

## **e-Santé : nouvelles perspectives pour la sécurité sanitaire ?**

*Marius Fieschi et Jean-Yves Robin*

*Open Health Institute Paris (France)*

### **Résumé**

*Contexte.* La généralisation progressive de l'informatisation des processus de soins, le développement des capteurs et des objets connectés, change le contexte de l'acquisition et de l'analyse des données pour la santé publique.

*Méthodes.* De très nombreux travaux témoignent de l'intérêt que suscitent ces masses de données pour la recherche en santé publique et la sécurité sanitaire. Elles posent des problèmes de traitement de l'information et ouvrent des perspectives sur de nouveaux paradigmes méthodologiques. L'analyse de la littérature et des résultats obtenus ces dernières années illustrent la création de valeur rendue possible grâce aux données de la e-santé.

*Résultats.* Les auteurs présentent quelques exemples de résultats touchant à des recherches sur la sécurité sanitaire. Ils abordent dans la partie discussion, quelques questions que soulève l'avènement de ces données massives.

### **Abstract**

*Background.* The progressive generalization of the computerization of the care delivering, the development of sensors and connected objects, changes the context of the acquisition and analysis of data for public health.

*Methods.* Numerous studies demonstrate the interest in these datasets for research in public health and health security. They underline information processing problems and open up perspectives on new methodological paradigms. The analysis of the literature and the results obtained in recent years illustrate the value creation made possible thanks to the e-health data.

*Results.* The authors present some examples of results related to research on safety. They address in the discussion section, some issues raised by the advent of these massive data.

### **Mots clés**

Prévention ; vigilance; e-santé; santé publique

### **Keywords**

Prevention; vigilance; e-health; public health

## 1 Introduction

La sécurité sanitaire a pour objet d'assurer la sécurité face aux risques liés à l'activité du système de santé. Plusieurs définitions et classifications permettent de décrire la sécurité des patients. Rappelons la Patient Safety Taxonomy de la *Joint Commission on Accreditation of Healthcare Organizations* (JCAHO) avec ses cinq domaines d'analyse:

- L'impact de l'erreur médicale
- Les types de processus défaillants
- Le domaine, c'est à dire le contexte dans lequel un incident se produit
- Les causes ou les facteurs conduisant à l'incident de sécurité
- Les facteurs de prévention et d'atténuation pour réduire le risque de récurrence et/ou améliorer les résultats dans le cas d'un nouvel incident.

Selon Didier Tabuteau [1], les objectifs de la sécurité sanitaire sont la prévention ou la réduction des risques iatrogènes, des accidents liés aux traitements, aux actes de prévention, de diagnostic et de soins ainsi qu'à l'usage des produits de santé. Ainsi, toute réflexion sur l'efficacité d'un système de santé devrait s'intéresser à la place et l'importance qu'il attribue à la prévention.

La sécurité sanitaire appartient au registre de la prévention et se place en amont des problèmes, puisqu'elle vise à prévenir les crises, à détecter des phénomènes émergents, qu'ils soient de nature sanitaire, environnementale ou sociétale. Les crises sanitaires de ces dernières années montrent qu'il est nécessaire d'anticiper les situations potentiellement porteuses de risque afin d'aider à l'amélioration de la décision.

Il s'agit donc de prendre les mesures dont on connaît l'efficacité sur des risques connus, attendus (la surveillance, par exemple de la grippe), mais aussi d'être à même de repérer le plus tôt possible les situations nouvelles, inattendues pouvant conduire à des problèmes sanitaires (la vigilance, par exemple en pharmacologie). Dans les deux volets concernés, surveillance et vigilance, les données et les développements de la e-santé apportent aujourd'hui une contribution manifeste.

La santé numérique a ouvert depuis près de vingt ans, d'innombrables opportunités en santé en général et les technologies de l'Internet tout particulièrement, sont devenues parties intégrantes de la surveillance en santé publique. Le numérique apporte une contribution déterminante pour faire face aux nouveaux enjeux de la prévention notamment en permettant de placer la santé dans son contexte social et environnemental.

