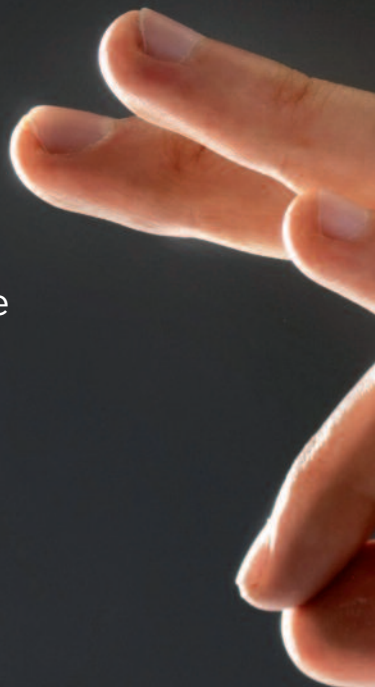


Big data en santé

Ceux qui donneront
le « la »

Depuis 2002, l'Assurance maladie collecte chaque année les données issues de 1,4 milliard de feuilles de soins. Depuis la fin des années 2000 se multiplient les objets connectés - balances, traqueurs d'activité, tensiomètres, piluliers électroniques - qui, petit à petit, constituent de nouvelles bases de données, privées celles-là, sur l'activité physique, les constantes biologiques (poids, tension, pouls, masse grasseuse), l'observance, etc. Cette masse d'informations accumulées est comparée à de « l'or noir » : un gisement de matière première numérique brute qui, s'il est exploité, permettrait une analyse plus fine de nos besoins en santé.

| Par Caroline Coq-Chodorge





Big data en santé

Ceux qui donneront le « la »

Le mouvement est global, et dépasse largement le secteur de la santé : « Tous les 18 mois, le monde génère plus de données que depuis le début de l'humanité, rappelle Vincent Genet, directeur de la division santé du cabinet de conseil en stratégie Alcimed, spécialisé sur les nouveaux marchés. L'enjeu est d'être capable de recouper et d'analyser ces données. L'exploitation du big data nécessite le développement d'algorithmes qui permettent de faire, de manière automatisée et à grande échelle, ces recoupements. » Google est déjà capable de prédire l'arrivée d'une épidémie de grippe en repérant, dans l'immense toile tissée par Internet, les premiers pics de recherche sur les symptômes grippaux. Il est difficile d'anticiper les usages futurs du big data en santé, mais il sera sans doute possible de repérer, grâce aux objets connectés, les signes précurseurs de la dépendance. Des études épidémiologiques de grande ampleur se préparent déjà (voir interview p. 29). Elles croiseront des données publiques et privées : sur les hospitalisations, la consommation de soins de ville et de médicaments, notre activité physique quotidienne, l'évolution en temps réel de nos constantes. L'exploitation de données est, depuis 15 ans, le métier de Celtipharm. Cette entreprise bretonne constitue depuis 2000 une base de données sur



« Grâce à l'analyse des données de la CNAM, nous voulons aider les patients à choisir leur médecin traitant. »

Marianne Binst (directrice générale de Santéclair)

les ventes de médicaments, à partir des données renvoyées régulièrement par 4 600 pharmacies. Celtipharm, qui revendique un CA de 12 M€, permet aux officines de se situer par rapport à leurs concurrentes. Mais la société exploite aussi ces données pour des acteurs publics et privés. Pour le compte de laboratoires pharmaceutiques, elle étudie l'évolution des ventes de médicaments et repère les anomalies. Pour l'Agence nationale de

sécurité du médicament (ANSM), elle a suivi les ventes des contraceptifs, en pleine crise des pilules de 3^e et 4^e générations.

Son P-DG, Patrick Guérin, a bien d'autres projets : « Le big data sera à l'origine des plus grands progrès en matière de santé publique. En croisant nos informations avec celles des objets connectés, nous pourrions déceler les effets secondaires insoupçonnés des médicaments. On saura par exemple que des personnes de tel âge, tel sexe, prenant tel médicament, ont plus souvent des problèmes cardiovasculaires. » Pour affiner ses études, Celtipharm cherche à obtenir un accès à la base de données de l'Assurance maladie. « On pourrait par exemple suivre l'état de santé des jeunes filles vaccinées contre le cancer du col de l'utérus. Ce serait une formidable occasion de rationaliser les débats autour des vaccins. » Très actif au sein du mouvement « Transparence santé », Celtipharm plaide pour « une ouverture aussi large que possible aux données de santé, pour démultiplier la connaissance et sortir de l'obscurantisme. Les seuls qui n'y ont pas intérêt sont les laboratoires les moins scrupuleux. »

