

# L'engouement de la médecine pour le big data

La gigantesque accumulation de données médicales rendue possible par les avancées informatiques permet d'identifier des maladies et des épidémies normalement invisibles.

Mais elle comporte aussi des risques.

TEXTE | *Julie Zaugg*

A 57 ans, Michael Snyder était en pleine forme. Sportif et svelte, le chef du département de génétique à l'Université de Stanford faisait attention à son alimentation, si l'on excepte une faiblesse pour les cheeseburgers. En 2011, il a décidé de livrer son corps à une expérience médicale inédite. Tous les deux mois, il a donné un échantillon de sang qui a été analysé de façon extrêmement précise. Cela a permis de générer son profil moléculaire personnalisé (iPOP), un instantané de l'ensemble des processus biologiques se déroulant dans son corps. Il mesurait 30 téraoctets, soit l'équivalent d'un CD qui jouerait sans interruption durant sept ans. Cette masse de données a ensuite été soumise à un algorithme informatique, appelé RiskOgram, qui permet de calculer la probabilité qu'un individu développe telle ou telle maladie.

Quelle ne fut pas alors sa surprise de découvrir qu'il avait 47% de risques de développer un diabète de type 2, soit plus du double du risque normalement associé à un homme de son âge. Une prédiction confirmée par une analyse du niveau de sucre dans son sang. Le RiskOgram a aussi démontré qu'il avait une prédisposition pour le carcinome basocellulaire (une sorte de cancer de la peau) et pour les problèmes cardiaques. Il a aussitôt réduit

sa consommation de sucre et a entamé un traitement anticholestérol. Son taux de sucre est aujourd'hui revenu à la normale.

L'histoire du professeur Snyder illustre l'émergence du big data en médecine, une accumulation de données jamais vue auparavant, dont la récolte et l'analyse ont été rendues possibles par les avancées informatiques. «Le big data permet de faire émerger des signaux normalement invisibles, indique Patrick Ruch, professeur à la Haute école de gestion de Genève - HEG-GE et spécialiste de l'analyse de ce genre de données. Ce qui n'apparaît pas lorsqu'on analyse une population de 500'000 personnes, va devenir visible si on élargit l'échantillon pour inclure plusieurs millions de personnes.» On pourra, par exemple, diagnostiquer plus facilement les maladies rares, qui ne représentent que 8% en moyenne des affections traitées par un hôpital.

La méthode est également utilisée pour comparer les gènes d'un patient avec une série de mutations qui peuvent causer des maladies. «Cela fonctionne assez bien pour certains cancers du sein, détaille Patrick Ruch. La détection d'une prédisposition permet d'effectuer une ablation préventive du sein.» L'actrice Angelina Jolie a choisi cette voie en 2013. A l'avenir, ces tests génétiques permettront d'anticiper

Patrick Ruch, spécialiste de l'analyse des données, explique que le big data permet de faire émerger des signaux normalement invisibles dans le domaine de la santé.

### Les traitements du cholestérol contestés

Une polémique existe actuellement concernant les traitements préventifs du cholestérol, les statines. Ces derniers freinent la production de cholestérol en inhibant une enzyme clé de sa synthèse, l'HMG CoA réductase. En 2013, le pneumologue français Philippe Even a sorti un livre «La vérité sur le cholestérol», dans lequel il accuse les cardiologues et l'industrie pharmaceutique de faire injustement du cholestérol un ennemi public et de gagner des milliards avec les médicaments préventifs. D'autres médecins se sont unis à son appel, mais ils sont encore loin de représenter la majorité. Le Swiss Medical Board a publié fin 2013 un rapport sur le traitement par statine, dans lequel il conclut que «la prescription de statines en prévention primaire est contre-indiquée en cas de risque d'accident cardiovasculaire mortel de moins de 10%».

également des affections comme Parkinson's ou Alzheimer. Cela a été rendu possible par la baisse spectaculaire du coût du séquençage du génome, passé de 100'000 francs au début des années 2000 à moins de 1'000 francs aujourd'hui.

La gigantesque accumulation de données de patients individuels comme Michael Snyder peut faire avancer la recherche. Outre-Atlantique, le Personal Genome Project a récolté depuis 2004 plus de 3'000 profils génétiques. Il va s'en servir pour mieux comprendre l'émergence et la progression de certaines maladies. Le Centre hospitalier universitaire vaudois (CHUV) a, pour sa part, lancé début 2013 un projet unique en Europe: «La Banque institutionnelle lausannoise a pour but de récolter des données génétiques et cliniques sur tous les patients qui sont hospitalisés au CHUV, indique son directeur, Vincent Mooser. Quelque 8'500 personnes ont déjà accepté de participer à ce projet.» Ces données serviront à étudier les mutations génétiques derrière Alzheimer, certaines maladies du foie ou du cœur et Parkinson.