Après avoir rappelé les enjeux de la sécurité en santé et la contribution de la e-santé dans ce domaine, nous illustrerons notre propos par des exemples de la littérature de ces dernières années. L'ensemble de la littérature identifiée étant très diversifiée, nous avons choisi de réaliser une synthèse qualitative au travers de quelques travaux marquants. Dans la discussion nous reviendrons sur la nécessité d'intégrer ces outils dans une politique de santé publique qui demande de favoriser le partage des données de santé en recherchant un équilibre entre les réponses législatives et éthiques et le développement d'une dynamique de recherche et d'un écosystème dans ce domaine.

## 2 Les enjeux de la sécurité en santé

L'OMS distingue trois types d'actions de prévention :

- *La prévention primaire* qui a pour but d'éviter l'apparition de la maladie en agissant sur ses causes.
- *La prévention secondaire* qui vise à détecter la maladie ou la lésion qui la précède, à

un stade où la prise en charge efficace des individus malades peut intervenir utilement.

- *La prévention tertiaire* qui a pour objectif de diminuer la prévalence des récurrences et des incapacités conséquentes.

On peut y ajouter la prévention quaternaire. Elle se caractérise par l'ensemble des activités de santé destinées à atténuer ou à éviter les conséquences de l'intervention inutile, excessive voire iatrogène du système de santé. L'efficacité et l'innocuité des médicaments participent à la prévention sanitaire que doit accompagner une pharmaco- et matériovigilance opérante.

De récentes affaires ont révélé au grand jour les défaillances du dispositif français de sécurité sanitaire. L'affaire du Médiateur® en particulier a amené, dans un premier temps, une incompréhension de la part du public puis, et c'est là un effet délétère induit préoccupant, une suspicion vis-à-vis des messages de santé publique que délivrent les autorités. A fortiori après que la démonstration ait été faite que l'analyse des données disponibles permettait à tout le moins d'objectiver le mésusage massif du produit.

Les questions majeures retrouvées dans la littérature, concernent des variations importantes dans la qualité des soins et les risques considérables de dommages iatrogènes. Ceux-ci concernent la surveillance des effets indésirables connus chez les patients traités mais aussi la vigilance pour des effets non connus notamment en cas de mésusages du médicament. Ces défauts de suivi et/ou de connaissance contribuent d'une manière importante aux taux élevés de morbidité et de mortalité potentiellement évitables, et à l'augmentation des dépenses de santé.

### **3 La contribution de la e-santé**

#### **3.1 La e-santé productrice de données massives**

On a résumé la problématique des données massives (*Big Data*) aux 3 V pour Volume, Vitesse et Variété.

- volume, car les masses de données à traiter sont sans cesse croissantes.
- vitesse, car la collecte, l'analyse et l'exploitation des données doit se faire de plus en plus souvent en temps réel.
- variété, car les données sont de formes très variées, hétérogènes et parfois non structurées.

Le développement des objets connectés et les pratiques de géolocalisation accroissent également considérablement le volume de données collectées disponibles.

La contribution de la e-santé, avec la production et le traitement des données massives, introduit un véritable changement de paradigme pour les chercheurs en épidémiologie et en santé publique. Ce changement prolonge celui qui était déjà inscrit dans l'approche du *data mining* sur les entrepôts de données. Il est lié à l'origine des données dont la production précède la question à analyser. L'expérience n'est plus planifiée en relation avec l'objectif, avec la question posée. Les données sont issues d'entrepôts, de bases de données classiques ou de ressources sur le Web, collectées et stockées pour d'autres raisons. C'est une approche guidée par les données massives et multi sources.

#### **3.2 Des traitements accessibles et des modèles plus précis**

Les outils agrégateurs de données de différentes sources et de visualisation ont engendré une nouvelle génération d'applications de surveillance des maladies qui peuvent fouiller, filtrer et visualiser en temps réel, *online*, des données pour comprendre les phénomènes épidémiques. « Voir c'est croire » est une assertion qui n'est pas uniquement valable pour les techniques d'imagerie, elle est valable pour présenter et donner du sens aux données de

santé.