La future loi de santé pourrait créer un fichier public des données de santé

Claude Gissot, directeur de la stratégie, des études et des statistiques de la CNAMTS, observe avec circonspection les appétits que suscite le Système national d'information inter-régimes de l'Assurance maladie, la plus grande base de données de santé française : « Le Sniiram, ce n'est pas big brother, on ne sait pas tout sur tout. C'est une base de données qui se constitue pas à pas depuis 2002. Nous l'exploitons à notre niveau en réalisant des études sur les pratiques médicales, la consommation de soins, etc. Elle est également ouverte à l'ANSM, à l'Institut national de veille sanitaire et à 500 chercheurs. » L'article 47 du projet de loi de santé, adopté par l'Assemblée nationale en première lecture, prévoit d'ouvrir, sous conditions, aux acteurs publics et privés les données de la CNAMTS, mais aussi celles de l'hôpital, des établissements médico-sociaux et des collectivités locales sur les causes de décès. Cet article est l'objet d'un âpre lobbying : jugé trop restrictif au départ, il s'est assoupli au fil de son examen par l'Assemblée nationale. Va-t-on assister à une réelle ouverture des données, permettant une compréhension plus fine du système de santé ? Claude



Gissot est là encore prudent : « *Le Sniiram est une base de données difficile à utiliser techniquement. Les risques d'un accès aux données individuelles sont réels. Et les acteurs privés peuvent l'utiliser dans un but commercial. Il y a un équilibre à trouver pour l'utiliser au mieux dans le sens du bien public.* »

Mais pour Christian Saout, secrétaire général délégué du Collectif interassociatif sur la santé, ces arguments visant à encadrer l'ouverture des données de santé ne tiennent pas, justement parce que « *le big data va plus vite que l'open data. Les autorités se battent pour protéger, d'une manière excessive, les données personnelles publiques. Mais le vrai problème, ce sont les données privées. Dans le nouveau monde numérique, tout le monde donne ses données à tout le monde, dans un contexte très concurrentiel et sans réglementation* ». La protection des données personnelles est la mission de la Commission nationale de l'informatique et des libertés (CNIL). Pour sa directrice de la conformité, Sophie Nerbonne, le plus grand flou règne sur les données privées : « *Sur les smartphones, on autorise très souvent, et sans s'en rendre compte, la géolocalisation et l'utilisation de nos données.* » Quant à celles produites par les objets connectés, elle les qualifie de « *données de bien-être. Pour l'instant, nous n'en savons pas grand-chose. Sont-elles anonymisées ? Quelle est la durée de leur conservation ? Où sont-elles hébergées ?...* ».

Alexis Normand, responsable du développement santé de la start-up française Withings, assure pourtant : « *Notre politique est claire. Les données nominatives et réidentifiables ne sont pas exploitables, sauf accord explicite. Nous traitons des données agrégées et anonymes pour suivre l'usage de nos produits et les améliorer. Nous développons aussi des modèles prédictifs : tel comportement annonce une prise ou une perte de poids, etc.* » Cette entreprise française créée en 2008 est l'un des grands acteurs mondiaux de l'esanté grâce à ses pèse-personnes, traqueurs d'activité, tensiomètres ou montres. Les données de chaque utilisateur sont collectées, analysées, synthétisées, et souvent restituées sur les applications mobiles des smartphones. « *Le suivi de l'activité est dynamique, pas statique, avec des compte rendus hebdomadaires, des commentaires sur l'évolution. Nous voudrions intégrer ces données dans le système de soins afin qu'elles soient utilisées par les médecins et les pharmaciens.* » Ces « données de bien-être », récoltées

dans la vie quotidienne, font en effet le lien avec le domicile. « *Le CHU de Toulouse utilise déjà nos objets connectés pour suivre à distance l'évolution du poids et de la tension des patients diabétiques.* »

Les pharmaciens devront se former à l'utilisation des données

Le groupement de pharmaciens PHR est bien décidé à ne pas rater cette occasion : 460 de ses adhérents vendent des objets connectés : « *Des traqueurs d'activité, des balances, des dispositifs de surveillance du sommeil des bébés,* énumère Lucien Bennatan, président de PHR. *Si nous ne voulons pas rester de simples logisticiens, nous devons nous adapter à la profusion des données de santé. Dans une mission de prévention, nous devons devenir les interprètes des données collectées, voire leurs collecteurs.* »

