



Big data et prédiction

La génétique peut être considérée comme le fer de lance de la « médecine des données » et pourrait apporter d'importants éléments en matière de santé publique. Mais quelle information transmettre au patient ? Avec quel encadrement juridique ?

Médecine génomique : vers une médecine prédictive ?

Sandra Mercier

Professeure des universités, praticienne hospitalière, service de génétique médicale, CHU de Nantes

Les références entre crochets renvoient à la Bibliographie générale p. 57.

Remerciements pour sa relecture à Guillaume Durand, maître de conférences en philosophie, département de philosophie, UFR Lettres et langages, Centre atlantique de philosophie (Caphi, EA 7463), université de Nantes, MSH Ange-Guépin (USR 3491, CNRS), Nantes.

« **G**énétique, jusqu'où a-t-on le droit d'aller ? » Cette question posée par le Comité consultatif national d'éthique (CCNE) en 2016 est de plus en plus d'actualité [16].

L'avènement du séquençage de nouvelle génération

Nous assistons en effet à une révolution technologique en génétique moléculaire avec l'avènement du séquençage haut débit ou séquençage de nouvelle génération (NGS, Next Generation Sequencing). Alors que l'analyse d'un seul gène pouvait prendre plusieurs mois (voire années) avec la technique de référence dite Sanger, le NGS permet aujourd'hui de séquencer, en une seule réaction, un panel de gènes (dizaines ou centaines de gènes) jusqu'à l'exome (région codante de nos 21 000 gènes, soit 1,5 % de notre ADN), voire l'ensemble de notre génome (100 % de notre ADN, 3×10^9 paires de bases).

L'analyse bio-informatique des données générées est fondamentale et complexe. Nous sommes en effet capables aujourd'hui de séquencer l'ensemble du génome d'un individu à des coûts qui n'ont jamais été aussi bas (moins de 1 000 dollars, analyse des résultats en sus). Les progrès liés au séquençage haut débit apportent un meilleur rendement diagnostique, fondamental pour les

patients, en particulier porteurs d'une maladie rare, mais montrent également une grande complexité dans l'interprétation du génome humain. En effet, nous identifions beaucoup de variants de signification inconnue. L'interprétation des résultats reste délicate ; elle nécessite une confrontation clinico-biologique entre le médecin prescripteur et le généticien moléculaire et peut évoluer avec le temps en fonction des connaissances. La corrélation entre le génotype et le phénotype, ainsi que la prise en compte des antécédents familiaux connus chez le patient sont indispensables pour établir le bon diagnostic moléculaire.

Vers une médecine personnalisée

Le plan France médecine génomique 2025 prévoit l'analyse génomique de milliers de patients sur des plateformes nationales. Deux premières plateformes ont débuté les analyses, dans les régions Ile-de-France (SeqOIA) et Auvergne-Rhône-Alpes (Auragen). Nous aurons accès aux données des génomes de plus en plus de patients d'ici les prochaines années.

La médecine 4 P (prédictive, préventive, personnalisée, participative) vise à donner le bon traitement au bon patient au bon moment. Les progrès génétiques liés au NGS et la plus grande accessibilité de ces tests

