

# Données de santé dans les maladies rares et séquençage génétique : les malades s'interrogent.

*Note de réflexion du Comité éthique et société de l'Alliance Maladies Rares - 2017*

Au cours de l'année 2016/2017, le Comité Éthique et Société\* de l'Alliance Maladies Rares a mené une réflexion, à la demande du Comité national de l'Alliance, sur la problématique suivante : quels sont les enjeux, les opportunités et les risques (du point de vue des malades, des familles et de leurs associations) au point de vue éthique et sociétal, du développement des données de santé\* sur les maladies rares, notamment par le séquençage génétique ?

Celles-ci, y compris leurs données génétiques, s'incluent dans le vaste ensemble des big-data\*. Numérisées, elles sont gérées par des opérateurs multiples et mondialisés, comme Google. Toutes ces données, directement fournies par les malades, constituent un matériel précieux pour la recherche et ouvrent de nouvelles perspectives thérapeutiques.

Mais, particulièrement sensibles car touchant à l'intimité des personnes, elles peuvent donner lieu à des utilisations inappropriées ou malveillantes susceptibles de conséquences graves pour la vie privée. Des intérêts financiers énormes sont en jeu.

Parmi ces données, celles concernant les caractéristiques génétiques sont particulièrement sensibles. Or le séquençage à haut débit qui se développe rapidement et représente un espoir pour bien des familles pose des questions inédites : en raison même de sa puissance à traiter les données, il est en mesure de révéler aux malades des pathologies incidentes qui peuvent s'avérer d'autant plus perturbantes qu'elles ont été inattendues.

Les préoccupations des malades tiennent aux nombreuses questions nouvelles qu'ils se posent. Celles-ci se regroupent autour de trois thématiques sensibles : la première sur la portée réelle du consentement éclairé et de l'information préalable dus aux malades, la seconde sur la protection de leurs données et sur l'absence de garantie de l'anonymisation et la dernière sur le respect de leur volonté dans le rendu des résultats du séquençage à haut débit\* de leur génome.

## I. LES ENJEUX

La production massive de données de santé et les données fournies par le séquençage génétique, notamment le séquençage à haut débit vont permettre des progrès dans la connaissance et le traitement des maladies.

Mais elles font surgir, dans le même mouvement, la nécessité de protéger la vie privée et d'empêcher une utilisation malveillante ou inappropriée des données médicales, notamment génétiques fournies par les malades dans le cadre du soin ou de la recherche.

Le développement très rapide des technologies et les intérêts financiers qui les accompagnent ne doit pas prendre le pas sur la réflexion et la définition de bonnes pratiques.

\* Voir en annexe

Il y a là un enjeu pour les associations de malades qui doivent être très présentes sur ces sujets.

Le défi est d'aller vers une médecine davantage préventive, prédictive, individualisée et peut-être économe de moyens, le malade étant placé au cœur du dispositif et devenant davantage acteur.

L'enjeu fondamental est donc l'ouverture de nouvelles perspectives de guérison dans le respect des droits des malades et des citoyens.

## 2. LES OPPORTUNITÉS

La production massive et numérisée de données de santé y compris celles provenant du séquençage génétique est installée dans le paysage et avance très vite.

Cela crée des opportunités positives qu'il faut saisir, tout en étant vigilants sur les dérives qui peuvent exister.

En effet ces nouvelles possibilités permettent, au profit des malades :

- de surmonter l'obstacle structurel du faible nombre de malades et de leur dispersion sur toute la surface du globe. C'est un intérêt majeur pour les maladies rares.

- d'élargir le champ et le nombre des données qui permettent de mieux connaître les maladies : données cliniques, données génétiques et lien entre ces domaines. Soulignons ici que la composante génétique est plus importante que dans les autres pathologies (80 % des maladies rares ont une composante génétique).

