous remercie. J'espère que ce Forum vous aura été utile et, surtout, vous permettra de vous rapprocher davantage encore les uns des autres !



Crédits photo : Sophie Höhn (Orphanet), Marc Hanauer (Orphanet)