Traitement des données, applications, nouveaux moyens de pilotage et de détection... le numérique est un des principaux leviers pour mettre en œuvre une politique de prévention plus efficace et moins coûteuse en santé. Ces données peuvent faciliter la prévention des maladies :

- En augmentant la capacité que nous avons de comprendre les déterminants comportementaux, sociaux et environnementaux de la santé en population
- En rendant possible les efforts de prévention mieux ciblés vers les sous-populations pour lesquelles ces efforts sont les plus efficaces
- En fournissant à la recherche de nouveaux moyens pour identifier les facteurs de risque modifiables des maladies

La disponibilité des données massives peut permettre de produire des indicateurs à un niveau de granularité plus fine, par exemple à l'échelle d'un territoire, sur des sous-populations, voire sur des données au niveau individuel.

Par ailleurs, les *Big Data* en santé permettent un alignement temporel plus précis des épidémies avec des co-variables qui peuvent améliorer la précision des modèles. Elles diffèrent des outils d'aide à la décision traditionnels dans ce domaine en permettant la collecte et l'analyse des données en temps réel.

Il est incontestable que l'exploitation des données massives procurées par les développements de la e-santé est un important levier pour une meilleure compréhension des maladies, du développement de médicaments et du traitement et de la sécurité des patients.

## 4 Quelques exemples de la littérature

Nous nous proposons ici, sans prétendre à l'exhaustivité des types d'applications ou des méthodes développées dans ce domaine, d'illustrer notre propos sur les apports de la e-santé et de montrer son potentiel de création de valeur et de connaissance.

### 4.1 Exemple de la détection d'une maladie d'origine alimentaire (infection entéro-hémorragique à *Escherichia coli*)

C'est un exemple de ce que l'on appelle le « *zero patient problem* ». Il consiste à déterminer la source initiale d'une épidémie étant données des informations fournies en temps différé.

Le changement des habitudes alimentaires, la globalisation des chaînes de production, augmentent les difficultés de retrouver l'origine de ce type de maladie. Elles se développent après un délai qui ajoute à la complexité de l'enquête et à la difficulté de l'établissement de la preuve. Il est généralement impossible d'estimer l'origine géographique d'un tel phénomène en se basant sur la distribution spatiale des cas rapportés.

L'exemple porte sur la détection de l'origine de l'intoxication alimentaire de 2011 en Allemagne [2] qui a affecté près de 4 000 personnes avec des taux de gravité (22%) et de mortalité (1,4%) très importants. Cette intoxication a été faussement attribuée initialement à des légumes importés alors que la cause était dans une région de la basse Saxe.

Ce travail est intéressant à plusieurs titres : Il montre l'intérêt du modèle mathématique élaboré d'une part et de l'utilisation de données multi sources pour ce type de problème de santé publique d'autre part. (Ces investigations demandent l'analyse de données de différentes sources et une analyse de données spatiales de transport de nourritures).

## **4.2 La modélisation de la propagation des épidémies pour comprendre, savoir où vacciner [3]**

Quel est l'impact prévisible d'une stratégie de vaccination ? Peut-on prévoir la diffusion de l'épidémie de grippe durant l'hiver en France ? Ce type de questions, intervient souvent dans le cadre d'une actualité sensible et l'exemple de la vaccination contre la grippe H1N1 est encore dans tous les esprits. Ces problèmes génèrent de nombreux travaux multidisciplinaires qui se sont multipliés depuis quelques années, notamment depuis l'avènement des systèmes d'information en temps réel. Grâce à ces systèmes d'information et à la modélisation mathématique, il est possible de proposer des prévisions épidémiologiques qui contribuent à une prévention et à un contrôle plus efficace des pathologies, notamment infectieuses. Ces résultats, pour être validés, exigent des observations recueillies à partir de longues séries constituées dans le temps et l'espace.