- de connaître des informations concernant la parentèle\* ce qui élargit la dimension de santé publique des maladies rares. Plusieurs études et initiatives associatives montrent que les malades sont favorables à fournir leurs données de santé mais sont soucieux de ne pas en subir pour eux-mêmes et leurs familles des conséquences négatives pour leur qualité de vie et leur dignité.

Il faut donc s'engager dans ces nouvelles techniques mais le faire avec prudence pour en tirer le meilleur bénéfice possible au niveau collectif comme au niveau individuel.

## 3. LES PRÉOCCUPATIONS

Les données de santé sont reconnues comme données « sensibles » et protégées par des lois dont les plus récentes sont :

- la loi du 5 mars 2012 (Loi Jardé\*) relative aux recherches impliquant la personne humaine,

- la loi de santé publique du 26 janvier 2016 sur le système national des données de santé et les interdictions de promotion commerciale de produits de santé et de sélection du risque par les assurances\*

- et le règlement européen 2016/679 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.\*

L'intensité de l'activité juridique concernant les données personnelles et les données de santé témoigne des efforts

\* Voir en annexe

d'adaptation des textes à l'évolution des technologies numériques et à leurs usages.

En complément de cet encadrement général, la législation française sur les tests génétiques est particulièrement stricte et impose des obligations pour tous les acteurs y compris les malades. Mais ce cadre qui prévoit toutes les étapes du processus, à savoir les règles de prescription, les personnes concernées, les laboratoires ainsi que les modalités d'information et du consentement éclairé des malades, est susceptible d'être mis en tension par le séquençage à haut débit.

Ces tensions sont renforcées par la porosité croissante des frontières entre la clinique et la recherche et par l'évolution même des pratiques de recherche, de plus en plus centrées sur le traitement des données. La production massive de données par le séquençage à haut débit entraîne un accroissement du nombre des bases de données\* et des banques de ressources biologiques. La circulation, le partage, l'analyse de ces données génèrent de nouvelles approches qui vont modifier en profondeur la recherche et la clinique génétique.

## 4. L'INFORMATION ET LE CONSENTEMENT DES MALADES MIS EN TENSION

Face à ces changements, les préoccupations des malades concernent le recueil/transmission des données, leur protection et le rendu des résultats.

### 4-1 - LA FOURNITURE DES DONNÉES ET LE CONSENTEMENT

Qu'il s'agisse du soin, de la recherche ou du diagnostic génétique, les malades transmettent leurs données de santé (données personnelles, imagerie ou données issues d'exams biologiques) qui s'accompagnent parfois de prélèvements ou de sang (ressources biologiques). Les modalités de recueil ont été récemment mises à jour par la loi Jardé. Elles sont différentes selon que l'on se situe dans le cadre du soin (recherches non interventionnelles ou interventionnelles à risque minime) ou dans le cas de la recherche interventionnelle. Dans le cas du soin (recherches non interventionnelles), le consentement explicite n'est pas requis, il suffit de ne pas s'opposer à l'intervention, dans le cas du soin, interventionnel à risque minime, le consentement requis est simple, il peut être écrit ou oral. Dans celui de la recherche il doit être écrit. Par ailleurs tous les exams de caractéristiques génétiques requièrent un consentement écrit.

Le séquençage génétique à haut débit crée une situation

nouvelle pour le malade, qui donne tout ou partie de son génome, si bien que le patrimoine génétique qu'il donne ne se limite plus à sa seule pathologie et peut révéler d'autres problèmes de santé indépendants de sa pathologie d'origine. Les malades prennent conscience que ces évolutions issues notamment des technologies les conduisent à réexaminer un certain nombre de questions pour eux essentielles sur le consentement libre et éclairé :

- À quoi est-ce que je donne mon consentement ?
- Que se passe-t-il si mes données sont utiles à d'autres finalités de recherche ? Revient-on vers moi ?
- Dans le cas de séquençage génétique de mon génome entier, combien de temps mon consentement est-il valable ? Puis-je refuser ou revenir sur mon consentement et m'opposer à l'utilisation de mes données personnelles ? Pratiquement comment mes données personnelles peuvent-elles être retirées du circuit dans lequel elles ont été intégrées ? Quelles conséquences tant personnelles que générales ce droit de retrait risque-t-il d'avoir ?