« En croisant nos informations avec celles des données collectées, nous pourrions déceler les effets secondaires insoupçonnés des médicaments. »

Patrick Guérin (P-DG de Celtipharm)

Mais les pharmaciens devront se former à l'utilisation de ces données. « *Malgré une distribution numérique d'Alcura sur 50 % des pharmacies du territoire, seulement une centaine ont souscrit à l'implantation d'un corner santé connectée, soit, en théorie, moins de 1 000 patients sur les 18 derniers mois,* » nuance Romain Watremez, directeur commercial et marketing d'Alcura. Alcura est l'un des précurseurs dans la distribution de ce type de produits en pharmacie en assurant, dès début 2014, la distribution de la marque Withings et en y ajoutant, fin 2014, les marques iHealth et Care. « *La répartition des ventes par type de devices est le suivant : 50 % de traqueurs d'activité, 22 % de balances et 16 % de tensiomètres connectés.* » Withings est la marque qui tire le plus son épingle du jeu, représentant 81 % des ventes. Withings développe également des projets avec les assureurs. Axa a offert en juin dernier Pulse, un traqueur d'activité, aux 1 000 premiers souscripteurs de la complémentaire santé Modulango. Il mesure le nombre de pas effectués dans la jour-



Big data en santé

Ceux qui donneront le « la »

●●● née, le rythme cardiaque, le taux d'oxygène dans le sang et le cycle du sommeil. A ses clients « gros marcheurs » (plus de 7 000 pas par jour), Axa a offert un chèque cadeau de médecine douce d'une valeur de 50 euros. Pour ceux qui marchent plus de 10 000 pas, c'est deux chèques. « C'était une petite opération commerciale dans un but de prévention », minimise le premier assureur mondial.

De cette expérience, les observateurs les plus avertis ont une tout autre analyse : « les assureurs poursuivent deux types de réflexion, décrypte Isabelle Hilali, vice-présidente en charge du marketing et de la stratégie d'Orange Healthcare. La première est de proposer de nouveaux services pour leurs assurés. » Et d'améliorer leur santé. Donc de diminuer le risque... et les coûts. La seconde réflexion est « d'analyser la consommation des soins et les comportements de chaque assuré afin de concevoir des offres mieux ciblées comme cela existe déjà aux Etats-Unis ». En France, cela se pratique dans le domaine de l'assurance automobile, où certains contrats sont adaptés au nombre de kilomètres parcourus.



« Nous traitons des données agrégées et anonymes pour suivre l'usage de nos produits et les améliorer. »

Alexis Normand (responsable du développement santé de Withings)

Pour Marianne Binst, directrice générale du réseau de soins Santéclair, une telle évolution de l'assurance relève du « fantasme : la législation française interdit aux complémentaires santé toute tarification individualisée. Les seuls critères pour moduler les contrats sont l'âge et la zone géographique ». Très engagée dans le mouvement « Transparence santé », Marianne Binst a d'autres projets d'exploitation des données en big data. Filiale de plusieurs complémentaires santé, Santéclair construit des réseaux d'opticiens, de dentistes et d'audioprothésistes qui encadrent leurs tarifs. L'entreprise souhaite affiner les critères d'inclusion des professionnels de santé : « Par exemple, nous voulons connaître la longueur des traitements en orthodontie, très variables, afin de déterminer une durée raisonnable. » Santéclair veut également déve-

opper de nouveaux services pour ses assurés, et plus largement pour tous les patients : « Grâce à l'analyse des données de l'Assurance maladie, nous voulons aider les patients à choisir, de manière éclairée, leur médecin traitant, par exemple en rendant public leur respect des objectifs de santé publique. De même, les patients ont le droit de connaître le taux d'infections nosocomiales des hôpitaux. » « Il existe une tendance de fond, confirme Vincent Genet. Le patient surfe sur Doctissimo, il veut confronter les points de vue, sélectionner ses professionnels de santé... »

« Le meilleur carnet de santé est le téléphone portable »

Sommes-nous en train d'entrer dans l'ère du « quantified self », la mesure de soi ? Apple s'y prépare en développant HealthKit, qui synthétise toutes les données de santé collectées par iPhone ou iPad, ensuite mises en forme par l'application iHealth. La multinationale américaine vient également de lancer en fanfare Apple Watch, une montre connectée au téléphone portable qui donne l'heure, lit les mails, mais surtout mesure l'activité physique : le temps passé à rester assis, en mouvement, à faire de l'exercice. « Avec le temps, elle apprend à vous connaître », explique-t-on chez Apple. Elle est même capable de « suggérer tous les jours un objectif d'activité adapté ». La force d'Apple est bien sûr de parvenir à créer des interfaces intuitives qui déclenchent de nouveaux usages. Certains suggèrent même qu'il finira par relever le défi du dossier médical personnel (DMP).