La distribution spatiale des personnes et leur comportement, notamment concernant leur mobilité, représentent deux éléments clés de la diffusion spatiale d'une maladie infectieuse affectant la population hôte [4]. Les approches de modélisation de propagation des épidémies qui en tiennent compte améliorent les résultats de ces modèles, concernant la persistance d'une infection dans la population ou les conditions du développement d'une épidémie émergente. L'intégration de ces paramètres dans les modèles serait hors de portée sans les moyens que proposent les technologies de l'information et de la communication avec par exemple, les données fournies par la téléphonie mobile [5].

## **4.3 La pharmacovigilance, un processus observationnel [6]**

### *Améliorer la pharmacovigilance*

L'analyse des données issues de cohortes, généralement très importantes, ou des bases médico-économiques sur le long terme permet d'observer de nombreux phénomènes, et notamment de faire des rapprochements entre des traitements et la survenue d'événements en santé. Cette pratique permet de **repérer des événements indésirables graves** et d'alerter sur certains risques.

Un meilleur management des *Big Data* doit permettre une identification plus systématique de signaux comme la détection précoce des effets indésirables et des réactions lors de l'utilisation des médicaments en population. Les exemples, ci-dessous, illustrent de manière magistrale l'importance de ces problèmes.

Différents modèles de plateforme et d'organisation de traitement des données sont proposées pour améliorer l'interopérabilité sémantique des systèmes et faciliter le traitement automatique en continue de ces données. Parmi ces différents modèles, on peut citer à titre d'exemple le « *Dynamic Clinical Data Mining* » [7]. L'idée est de réaliser des recherches automatiques sur des bases de données de patient de-identifiés pour déterminer des résultats et des traitements chez des patients similaires de manière à fournir une aide à la décision individualisée pour le laboratoire producteur comme pour le patient.

## **4.4 L'importance de l'usage des données massives, exemples en pharmacovigilance**

### *4.4.1 Contraception orale et risque cardiovasculaire*

En 2013, les autorités sanitaires françaises ont été conduites à dé rembourser des pilules de troisième génération suite à des études montrant une légère mais significative augmentation du risque d'embolie pulmonaire, voire d'infarctus ou d'AVC<sup>1</sup> en cas de prise régulière de

---

<sup>1</sup> AVC : Accident vasculaire cérébral

contraceptifs hormonaux. Afin de confirmer ou d'infirmer ce sur-risque, une étude, publiée dans le BMJ, a été conduite en France sur 5.000.000 de femmes de 15 à 49 ans [8], à partir de données de l'assurance maladie, la base de données du SNIIRAM, et de la base de données du PMSI, pour étudier le risque d'embolie pulmonaire et thromboembolique chez les femmes utilisant une pilule contraceptive de 3ème génération. Les résultats donnent une estimation affinée de la fréquence globale des événements cardiovasculaires sous pilule. Elle semble légèrement plus élevée, au moins pour les accidents thromboemboliques veineux, que dans la population générale.

#### **4.4.2 L'analyse des données conduit au retrait de médicaments dangereux : exemple du Vioxx [9]**

80 millions de patients ont pris ce médicament entre 1999 et 2004. Les risques cardiovasculaires du Vioxx ont été signalés dans un certain nombre d'études sur de petits échantillons.

Il a fallu plus de cinq ans pour que le laboratoire Merck prenne une décision sur cette question. Il a fallu qu'une étude de DJ Graham de la FDA<sup>2</sup>, portant sur 1,4 millions de patients de la base de données Kaiser Permanente compare le risque d'événements cardiovasculaires indésirables pour les utilisateurs de Vioxx contre le risque pour les utilisateurs de Celebrex de Pfizer. L'étude a conclu que plus de 27.000 infarctus du myocarde (crise cardiaque) et morts subites ont eu lieu entre 1999 et 2003. Ils auraient pu être évités. Ainsi, ces résultats ont conduit à ce que l'on a pu appeler le plus grand rappel de médicament dans l'histoire de la médecine. Merck a retiré Vioxx de la vente en pharmacies.