## ■ 4-2 - LA PROTECTION DES DONNÉES

### → le circuit des données

- ↳ Où sont-elles stockées ? Dans des banques de ressources biologiques et dans des bases de données qui sont installées sur des hébergeurs publics ou privés ou encore dans les laboratoires qui ont les résultats des diagnostics génétiques ? Quelles garanties de confidentialité présentent ces hébergeurs ?

Un Web du génome a émergé. 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. La plus grande banque de données est Google (23 and Me, une société qui propose des analyses de code génétique aux internautes) qui dispose de plusieurs millions de génomes, sans qu'on sache quel usage il pourrait en faire. Le stockage de nos ADN dans les « clouds » a commencé. Google et Amazon proposent aujourd'hui d'héberger gratuitement des bases de données génomiques réputées comme « 1000 Genomes Project ».

Les réseaux sociaux comme Facebook, propriétaires de leurs données, disposent de nombreuses informations à partir des éléments fournis et échangés par les malades eux-mêmes.

- ↳ Combien de temps ces données sont-elles conservées ?
- ↳ Existe-t-il un droit à l'oubli ?
- ↳ Où sont transférées les données ?

L'hyper connexion leur permet de voyager, y compris dans des circuits commerciaux ou dans des pays où leur protection n'est pas garantie. Des accords commerciaux ont été conclus récemment entre ces banques et des laboratoires pharmaceutiques. L'exemple des plates-formes de malades (exemple de la SLA ou de la dépression profonde) montre avec quelle facilité les données peuvent circuler. Ces plates-formes, souvent très appréciées par des malades en détresse, permettent aux promoteurs de ces plateformes (des labs, des chercheurs, des associations) de recruter des malades, de collecter un grand nombre de données de santé livrées par les malades eux-mêmes et de les mettre en lien parfois directement avec des laboratoires et des centres de soins...

### → l'accès aux données, secret médical et anonymisation

- ↳ Puis-je accéder à toutes les données me concernant ? Les données sont-elles traitées dans le respect du secret médical ? Sont-elles, si nécessaire, codées et anonymisées ?

L'anonymisation semble ne pas pouvoir être irréversible. Lorsque dans un pays il existe moins de 20 malades d'une pathologie, l'anonymisation est impossible. L'absence de garanties sur le caractère irréversible de l'anonymisation est une source d'inquiétude majeure pour les malades.

Les règles et sanctions concernant le secret médical et maintenant paramédical et son partage sont stricts. Mais quid de leur application ?

La question du piratage des données de santé s'inscrit continuellement dans l'actualité : des données médicales détournées par des hôpitaux ont été piratées et ont fait l'objet de demandes de rançons. Le phénomène des hackers s'est répandu et ils sont habiles jusqu'à décrypter des codages et des systèmes d'anonymisation. Ceci n'est pas spécifique aux données de santé ni aux maladies rares, mais plus lourd de conséquences dans nos situations (assurances, banques, employeurs...).

## ■ 4-3 LE RENDU DES RÉSULTATS DES DONNÉES FOURNIES

Le séquençage à haut débit est une révolution technologique qui fournit rapidement, à des coûts matériels (séquenceurs et autres) qui ne cessent de baisser, une foule de résultats génétiques susceptibles de donner des pistes thérapeutiques pour une pathologie donnée mais aussi de révéler des pathologies inattendues (résultats incidents). Plusieurs études quantitatives et qualitatives en cours ont pour objet l'impact

des modalités d'information et d'annonce des résultats et l'analyse des attentes et des réactions des malades dans les handicaps rares et les déficiences intellectuelles.