La version complète  
de la revue est en vente  
sur le site

[www.revuehemispheres.com](http://www.revuehemispheres.com)

### Mesurer le flux des travailleurs et des idées

Le *big data* peut aussi être exploité dans les sciences sociales. Professeur de sociologie à l'ETHZ, Dirk Helbing a analysé les profils LinkedIn d'un grand nombre de salariés pour étudier leur provenance et leur lieu de travail. «Depuis le 11 septembre 2001, le flux des travailleurs vers les Etats-Unis s'est tari, alors qu'ils sont toujours plus nombreux à émigrer en Asie», constate-t-il. L'attractivité de ce continent a également crû sur le front des idées. «En examinant qui citait qui dans la littérature scientifique, j'ai pu faire émerger des pôles d'innovation», note-t-il. Ces derniers produisent les idées 'consommées' par d'autres régions du monde. «Or si quelques universités américaines, britanniques, allemandes et suisses ont longtemps dominé ce champ, ce n'est plus le cas aujourd'hui, poursuit-il. De plus en plus d'idées sont générées par l'Asie.» Un membre de l'équipe de Dirk Helbing, Thomas Chadefaux, a en outre procédé à une analyse du vocabulaire utilisé dans les journaux durant les cent dernières années. «Plus on y trouve de mots qui expriment la tension ou la colère, et plus le risque de guerre augmente», explique le chercheur. Il pense qu'une telle méthode pourrait servir à anticiper les conflits avant qu'ils n'éclatent.

Ces informations peuvent aussi servir à repérer les effets secondaires d'un traitement. Des chercheurs des universités de Stanford et de Columbia ont passé au peigne fin les dossiers électroniques des patients des hôpitaux de Stanford, de Harvard et de Vanderbilt pour étudier l'interaction de deux médicaments (Paxil et Pravastatin). Les 130 patients qui avaient pris ces deux préparations simultanément présentaient presque tous un niveau glycémique trop élevé. Ces mêmes chercheurs étaient déjà arrivés à une conclusion similaire en analysant les recherches sur Google, Microsoft et Yahoo comprenant les noms de ces deux médicaments, ainsi que des mots comme «niveau de glycémie trop élevé» ou «vision floue».

L'internet, Facebook ou Twitter représentent en effet une vaste source d'informations pour les adeptes du big data. «L'Organisation mondiale de la santé analyse certains mots clés postés sur les réseaux sociaux pour repérer les épidémies et étudier leur progression, indique Henning Müller, responsable de l'unité eHealth auprès de la Haute Ecole de Gestion et de Tourisme du Valais. Cela permet de mettre en place un système de veille et d'alerte rapide.» Google a développé un système analogue pour traquer la grippe saisonnière. Marcel Salathé, un professeur de biologie d'origine suisse à l'Université Penn, a pour sa part analysé les tweets postés durant l'épidémie de H1N1 pour dresser une carte des Etats-Unis en fonction des sentiments pro- ou anti-vaccins exprimés sur le réseau social. «Cela nous a permis d'identifier des concentrations de personnes hostiles à la vaccination et donc susceptibles de propager la maladie», explique-t-il.

Mais malgré ses promesses, le big data a ouvert une brèche dans la protection de la sphère privée et du secret médical. «Même si les informations ont été anonymisées, on peut retrouver la personne qui se trouve derrière, en effectuant une série de recoupements», note David Billard, professeur d'informatique à la HEG-GE. Il cite une étude de l'Université de Harvard, qui est parvenue à identifier 84% des personnes qui avaient livré leur génome au Personal Genome Project en recoupant les dates de naissance, âges et codes postaux qui y figuraient avec une liste d'électeurs américains.

La situation est d'autant plus critique dans un petit pays comme la Suisse. «A Genève, nous avons un registre qui recense, sous une forme anonyme, tous les cas de cancer depuis 40 ans, poursuit-il. Mais il existe quantité d'organismes qui collectent des données médicales et peu de critères suffisent parfois à déterminer un individu.» Sans oublier que ces données peuvent être hackées ou volées par un employé malfaisant. Voire mises en ligne volontairement par un enthousiaste des réseaux sociaux. «Il existe un vrai risque: un employeur ne va sans doute pas embaucher quelqu'un qui est malade et un assureur ne va pas lui accorder de police d'assurance, met en garde David Billard. Quelqu'un pourrait aussi se servir de données médicales sensibles pour faire chanter un politicien, un entrepreneur ou un haut-fonctionnaire.»

Aux Etats-Unis, il s'agit déjà d'une réalité. La psychiatre américaine Deborah Peel a fondé une ONG, Patients Privacy Rights, pour investiguer ce qu'il advient de la masse de données générée par le dossier électronique du patient. «Toute cette information circule librement entre le cabinet du médecin, l'hôpital, les pharmacies et les laboratoires, indique-t-elle. Or, nous avons constaté qu'une bonne partie de ces données sensibles sont ensuite revendues. J'ai vu des cas où des patients atteints d'un cancer se sont vu refuser un emprunt à la banque, car cette dernière avait eu accès à leur dossier médical.»

Les  
traitements  
du cholestérol  
contestés  
Une polémique  
existe  
actuellement  
concernant les  
traitements  
préventifs du  
cholestérol, les  
statines. Ces  
derniers traitent  
la production de  
cholestérol en  
inhibant une  
enzyme clé de sa  
synthèse, l'HMG  
CoA réductase.  
En 2013, le  
neurologue  
français Philippe  
Even a sorti un  
livre «La vérité  
sur le  
cholestérol», dans  
lequel il accuse  
les cardiologues  
et l'industrie  
pharmaceutique  
de faire  
l'instrument du  
cholestérol un  
ennemi public  
et de gagner  
des milliards avec  
les médicaments  
préventifs.  
D'autres  
médecins se sont  
unis à son appel,  
mais ils sont  
encore loin de  
représenter la  
majorité. Le  
Swiss Medical  
Board a publié  
fin 2013 un  
rapport sur le  
traitement par  
statine, dans  
lequel il conclut  
que «la  
prescription  
de statines en  
prévention  
primaire est  
contre-indiquée  
en cas de risque  
d'accident  
cardiovasculaire  
modéré de moins  
de 10%».