#### **4.4.3 Les prescriptions dangereuses et le scandale de la Dépakine chez les femmes enceintes [10] [11]**

La Dépakine® (valproate de sodium) est commercialisée depuis 1967. Ce médicament a permis, avec le Tégrétol® (carbamazépine), un progrès considérable dans la prise en charge des épilepsies avec une meilleure tolérance et une efficacité thérapeutique accrue par rapport aux deux alternatives existantes à l'époque, le Gardéнал® (phénobarbital) et le Dihydan® (phénytoïne). La possible tératogénicité des anticonvulsivants est évoquée pour la première fois en 1968. Il faudra attendre les années 80 pour que de premières publications incriminent précisément le valproate de sodium dans des cas de malformations. Il faut attendre 1986 [12] pour que la première cohorte prospective atteigne un effectif statistiquement suffisant permettant de confirmer le risque significativement élevé d'anomalie de fermeture du tube neural chez des enfants après exposition de la mère au valproate de sodium durant le premier trimestre de la grossesse.

Si de nombreuses études ponctuelles portant le plus souvent sur des effectifs très réduits documentent les risques du valproate de sodium, l'absence de larges bases de données contenant à la fois des données de la mère et des enfants a constitué un frein certain. Le retard d'informatisation des prises en charge psychiatriques (le valproate est aujourd'hui plus utilisé dans ces indications que comme anticonvulsivant) et des établissements médicaux-sociaux (où sont souvent pris en charge les enfants porteur d'un handicap) ne facilitent sans doute pas les études rétrospectives. Il faudra ainsi attendre 2014 pour qu'une cohorte (registre REMERA) mette en évidence 33 enfants présentant des malformations sur 9 ans de suivi. La première mention du risque de malformations est insérée au Résumé des Caractéristiques du Produit (RCP) en 1997 en minimisant toutefois ces risques (fréquence et réalité non établies). La prise de valproate pendant la grossesse sera simplement

---

<sup>2</sup> FDA : Food and Drug Administration

déconseillée en France à partir de 2006. Des mesures restrictives de prescription et de dispensation sont prises en 2015 avec retard. Les premières études mobilisant les données du SNIIRAM<sup>3</sup> ont été diligentées en 2015 et ne permettent pas encore d'évaluer le nombre des grossesses menées sous valproate et la prévalence des complications tératogènes plus de 40 ans après les premières publications. Pas plus que n'est évalués à ce stade l'impact des mesures prises pour réduire ces risques.

## **5 Discussion**

Le processus classique d'évaluation du risque repose sur des données d'exposition qui comportent des limitations (systèmes de déclaration, suivi de cohortes limitées sur quelques données choisies et collectées,...). Il nécessite de faire des hypothèses et des choix subjectifs sur les données à recueillir, les indicateurs pertinents,... Ces limites doivent être considérées avec attention quand il s'agit de guider la politique de la santé. L'obligation de moyens prend ici tout son sens au niveau des interventions collectives de santé. Si l'essai contrôlé randomisé et le suivi longitudinal de cohortes restent le gold standard et la méta analyse de ces études la méthode à plus forte valeur probante, les délais nécessaires peuvent être très préjudiciables.

On ne peut nier l'intérêt, dans le contexte de données massives et multi-sources, de la possibilité de les combiner. Disposer de nombreuses observations peut permettre de prendre en compte des interactions entre variables. L'exploration des données peut s'avérer très intéressante pour faire émerger des relations que l'on ne suspectait pas initialement. L'exploitation de la masse de données conduit à détecter des régularités inconnues jusqu'alors ou des signaux dont la réalité mérite d'être confirmée par d'autres méthodes, notamment celles éprouvées de l'épidémiologie traditionnelle.

Les données massives permettent des segmentations particulières des populations et de répondre aux besoins plus précisément en personnalisant les actions.