Pour l'heure les questions tournent, pour le généticien, autour de la manière de gérer les données, de les interpréter et d'informer sur les résultats, dire ou ne pas dire, quoi dire comment dire ? Pour le malade, elles ont trait symétriquement autour de l'incertitude des résultats fournis (en général des prédispositions, avec tel ou tel pourcentage de risques) et autour de la question du « savoir ou ne pas savoir ». Le problème se pose de manière particulièrement aigüe dans le cas de résultats incidents qui révèlent des pathologies inattendues (ex prédispositions fortes à des cancers, maladies cardiaques...). La difficulté est encore plus grande lorsque cela concerne des enfants, car ces pathologies peuvent ne se révéler qu'à l'âge adulte, un âge où ils seraient normalement capables de décider eux-mêmes s'ils veulent savoir ou ne pas savoir. De plus ces résultats peuvent intéresser la parentèle.

Des recommandations de bonnes pratiques n'existent pas encore en France. Actuellement, les généticiens se réfèrent à une liste établie aux Etats unis qui identifient 59 gènes dont les variants (variations dans la séquence de l'ADN) ont des conséquences pathogènes, graves, suffisamment certaines, et « actionnables » (c'est-à-dire dont la connaissance permet de mettre en place une prévention ou des traitements) pour informer le patient. Cette liste est par nature évolutive. Les variants qui ne répondent pas à ces critères d'actionnabilité ne sont pas communiqués au malade. A noter cependant que les critères d'actionnabilité évoluent avec le temps. Il est essentiel en conséquence d'informer le malade,

lorsqu'il donne son accord pour un séquençage de son génome, que des résultats incidents, c'est-à-dire des variants insoupçonnés et non liés à sa maladie, mais potentiellement graves peuvent être découverts. Il importe de lui demander dès l'étape du consentement s'il veut connaître ou non les résultats incidents.

- Plusieurs positions peuvent être exprimées par le malade :*
- 1. Tout savoir** : Le malade veut être informé systématiquement de toute anomalie, soupçonnée ou certaine, décelée lors du séquençage. Cette position entraînerait de la part des généticiens une exigence de tout dire, ce qui peut converger avec une volonté de se protéger juridiquement en informant systématiquement au risque de nuire et d'angoisser.
  - 2. Le malade ne veut être informé que si la variation décelée est certaine, pathogène et actionnable.**
  - 3. Le malade ne veut pas être informé des résultats du séquençage de son génome entier mais ne connaître seulement que les résultats concernant sa maladie.**

Quelle que soit la volonté exprimée des malades de **tout savoir ou ne savoir** qu'en fonction du caractère actionnable du variant, se pose la question du retour vers eux, au fur et à mesure de l'évolution des connaissances, ce qui implique une nouvelle organisation de l'information des malades

Sur le plan éthique, il existe des tensions entre plusieurs exigences : le droit de savoir ou non du patient, le devoir du médecin de ne pas nuire au patient (ne pas lui donner des informations inutiles ou angoissantes) et le devoir de soin (si le malade court un risque grave).

## CONCLUSIONS : ENTRE ATTENTE ET RÉSERVE

Selon leurs pathologies, l'existence ou non de traitement et de suivi médical, ou tout simplement le fardeau que représente leur maladie et la crainte d'apprendre qu'ils en ont une autre, amènent chez certains une réserve sinon un refus de faire séquençer leur génome entier. D'autres au contraire, notamment ceux qui sont en errance diagnostique sont très en attente des résultats que leur apporteront le séquençage. Il existe ainsi une hétérogénéité des attitudes par rapport à l'intérêt pour le malade du séquençage du génome entier.

La question est posée également de savoir si la génétique tiendra ses promesses. On met en place d'énormes dispositifs dont on ne connaît pas encore vraiment les bienfaits.