permettent l'identification de « nouveaux gènes » ou « nouveaux variants » pour les patients porteurs de maladies rares et contribuent sensiblement à améliorer le rendement diagnostique pour ces patients. Il s'agit bien d'une médecine personnalisée où le test génétique va permettre de porter un diagnostic précis pour le patient avec une prise en charge et un conseil génétique adaptés. On entend aussi donner beaucoup de poids à la génétique pour une médecine « prédictive » et « préventive ». Nous sommes maintenant capables d'identifier des anomalies ou des prédispositions génétiques à certaines pathologies dans le cadre d'un diagnostic présymptomatique : par exemple, des prédispositions à certains cancers, certaines maladies neuromusculaires, cardiaques ou neurodégénératives. Dans les familles où la pathologie est déjà connue, il est possible de renseigner les apparentés qui le souhaitent sur leur statut génétique et de déterminer s'ils sont porteurs ou non de l'anomalie génétique. Selon les cas, un conseil génétique, une surveillance ou des mesures prophylactiques sont proposés. Le résultat d'un test génétique est rarement « prédictif » car, finalement, peu de pathologies génétiques recherchées en diagnostic présymptomatique ont une pénétrance complète ; il s'agit plus souvent d'une prédisposition : la personne ne va pas obligatoirement développer des signes de la maladie durant sa vie et, si elle en développe, il est souvent difficile de déterminer précisément l'âge de survenue de la maladie, son évolution et le pronostic. La participation du génotype dans cette médecine « prédictive » ne doit pas être considérée isolément sans prendre en compte l'individu dans sa globalité. En effet, le mode de vie, l'alimentation, l'environnement ont des répercussions sur l'épigénétique, qui régule l'expression de nos gènes, et sont donc des éléments déterminants dans l'expression de ces pathologies. Il serait réducteur de considérer que le phénotype d'un individu est déterminé essentiellement par son génotype. Cela pourrait conduire à des dérives, comme l'illustre notamment le film de science-fiction *Bienvenue à Gattaca*, réalisé par Andrew Niccol en 1997. Ces problématiques ne paraissent plus si futuristes aujourd'hui. Les avis n° 124, n° 129 et n° 130 du CCNE exposent parfaitement l'ensemble de ces problématiques liées à l'évolution des tests génétiques et mettent en garde contre des dérives potentielles, notamment des risques de discrimination [16].

Intérêt médical, risques de dérive et précautions à prendre

Par l'analyse de l'exome ou du génome d'un patient, nous avons potentiellement accès à l'identification de variants dans des gènes qui ne correspondent pas à la recherche diagnostique initiale. Il s'agit d'une donnée dite « secondaire ». L'ACMG (American College of Medical Genetics) a donné des recommandations concernant ces données secondaires et a établi une liste de 66 gènes à analyser si le patient le souhaite [34]. Ces gènes sont dits « actionnables », c'est-à-dire qu'une

prévention et/ou un traitement peuvent être proposés si un variant pathogène est identifié. Il s'agit principalement de gènes impliqués dans les prédispositions aux cancers et à certaines pathologies cardiovasculaires (troubles du rythme cardiaque, cardiomyopathie, dissections artérielles). Cette question est très débattue en Europe et en France [16]. En effet, est-ce que le rendu de cette information sera bénéfique pour une famille déjà fragilisée par la maladie ? Sans doute mieux vaut-il prévenir que guérir, mais il ne faudrait pas non plus tomber dans l'excès inverse.

Nous connaissons le risque potentiel de telles prédispositions dans les familles où la maladie s'exprime, mais nous n'avons pas encore de données fiables sur un éventuel « surrisque » que conférerait la présence d'un même variant dans une famille sans antécédents familiaux. Une étude portant sur une cohorte de personnes âgées (plus de 80 ans) et sans maladies chroniques n'a pas retrouvé de différence significative de fréquence des variants dans la liste de ces gènes « actionnables » alors que l'on se serait attendu à une fréquence beaucoup plus faible [24]. Il existe des facteurs modificateurs, éventuellement protecteurs (génétiques, environnementaux...), qui vont interférer et peuvent modifier sensiblement le risque en population générale. À ce jour, la recherche des données secondaires n'est pas proposée dans le cadre diagnostique en France, mais elle l'est dans le cadre de la recherche (exemple du protocole Defidiag). Notre but en tant que médecin généticien est avant tout d'aider les patients et de ne pas être délétères dans l'annonce d'une prédisposition si celle-ci n'est pas avérée. Dans la balance bénéfices/risques, il faut prendre en compte le caractère anxiogène de ces données pour les patients. Le risque serait de générer à tort une anxiété en rendant toute la population « potentiellement » malade et de prévoir une prise en charge inadaptée.