L'approche *Big Data* n'est pas exempte de critiques méthodologiques. La difficulté par exemple qu'il peut y avoir à reproduire certains résultats quand ils mobilisent des algorithmes complexes sur des flux de données parfois volatiles met à mal le principe de la reproductibilité des résultats qui fonde les principes de l'expérimentation. Les limites de l'interopérabilité peuvent conduire parfois à des approximations sémantiques entre référentiels de données différents. Et bien sûr, les risques de confusions entre corrélation et causalité sont amplifiés par ces méthodes de fouille de données multi sources.

Tirer le meilleur parti des apports des différentes approches doit être la ligne à suivre pour le bénéfice de la santé publique. On doit considérer que ces approches se complètent plus qu'elles ne s'opposent.

Outre le bénéfice de la création de valeur et de l'amélioration des connaissances, ces approches participent à l'amélioration de la transparence du travail sur les données et contribuent à les rendre plus accessibles. La recherche en épidémiologie traditionnelle implique des équipes ayant le contrôle exclusif de leurs données. Les données brutes sont rarement disponibles. Cette pratique peut nuire à la reproductibilité, la responsabilisation et l'efficacité [13] : Optimiser l'utilisation des ressources pour les études et favoriser le partage des données, telle est la vision dans laquelle s'inscrit le développement de la e-santé.

L'enjeu crucial est aujourd'hui d'accompagner ces techniques et de les intégrer dans un système de santé cohérent qui se complète au fur et à mesure que les techniques ouvrent de nouveaux champs d'investigation, de données, de connaissances.

---

<sup>3</sup> SNIIRAM : Système national d'information inter-régimes de l'Assurance Maladie

## 5.1 Des nouveaux outils de recherche

Les données massives transformeront-elles nos manières d'étudier ? Sans succomber aux discours qui annoncent régulièrement une révolution, il faut reconnaître que l'abondance de données tend à modifier en profondeur certains aspects de la recherche.

Il est désormais quasiment possible d'étudier la population entière plutôt qu'un échantillon restreint. Le changement est significatif, puisque c'est la question de la représentativité qui cesse de se poser quand l'un et l'autre sont similaires.

La répétition de tests statistiques pose des problèmes méthodologiques qui limitent l'interprétation avec l'inflation du risque alpha.

L'introduction des modèles développés avec les données massives a permis d'élargir la perspective de modélisation traditionnelle pour englober un grand nombre d'individus et par exemple, les données de téléphonie mobile ont été utilisées pour créer des modèles réalistes de la mobilité humaine [14].

## 5.2 Ces outils ont leurs limites

En grande dimension, la probabilité, d'observer des corrélations importantes entre variables n'ayant en réalité aucun lien, augmente. On peut être tenté d'interpréter à tort une corrélation fortuite. Un exemple concret : La défaillance de l'application Google Flu, est sans doute liée à des corrélations fortuites. Cet outil, lancé en 2008 aux Etats-Unis, avait pour objectif de prévoir la progression de l'épidémie de grippe. A ces débuts, l'outil parvenait à fournir un indicateur une à deux semaines avant les publications officielles aux Etats-Unis. Par la suite, il a été mis en défaut notamment lors des hivers 2011-2012 et 2012-2013, en surestimant les épidémies de grippe aux États-Unis. Les termes de recherche sélectionnés pour prévoir l'épidémie étaient corrélés aux données sanitaires mais ne présentaient pas de rapport direct avec la maladie.

## 5.3 La place particulière des réseaux sociaux [15] [16]

La détection des effets indésirables et des accidents thérapeutiques souffrent d'une sous déclaration importante de la part des patients (86% [17]). Ils passent de plus par l'intermédiation des professionnels de santé, des laboratoires, des services de pharmacovigilances, ce qui peut être potentiellement la source de perte d'information ou d'erreurs. Les autorités sanitaires dans la plupart des pays et notamment la HAS en France, l'EMA<sup>4</sup> et la FDA à travers le *Safety and Innovation Act* soulignent l'importance de la surveillance post AMM des produits et la nécessité d'identifier de nouvelles sources de données. Parallèlement, les réseaux sociaux bénéficient d'un usage croissant voire massif et sont utilisés par les patients pour partager leurs retours d'expérience lors de l'usage de médicaments. Ces réseaux sociaux constituent un nouveau media permettant une expression directe des patients et potentiellement utile pour la détection d'effets indésirables médicamenteux.