*Ces dispositifs ne vont-ils pas diminuer les efforts et les ressources pour l'accompagnement et les soins ?  
Aura-t-on les inconvénients sans avoir les avantages ? Faut-il se résigner à donner sans retour, faut-il accepter de perdre ?*

Cette hétérogénéité des attentes individuelles des malades renforce l'exigence du respect de leur choix, et l'importance d'organiser leur information en définissant les moyens et les modalités de celle-ci.

[www.alliance-maladies-rares.org](http://www.alliance-maladies-rares.org)

## ANNEXE

### Le Comité Éthique et Société



En 2010, l'Alliance créait une Commission Ethique pour mener une réflexion sur les différentes étapes du parcours de vie de la personne atteinte de maladie rare. La 1ère Journée de réflexion Éthique du 24 janvier 2014 était orientée sur l'accompagnement des associations dans l'annonce du diagnostic et dans l'écoute. En lien avec cette question, la Commission Éthique a proposé, le 24 janvier 2015, une réflexion sur la place de l'éthique dans le désir de parentalité des personnes atteintes ou concernées au sein de leur famille par une maladie rare.

### Le Comité Éthique et Société est composé de 12 membres :

Badia ALLARD; Françoise ANTONINI, personnalité qualifiée; Catherine AVANZINI; Juliette DIEUSAERT; Dominique DONNET-KAMEL; Alain FONTAINE; Anne-Sophie LAPOINTE; Paulette MORIN; Marie-Christine de La MORLAIS; Bernadette ROUSSILLE; Ginette VOLF; Viviane VIOLLET, Responsable du Comité éthique et société;

### Définitions :

#### Données de santé :

Le nouveau règlement européen (mai 2016) relatif à la protection des données personnelles introduit une définition des données de santé : «données à caractère personnel relatives à la santé physique ou mentale d'une personne physique, y compris la prestation de services de soins de santé, qui révèlent des informations sur l'état de santé de cette personne». Il précise qu'elles comprennent «toute information concernant, par exemple, une maladie, un handicap, un risque de maladie, un dossier médical, un traitement clinique ou l'état physiologique ou biomédical de la personne concernée, indépendamment de sa source, qu'elle provienne par exemple d'un médecin ou d'un autre professionnel de la santé, d'un hôpital, d'un dispositif médical ou d'un test de diagnostic in vitro».

**Big data** signifie volumes massifs de données. Ces données proviennent de partout : de capteurs utilisés pour collecter les informations climatiques, de messages sur les sites de médias sociaux, d'images numériques et de vidéos publiées en ligne, d'enregistrements transactionnels d'achats en ligne et de signaux GPS de téléphones mobiles, pour ne citer que quelques sources. Le séquençage à haut débit génère des quantités très importantes de données qui se chiffrent en teraoctets. Ils participent des big data qui requièrent des volumes de stockage considérables, une puissance informatique inégalée et la conception de logiciels/algorithmes appropriés. <https://www.cnil.fr/fr/definition/big-data>

**Une base de données**, est un ensemble structuré et organisé permettant le stockage de grandes quantités d'informations afin d'en faciliter leur utilisation (ajout, mise à jour, recherche et éventuellement analyse)

**Séquençage** : lecture de l'enchaînement linéaire des bases (A, T, G ou C ) composant la séquence d'un fragment d'ADN (l'ADN humain contient 6 milliards de base). Le séquençage comprend trois étapes, l'acquisition de la séquence, l'assemblage et l'annotation/interprétation.

**Séquençage à haut débit** permet d'acquérir la séquence du génome entier, de l'exome ou d'une sélection de gènes (panels de gènes). Les techniques de séquençage à haut débit s'appliquent à l'ADN du génome humain, mais aussi à l'ARN, à l'ADN des bactéries du microbiote intestinalect. L'évolution des technologies amplifie les capacités d'acquisition des séquences (analyse, en simultané et en un temps très court, de millions de fragments d'ADN appartenant aux génomes de plusieurs individus), de leur assemblage par les technologies de la bio-informatique et de l'annotation pour rechercher les informations pertinentes dans les séquences obtenues.