Si l'on considère à l'avenir que cette recherche de données secondaires peut être proposée aux patients – donc que l'on estimerait un bénéfice médical à connaître ces données –, faudrait-il la proposer en population générale et pas seulement aux personnes malades par souci d'équité ? Dans ce cas, il faudrait pouvoir proposer une prise en charge adaptée, c'est-à-dire calquée sur le diagnostic présymptomatique comprenant un entretien psychologique et un délai de réflexion en amont du test génétique afin de préparer au mieux la personne à recevoir ses résultats. On touche là à une problématique plus large de santé publique concernant l'accessibilité, la faisabilité à grande échelle et le coût médico-économique de ces tests et de leurs conséquences en population générale.

Par ailleurs, on prend en compte également dans les données secondaires les tests préconceptionnels, à savoir la recherche pour un couple du risque de transmettre une maladie à sa descendance pour laquelle les membres de ce couple seraient porteurs sains (hétérozygotes). Le CCNE a proposé dans sa contribution à la révision de la loi de bioéthique (avis n° 129,



septembre 2018) qu'un test soit prescrit à toutes les personnes en âge de procréer qui le souhaitent, après une consultation de génétique pour les maladies héréditaires monogéniques graves. Le CCNE préconise aussi d'examiner les possibilités de l'extension du dépistage génétique à la population générale [16]. La révision de la loi de bioéthique en cours ne semble pas suivre ses propositions, mais ces tests préconceptionnels sont mis en place en Belgique et dans d'autres pays.

Enfin, les tests génétiques en libre accès se multiplient. Ces tests directement proposés aux consommateurs sont interdits par la législation française, mais sont autorisés dans d'autres pays, notamment aux États-Unis, et accessibles sur Internet. Le produit est payé en ligne et l'analyse génétique se fait à partir d'un prélèvement salivaire. Les résultats sont transmis sous forme d'un fichier numérique. La société 23andMe, par exemple, propose des analyses génétiques portant sur la prédisposition au cancer du sein et de l'ovaire, le risque de maladie d'Alzheimer, de Parkinson, etc. La FDA (Food and Drug Administration) avait interdit la commercialisation de ces tests en 2013 à la fois pour des raisons réglementaires, mais également en raison du manque de fiabilité et d'exhaustivité des tests, sans oublier les conséquences délétères potentielles de ces résultats controversés sur les individus. Ces tests sont de nouveau autorisés et les mises en garde sont un peu plus explicites, comme la part liée au mode de vie, à l'environnement et à d'autres facteurs génétiques. Mais nous nous interrogeons sur l'information qui sera comprise à partir de ces éléments bruts, sans contextualisation par rapport aux antécédents personnels et familiaux, et sur l'autonomie de la décision de la personne qui commande de telles

analyses depuis son ordinateur [22]. Par exemple, cette personne va-t-elle comprendre que, dans le cas de la prédisposition au cancer du sein et de l'ovaire, seulement trois variations sont recherchées sur plus de trois mille variations pathogènes connues dans les gènes BRCA1 et BRCA2 et donc que l'analyse, ainsi que les résultats, sont très partiels ? La recherche de cette prédisposition en consultation d'oncogénétique passe par une information exhaustive, délivrée par un médecin spécialiste, un séquençage complet de ces gènes et un rendu des résultats en consultation ainsi qu'un accompagnement psychologique. Plus généralement, la prescription et le rendu de ces tests nécessitent un accompagnement global de la personne concernée par des professionnels spécialisés dans le domaine de la génétique. Une mauvaise interprétation des résultats a des conséquences désastreuses pour les patients et leur projet de vie. C'est le cas d'environ 40 % des tests en libre accès dont les résultats correspondent à des faux positifs liés à des artefacts techniques ou une mauvaise interprétation [53].

