

# LE « BIG DATA », RÉSERVOIR DE PROGRÈS POUR LA MÉDECINE

Pour affiner un diagnostic ou le choix d'un médicament, chercheurs et médecins recourent de plus en plus au traitement informatique massif des données de patients.

Alliée aux avancées en génétique moléculaire, cette médecine personnalisée soulève à la fois des espoirs et des questions éthiques.

**D**e l'admission à la sortie de l'hôpital, un patient laisse derrière lui un nombre important de données : nom, prénom, date et lieu de naissance, mais aussi allergies, antécédents familiaux et divers résultats d'analyses médicales – prises de sang, radiographies, etc. Autant d'éléments, qui se comptent en millions, et dont le traitement informatique, allié aux progrès de la génétique, a fait avancer la médecine. En parallèle, le séquençage de l'ADN « à haut débit » a permis de décoder tout ou partie du génome d'un individu très rapidement et à un coût de plus en plus faible. L'évolution est spectaculaire : en 2017, selon l'institut américain National Human Genome Research Institute, le séquençage complet du génome d'un individu coûtait environ 1100 dollars<sup>1</sup> (soit quelque 890 euros), contre 95 millions de dollars au début des années 2000.

## Un espoir face au cancer

Cette profusion de données, que l'on regroupe sous l'appellation « big data », ouvre aux médecins et chercheurs de nouvelles perspectives en matière de diagnostics et de soins. « Nous pouvons faire des diagnostics beaucoup plus précis en croisant de nombreuses données sur ce qu'on appelle des biomarqueurs [la signature biologique d'une maladie, NDLR] ou des marqueurs génétiques acquis par séquençage partiel ou complet d'un génome », explique Anne Cambon-Thomsen, directrice de recherche au Centre national de la recherche scientifique (CNRS) à Toulouse et spécialiste en immunogénétique

humaine, l'étude des facteurs génétiques intervenant dans les mécanismes de l'immunité. Pour les soins, explique-t-elle, le séquençage ciblé de certains gènes d'un patient permet de prévoir quel médicament sera le plus efficace et avec le moins d'effets secondaires possible. Les cancers sont les premières maladies ciblées par cette médecine de précision. Lors d'un cancer, la tumeur évolue et mute

Une équipe de l'hôpital Pitié-Salpêtrière, à Paris, a pour projet de regrouper les données des patients stockées par les hôpitaux et de les recouper avec les données biologiques et génétiques

génétiquement ; certaines de ces mutations échappent au contrôle du corps. « Le séquençage de l'ADN de la tumeur permet déjà d'adapter le traitement, explique Bastien Rance, chercheur à l'Hôpital européen Georges-Pompidou, à Paris, au sein du programme de recherche Carpem (Cancer research for personalized medicine). Il est possible de comparer les mutations de certains gènes précis de la tumeur avec des références internationales et d'en déduire la réponse au traitement. » Reste que l'analyse de ces résultats demeure très chronophage ; pour chaque dossier, toute une équipe de médecins doit être mobilisée. Aussi, l'unité de bioinformatique d'Anita Burgun où travaille Bastien Rance

espère accélérer le processus en proposant aux médecins un outil informatique d'aide à la décision. L'idée est de regrouper les données des patients stockées par les hôpitaux et de les recouper automatiquement avec les données biologiques et génétiques. Soit une précieuse base de données qui pourra aussi être utilisée par les chercheurs. L'une des principales difficultés est d'arriver à extraire informatiquement les informations importantes des comptes rendus médicaux textuels – une mine d'informations que la machine ne sait pas encore « traiter » efficacement. Quoi qu'il en soit, le projet soulève bien des espoirs. Outre le cancer, l'équipe d'Anita Burgun travaille aussi sur les maladies rares, sur lesquelles, par définition, les médecins ont peu d'information. En regroupant les données des malades, les chercheurs espèrent mieux cerner les particularités de chaque pathologie.

## Sécuriser les données médicales

D'autres initiatives de regroupement des données de patients existent un peu partout dans le monde, notamment dans les universités américaines Vanderbilt à Memphis (Tennessee), Stanford (Californie), Harvard (Massachusetts), ou encore à l'université chinoise Fudan, à Shanghai. La force du big data étant liée à la quantité de données traitées, des réseaux de coopération internationale se créent. Un projet s'est notamment lancé en 2015 : l'Observational Health Data Sciences and Informatics (OHDSI), qui regroupe, entre autres, des chercheurs américains, européens, hongkongais et coréens. La

constitution de ces réseaux ne va pas sans poser de problèmes d'éthique et de confidentialité : les données de santé des patients sont extrêmement personnelles et sensibles ; leur regroupement au niveau mondial risquerait de mener à la pire des fuites de données jamais connue. Pour l'heure, les réseaux constitués se limitent donc aux partages de conclusions globales sur les différents échantillons.

## Une éthique différente des deux côtés de l'Atlantique

En France, les données des patients hospitalisés sont conservées sur des serveurs non reliés à internet. Pour y accéder, chercheurs et cliniciens doivent adresser une demande à un comité d'éthique, qui doit

examiner l'étude, jauger les risques potentiels pour les patients (cliniques ou éthiques) et autoriser l'accès aux données strictement nécessaires. De l'autre côté de l'Atlantique, des sociétés de tests ADN, comme le précurseur 23andMe, voient, eux, leur avenir dans le marché plus que prometteur de la médecine personnalisée. Cette société a lancé en décembre une étude sur le lien entre génétique et efficacité des régimes : 23AndMe récolte les données de 100 000 personnes dont elle avait déjà séquencé l'ADN et leur fait suivre un régime, enrichissant par la même occasion sa base de données. Pour l'instant, la médecine personnalisée se développe sans réel cadre juridique mais l'institut de recherche new-yorkais Data &

Society vient de publier un rapport<sup>2</sup> pointant les risques éthiques liés à son essor. L'anthropologue de l'institut Kadija Ferryman et sa collègue psychologue Mikaela Pitcan y insistent notamment sur les biais liés à la discrimination sociale et les problèmes que pourrait engendrer une information mal comprise. En France, les tests ADN en accès libre sont interdits mais la loi ne s'est pas encore emparée de toutes les questions éthiques liées à la médecine de précision. Elles seront au cœur des nouvelles lois de bioéthique dont la discussion doit commencer cet été. ♦

<sup>1</sup> [www.genome.gov/sequencingcostsdata/](http://www.genome.gov/sequencingcostsdata/)

<sup>2</sup> *Fairness in Precision Medicine*, Kadija Ferryman and Mikaela Pitcan, 26 février 2018, <https://datasociety.net/output/fairness-in-precision-medicine>