L'analyse des posts via le réseau Twitter reste délicate et l'intervention humaine en complément de l'usage d'ontologies permettant une automatisation de l'analyse des messages reste nécessaire. Mais les études montrent une corrélation encourageante avec les déclarations effectuées via les canaux classiques. Si l'étude des données provenant de réseaux dédiés à la santé contiennent une proportion plus élevée de messages identifiés comme relatifs à des effets secondaires, le volume considérable de données disponibles via les réseaux sociaux généralistes permettent l'obtention d'un volume plus élevé de données pertinentes. Bien que des travaux restent nécessaires dans les techniques de traitement de

---

<sup>4</sup> EMA : European Medicines Agency



ces données, la croissance rapide de l'usage de ces media devraient avoir à terme un impact important dans le domaine de la pharmacovigilance.

#### **5.4 Les implications juridiques et éthiques**

Un pas majeur a été franchi avec le nouveau règlement européen sur la protection des données personnelles adopté le 27 avril 2016 qui vient remplacer une directive de 1995 alors que l'Internet n'en était qu'à ses débuts. Il introduit notamment la notion de « finalité compatible » qui facilite la réutilisation des données pour des traitements différents (mais compatibles) avec les motifs de leur collecte initiale. Ce point s'avère majeur en matière de sécurité sanitaire. Le nouveau règlement adopte également une série de mesures renforçant le droit des personnes sur leurs données personnelles ce qui est de nature à renforcer la confiance des citoyens dans le numérique en général et la protection de leur vie privée en particulier. La protection des données personnelles et de la vie privée reste de la responsabilité des Etats et des autorités de régulation (la Commission Nationale Informatique et Libertés en France) même si demain son cadre sera européen. Mais des solutions et des approches techniques maladroites, comme d'imposer des niveaux d'agrégation élevés ou des durées de conservation trop courtes, ont souvent porté préjudice à l'intérêt général. Les études en santé vont nécessiter de disposer de données individuelles conservées sur de longues durées. Il appartient donc au législateur d'apporter des réponses adaptées qui permettent de prendre en compte des équilibres bénéfiques / risques en constante évolution.

L'accès à de nouvelles connaissances du fait de la disponibilité de ces données et l'amélioration potentielle de la conduite de la politique de santé qu'elle offre, va sans aucun doute modifier la responsabilité des acteurs. De même, grâce aux données, l'évaluation des politiques publiques devrait se systématiser dans les prochaines années modifiant sans aucun doute de façon vertueuse, les processus de décision et poussant progressivement à une obligation de moyens plus transparente, plus objective et plus exigeante. Ainsi des mesures de santé publique qui n'auraient pas l'effet attendu mesuré au décours de leur mise en œuvre ne manqueront pas d'être analysées ce qui contribuera à une meilleure efficacité de l'action publique. Il en va d'ores et déjà ainsi par exemple de l'échec des campagnes de vaccination contre la grippe qui s'avèrent incapables année après année d'améliorer la couverture vaccinale en France et son corollaire en terme de surmortalité liée à l'épidémie grippale chaque année. Les exemples de ce type pourraient être multipliés et montrent également toute l'importance d'une ouverture des données permettant à une pluralité d'acteurs de procéder à ce type d'évaluation sans en laisser le monopole à la personne publique.

#### **5.5 Des obligations pour les industriels et les régulateurs**

Le développement de l'usage des données de santé repose sur la création d'un environnement facilitateur au sein duquel chaque acteur assume clairement son rôle. La loi dite de modernisation de notre système de santé adoptée en janvier 2016 prévoit une ouverture des données de santé sans se donner réellement les moyens de cette ambition.