En conclusion, un encadrement éthique et juridique des pratiques en génétique doit être renouvelé régulièrement en tenant compte de l'évolution technologique rapide dans ce domaine, mais aussi de l'accessibilité des analyses du fait de la mondialisation. Ces questions nous concernent tous et nous amènent à repenser la société à laquelle nous aspirons pour demain. Il serait de toute façon illusoire de penser que la médecine génomique permettra de prédire précisément la santé et l'espérance de vie d'une personne alors que, dans la grande majorité des cas, celles-ci dépendent essentiellement de déterminants socioéconomiques et non génétiques. ●

Données génétiques massives : quelles règles juridiques et éthiques pour leur usage en santé publique ?

**Emmanuelle
Rial-Sebbag**

Directrice de
recherche Inserm,
UMR 1027, Inserm
université Toulouse III
Paul Sabatier

Les nouvelles techniques génomiques (par exemple le séquençage complet du génome ou de l'exome) permettent dorénavant de générer des données massives pour un usage dans le soin (notamment grâce aux plateformes financées par le plan France médecine génomique 2025) mais également dans la recherche. Ce nouveau contexte technique soulève des enjeux très différents, que ce soit en termes d'acteurs (État, patients, population générale) ou d'obligations juridiques, mais également au regard de la production des données (leur qualité et leur fiabilité doivent être maximales afin d'atteindre un niveau de preuve suffisant pour concrétiser

leur usage). Un niveau de complexité supplémentaire est atteint quand il s'agit d'utiliser les données génétiques massives dans un contexte de santé publique. En effet, notre droit met à la charge de l'État un certain nombre de possibilités pour déployer des actions de santé publique, sans toutefois faire explicitement référence à la génétique. Ainsi plusieurs concepts coexistent, souvent considérés comme proches mais fondamentalement différents (médecine prédictive, médecine préventive ou dépistage). Juridiquement et éthiquement ces approches sont basées sur des dimensions visant soit à développer des stratégies médicales individuelles

(tests génétiques pour le diagnostic), soit des stratégies, plus collectives que ces dernières, ciblant des groupes à risque (dépistage de pathologies dans des populations ciblées) ou la population générale (dépistage néonatal). Nous proposons de clarifier ces différentes approches de la production de données génétiques massives dans leurs dimensions juridiques en insistant sur les spécificités de leur usage en santé publique.

Les règles juridiques communes de la production et de l'utilisation des données génétiques

Quel que soit le contexte d'usage, la génération de données génétiques massives répond aux mêmes exigences que ce soit pour les droits des patients ou pour les droits attachés à la protection des données.

Tests génétiques et droits des patients

Les tests génétiques sont le support des données génétiques. En ce sens, le droit français vise essentiellement à protéger l'individu (et par extension sa famille, dont certains membres peuvent avoir un intérêt médical à être informés de l'existence de maladies génétiques à caractère héréditaire) face aux potentiels mésusages de ses informations génétiques. Ainsi, le Code civil et le Code de la santé publique dotent la personne dont sont issues les données génétiques de droits fondamentaux à l'occasion de la réalisation d'un test génétique (consentement écrit, information complète, conseil génétique, communication avec la famille, etc.) et cela à tous les moments de la chaîne de réalisation du test (prescription, rendu des résultats). Ce cadre juridique strict est construit au prisme des maladies monogéniques héréditaires¹ et tend, précisément, à considérer la protection accordée par le droit comme une prérogative individuelle (voire familiale) pour un diagnostic spécifique mais sans véritable référence à sa dimension collective. Ainsi, au regard des droits des patients, le Code de la santé publique (CSP) n'a pas encore totalement incorporé la dimension « de masse », qui, on le sait, vient pourtant bouleverser les notions d'information et de consentement – alors que toutes les données ne sont pas interprétables ou que les scientifiques ne connaissent pas totalement leur usage futur pour le soin ou pour la recherche (informer sur quoi? comment?) –, de retour de résultats (quels résultats? communiqués à qui?), ou encore de protection de la vie privée (données personnelles, données anonymes)². En sus de la protection offerte par le droit des patients, le droit des données doit également être considéré.