Vincent Genet est plus mesuré : « Le DMP n'est accessible qu'aux professionnels de santé. Mais il est vrai que certains considèrent, dans une vision assez extrême, que le meilleur carnet de santé est le téléphone portable. La valeur du soin sera désormais centrée sur la capacité à exploiter les données de santé des individus. Dans une approche plus préventive, un patient sera alerté sur ses comportements à risque, de nombreuses maladies seront anticipées, le bon produit sera dispensé au bon patient, et le professionnel de santé s'assurera qu'il est bien observant. » Un nouvel âge d'or se profile-t-il, où la durée de vie en bonne santé va encore s'allonger ? Est-ce au contraire un âge sombre, de contrôle en temps réel de nos moindres faits et gestes ? La première occurrence nous permettra peut-être de connaître la seconde... ◉



INTERVIEW

L'ouverture des données de santé publique ouvre-t-elle la voie au big data ?

Nous disposons de la plus grande base de données de santé du monde. Son ouverture dessine des perspectives immenses. Mais il va falloir se retrousser les manches, car le Système national interrégimes de l'assurance maladie est une base de données d'une complexité extraordinaire. Des acteurs plus nombreux vont développer de nouveaux outils pour l'exploiter, dont du big data, qui permet de traiter des masses de données. Nous pourrions déceler des problèmes de santé à des niveaux territoriaux très fins, par exemple à proximité d'une usine d'incinération. Ou encore nous pencher plus précisément sur les inégalités sociales de santé. Et, bien sûr, nous pourrions confirmer ou infirmer très rapidement des soupçons sur un médicament, ou déceler les effets secondaires rares d'un autre. Le vrai enjeu est de savoir si les moyens seront suffisants. L'Assurance maladie va devoir renforcer ses équipes. Nous avons le potentiel pour

« La France dispose de la plus grande base de données de santé du monde »

MARCEL GOLDBERG,
PROFESSEUR DE BIOSTATISTIQUE ET D'INFORMATIQUE MÉDICALE, CHERCHEUR À L'INSERM ET MEMBRE DU HAUT CONSEIL DE LA SANTÉ PUBLIQUE

devenir les meilleurs au monde dans le traitement des données de santé. Nous pourrions attirer toutes les multinationales de la santé. Si j'étais ministre de la Santé ou de l'Economie, j'investirais dans ce sens.

L'article 47 du projet de loi de santé, au moins dans sa première rédaction, est souvent jugé trop restrictif. Partagez-vous cette analyse ?

Non. Est-ce que tout le monde est réellement d'accord pour que son dossier médical soit aisément accessible ? Les données de la CNAMTS réellement anonymes sont déjà en libre accès. Celles qui vont être ouvertes sont facilement réidentifiables : le numéro d'identification de l'assuré est crypté, mais si je connais l'âge d'une personne, le lieu et la date de son hospitalisation, j'ai 99 % de chances de retrouver son dossier médical. Je pense, comme beaucoup, que ce n'est pas une bonne chose. Il faut donc mettre des garde-fous. Mais l'accès à ces données est extrêmement facilité. Le statut du demandeur – public ou privé – n'entre plus en compte. Assureurs, laboratoires,

entreprises informatiques ou journalistes pourront avoir accès à ces données à partir du moment où ils présenteront le but de leur recherche, qui doit poursuivre un objectif d'intérêt public (celles poursuivant un but commercial sont interdites).

Une base de données de santé privées ou de bien être est-elle en train de se constituer ?

Nous assistons à une explosion des données. Mais je manque sans doute d'imagination, car je trouve qu'il y a beaucoup de charlatanisme et de fantasmes. C'est un peu prématuré de savoir ce qui va ressortir des données produites par l'e-santé, exploitées par Apple ou Google. Il y a aura sans doute des choses utiles, mais moins que ce que l'on imagine. Quelle est la fiabilité des données produites par une montre connectée ? Selon quels protocoles ont-elles été collectées ? Pour avoir une bonne vision de l'état de santé d'une population, j'aurai toujours plus confiance dans la base de données de l'Assurance maladie.
Propos recueillis par Caroline Coq-Chodorge