[http://www.anddi-rares.org/assets/files/plaquette-sequencage\\_haut\\_debit\\_exome.pdf](http://www.anddi-rares.org/assets/files/plaquette-sequencage_haut_debit_exome.pdf)

vidéo : <https://www.youtube.com/watch?v=fZQfpE67pcl>

**Séquençage du génome entier** : séquençage de toutes les bases de l'ADN, l'ADN du génome humain contient 6 milliards de bases.

**Séquençage de l'exome** : lecture de l'enchaînement des bases qui constituent l'exome, c'est-à-dire l'ensemble des exons présents dans les 23 000 gènes de l'ADN humain. A noter que les exons sont la partie dite codante des gènes et ne représentent que 1% de tout le génome (le reste de l'ADN étant formé de séquences régulatrices, de séquences répétées, d'introns et de séquences dont on ne connaît pas encore la fonction)

**Séquençage de panels de gènes** : seuls les gènes, connus ou suspectés comme pouvant être impliqués dans une pathologie, sont séquencés.

## Bibliographie

### CNIL :

Commission Nationale de l'informatique et des Libertés – Loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Article 2 - « La présente loi s'applique aux traitements automatisés de données à caractère personnel, ainsi qu'aux traitements non automatisés de données à caractère personnel contenues ou appelées à figurer dans des fichiers, ... ».  
<https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460>

La CNIL a publié 3 méthodologies de référence relative aux traitements de données à caractère personnel mis en œuvre dans le cadre des recherches dans le domaine de la santé :

MR-001 : Recherche interventionnelle nécessitant le consentement exprès ou écrit du participant :

<https://www.cnil.fr/sites/default/files/atoms/files/mr-001.pdf>

MR-002 : Recherche non interventionnelle dans le domaine des dispositifs médicaux de diagnostic in vitro.

MR-003 : recherche observationnelle ne nécessitant pas le consentement exprès ou écrit du participant.

<https://www.cnil.fr/sites/default/files/atoms/files/mr-003.pdf>

### Mise en place d'un nouveau fichier – TES :

Titres électroniques sécurisés (identité, domicile, couleur des yeux, photo, empreintes digitales...) avec pour objectif la simplification administrative. [http://www.senat.fr/espace\\_presse/actualites/201611/fichier\\_tes.html](http://www.senat.fr/espace_presse/actualites/201611/fichier_tes.html)

### Loi de santé publique du 26 janvier 2016 sur le système national des données de santé et les interdictions de promotion commerciale de produits de santé et de sélection du risque par les assurances :

<https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000031912641&categorieLien=id>

### Règlement Général sur la protection des données :

Règlement (UE) 2016/679 du parlement européen et du conseil 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 et abrogeant la directive 95/46/CE (Règlement général sur la protection des données (RGPD – General Data Protection Regulation)

Entrée en vigueur le 25 mai 2018. <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>

### Loi Jardé :

Le Décret n°2017-884 du 9 mai 2017 (paru au JO du 10 mai dernier) précise le champ des recherches impliquant la personne humaine soumises à l'avis des comités de protection des personnes et, le cas échéant, à l'autorisation de l'Agence Nationale de Sécurité du Médicament. Il explicite à cette fin la notion de « développement des connaissances biologiques ou médicales ». Il simplifie également le dossier de demande d'avis pour des recherches non interventionnelles qui consistent en des réponses à des questionnaires ou à des entretiens, dont une grande partie correspond à des thèses ou des mémoires d'étudiants professionnels de la santé.

<https://www.legifrance.gouv.fr/eli/decret/2017/5/9/AFSP1706303D/jo/texte>

<http://extranet.inserm.fr/actualites/recherche-sur-la-personne-humaine-decrets-de-la-loi-jarde>