La protection des données génétiques

Par essence, le concept de données génétiques est protéiforme sur le plan scientifique. En effet, celles-ci peuvent

être classées soit en fonction de leur nature (un gène, plusieurs gènes, séquençage complet) soit en fonction de leur usage (médical ou préventif dans un contexte familial par exemple ; recherche des empreintes génétiques pour un usage judiciaire). En droit, la doctrine tend également à différencier la donnée génétique (donnée brute) de l'information génétique (donnée interprétée), impliquant dès lors l'application de cadres juridiques spécifiques. Toutefois, la science et le droit partagent le sentiment que la donnée génétique est particulière, car disant quelque chose de l'intime de la personne testée et potentiellement de sa famille, voire de son identité (ou sa possible identification). De ce fait, tant la réglementation européenne (Règlement général pour la protection des données personnelles, RGPD, (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016, entré en vigueur en mai 2018) que le droit français (loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, LIL, modifiée) couvrent, par des dispositions particulières, la production et l'usage des données génétiques.

L'apport premier du RGPD n'est pas de doter les données génétiques d'un régime particulier (elles tombent sous le même régime juridique que les données de santé), mais plutôt de leur donner une définition. Ainsi, au sens du RGPD, les données, quand elles sont qualifiées de données personnelles (permettant une identification directe ou indirecte des personnes sources), sont des « données à caractère personnel relatives aux caractéristiques génétiques héréditaires ou acquises d'une personne physique qui donnent des informations uniques sur la physiologie ou l'état de santé de cette personne physique et qui résultent, notamment, d'une analyse d'un échantillon biologique de la personne physique en question » (RGPD, art. 4 al. 13). L'art. 9 du RGPD classe ces données parmi les données sensibles (au même titre donc que les données de santé), emportant par principe l'interdiction de leur traitement. Le RGPD pose, bien évidemment, un certain nombre d'exceptions à ce régime permettant de légitimer le recueil et l'usage des données génétiques pour la santé si le consentement de la personne est acquis, ou si le traitement est nécessaire pour des motifs d'intérêt public de recherche scientifique ou encore si les données sont nécessaires pour un traitement médical (y compris la prévention), une prise en charge sanitaire ou sociale, la gestion des systèmes et des services de soins de santé ou de protection sociale. Dans ce dernier cas (usage médical), une condition supplémentaire est requise puisque les personnes réalisant le traitement devront être soumises à une obligation de secret professionnel³.

La loi française a incorporé l'ensemble de ces dispositions tout en rappelant spécifiquement que « dans le cas où la recherche nécessite l'examen des caractéristiques génétiques, le consentement éclairé et exprès des per-

1. Cela devrait évoluer dans la prochaine loi de bioéthique, qui vise à mieux distinguer génétique constitutionnelle et somatique.

2. Comité consultatif national d'éthique, avis 124 : « Réflexion éthique sur l'évolution des tests génétiques liée au séquençage de l'ADN humain à très haut débit », janv. 2016

3. À ces conditions générales devra s'ajouter l'ensemble des conditions nécessaires à la licéité du traitement des données personnelles. RGPD, art. 6.



sonnes concernées doit être obtenu préalablement à la mise en œuvre du traitement de données » (loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, art. 75 al. 1). Ce cadre, bien qu'encore discuté en doctrine, semble clair pour la réalisation des tests dans le contexte médical individuel mais reste à clarifier pour une application en santé publique.

Une nécessaire clarification du cadre juridique des usages en santé publique

Si la réalisation des tests génétique est régie en France de manière uniforme, de nombreux défis restent à relever concernant leur usage en santé publique. L'action du législateur s'est concentrée sur l'application de principes de protection individuelle et reste timide pour son application en population. En effet, le spectre de la possible mise en œuvre de politiques eugénistes, de tri voire de sélection, n'a jamais engagé réellement les pouvoirs publics sur le déploiement de la génétique au service de la santé publique. L'apparition des données génétiques massives et de leur potentiel usage bénéfique au profit des populations vient rebattre les cartes de leur place au sein des politiques de santé publique. Parmi les actions de santé publique, rappelons que les actions de prévention relèvent de la compétence de l'État à titre principal, l'article L. 1411-1 du CSP en donne une définition large : « la prévention collective et individuelle, tout au long de la vie, des maladies et de la douleur, des traumatismes et des pertes d'autonomie ». Ces actions de prévention se déclinent, selon les définitions données par l'OMS, en prévention primaire (agit en amont de la maladie), secondaire (agit à un stade précoce de son évolution) et tertiaire (agit sur les complications et les risques de récurrence). La génétique n'est pas spécifiquement convoquée à l'appui de ces stratégies de santé publique mais en fait partie [2] tout en posant des questions de nature singulière notamment pour le dépistage. Le dépistage « consiste à identifier de manière présomptive à l'aide de tests, d'examens ou d'autres techniques susceptibles d'une application rapide, les sujets atteints d'une maladie ou d'une anomalie passée jusque-là inaperçue » [56].

Les tests génétiques peuvent donc être utilisés à des fins de dépistage mais le Conseil de l'Europe en rappelle les principales conditions de mise en œuvre. Ainsi, un programme de dépistage ne pourrait se déployer que conformément au respect d'un certain nombre de droits fondamentaux, notamment son approbation par une autorité compétente et après une évaluation indépendante portant sur son acceptabilité sur le plan éthique⁴. Le Conseil de l'Europe a établi à cet égard des conditions de validité pour le déploiement d'une politique de dépistage (2008) pouvant se résumer à son utilité, sa faisabilité et son acceptation sociale [6].

4. Conseil de l'Europe. Protocole additionnel à la Convention sur les droits de l'homme et la biomédecine relatif aux tests génétiques à des fins médicales, STCE n° 203, Strasbourg, 27 novembre 2008 (art. 19).

La France met déjà en œuvre des dépistages ciblés (dans le cadre du diagnostic préimplantatoire et prénatal) mais également du dépistage en population pour des indications ciblées. La nouvelle stratégie présentée par le Plan France médecine génomique 2025 suggère l'émergence d'un nouveau modèle en population, ou pour des groupes à risque, par l'établissement de dépistage précoce. Le Comité national consultatif d'éthique (CCNE) s'est également prononcé sur cet élargissement dans son avis 129 (contribution du CCNE à la révision de la loi de bioéthique 2018-2019, sept. 2018) en ciblant deux questions clés : l'extension des modalités de dépistage préconceptionnel (dans les groupes à risque pour d'autres maladies que celles déjà existantes au sein de la famille et extension à la population générale) et le dépistage génétique en population. Dans les deux cas, le CCNE n'est pas défavorable à l'extension de l'usage de la génétique dans ces champs d'action de la santé publique mais sous le respect de conditions strictes (accompagnement par des conseillers en génétique, information claire et consentement, validation par des instances *ad hoc*, liste de gènes et techniques employée, etc.). Sur le plan du droit, s'il n'existe pas d'impossibilité de réaliser ces tests à large échelle (dans le respect des règles existantes), il n'en demeure pas moins que des modifications législatives devront être effectuées pour incorporer dans notre droit les modalités nouvelles de leur réalisation. Il semblerait que cette orientation n'ait pas été totalement suivie par le législateur puisque le projet de loi de révision de la loi de bioéthique n'a pas donné suite à ces élargissements concernant le dépistage général en population ou le dépistage généralisé préconceptionnel (la commission spéciale bioéthique du Sénat avait proposé cette possibilité à titre expérimental, modification qui n'a pas été à ce jour retenue). Cependant, le Sénat a fait droit à la demande concernant le dépistage néonatal puisqu'il a proposé, en première lecture, que ces tests puissent être offerts aux titulaires de l'autorité parentale.

Malgré les évolutions législatives en cours, les données génétiques massives ont du mal à trouver un régime juridique cohérent, entre protection de l'individu et bénéfice pour la collectivité. Les nouvelles techniques, telle que l'intelligence artificielle, viennent s'inviter au débat promettant de passionnantes controverses. 🧠

Les références entre crochets renvoient à la Bibliographie générale p. 57.