Les professionnels de santé produisent une part essentielle des données de santé. Leur activité est largement financée par la solidarité nationale. Si un système d'incitation / pénalités a été mis en place au fil des ans par l'assurance maladie pour l'informatisation des feuilles de soins et que de premières incitations financières apparaissent pour la tenue d'un dossier médical informatisé, la production de données devrait être reconnue comme une de leurs missions essentielles faisant partie intégrante de la démarche de soin et de prévention. Il en va de la qualité des données recueillies.

Pour assurer sa qualité et sa faisabilité, la transmission des données telle qu'évoquée plus

haut, impose l'intensification de la recherche de l'interopérabilité sémantique des systèmes d'information et des applications traitant des données de santé. Cette nécessité doit être accompagnée par les pouvoirs publics de manière particulièrement énergique. Elle doit être opposable et faire l'objet d'une adhésion inconditionnelle des industriels.

Les patients devraient être clairement informés de l'utilisation de leurs données au service de l'intérêt général, cette contribution étant partie intégrante du contrat conduisant à rembourser les soins en préservant leur droit à s'y opposer pour des motifs légitimes.

L'Etat devrait se concentrer sur ses missions régaliennes que constituent les dispositions relatives à la protection des données personnelles et de la vie privée et l'organisation des mesures collectives de solidarité dont l'ouverture des données de santé fait partie. A souvent prétendre être un opérateur, les pouvoirs publics assument mal leurs missions dans la durée. Les lenteurs des pouvoirs publics illustrées ci-dessus pour faire face à des crises sanitaires voire les mesures restrictives abusives qu'ils peuvent être tentés de prendre comme l'arrêt relatif à l'accès aux données du SNIIRAM annulé par le Conseil d'Etat au terme de plusieurs années de procédure le démontrent [18].

La régulation est assurée dans tous les pays européens par des autorités administratives indépendantes (la CNIL en France). Leurs moyens devraient permettre de maintenir leur efficacité et leur indépendance. Dans le domaine de la santé, l'appréciation de la légitimité des traitements et notamment leur conformité à l'intérêt général devrait avantagusement relever de ces autorités indépendantes renforcées par des expertises sectorielles plutôt que par des commissions placées auprès des administrations centrales des ministères.

Les industriels du médicament ont une responsabilité croissante, et une obligation légale, de veiller au bon usage qui est fait de leurs produits par les professionnels de santé. Restreindre leur accès aux données de santé est contradictoire avec cette responsabilité. La gestion des éventuels conflits d'intérêt pourrait relever d'un contrôle sur le respect des bonnes pratiques notamment méthodologiques mais ne sauraient motiver une interdiction à accéder à des données.

Rappelons enfin que les données peuvent être la source de nouvelles connaissances et de création de valeur dont le développement économique et la création d'emplois font partie. Il est majeur, dans la compétition internationale, que les acteurs économiques aient un accès facilité aux données de santé dans le respect des deux règles citées plus haut, le respect de la vie privée et l'intérêt général.

Les collaborations publiques – privées sont la clé pour pouvoir disposer rapidement des moyens compétitifs de recherche et notamment de plateformes épidémiologiques nationales et internationales.

## Références

[1] Tabuteau D. La sécurité sanitaire", Berger - Levraut, 1994.

[2] Manitz J, Kneib T, Schlather M, Helbing D, Brockmann D. Origin Detection During Food-borne Disease Outbreaks – A Case Study of the 2011 EHEC/HUS Outbreak in Germany. PLOS Currents Outbreaks. 2014 Apr 1.

[3] [Guo Z](#), [Xiao D](#), [Li D](#), [Wang X](#), [Wang Y](#), [Yan T](#), [Wang Z](#). Predicting and Evaluating the Epidemic Trend of Ebola Virus Disease in the 2014-2015 Outbreak and the Effects of Intervention Measures. PLoS One. 2016 Apr 6;11(4): e0152438.

[4] Paz-Soldan V, Morrison AC, Elder JP, Kochel TJ, Scott TW, Kitron U. Usefulness

of commercially available GPS data-loggers for tracking human movement and exposure to dengue virus. *Int J Health Geograph* 2009; 8: 68.

[5] Lu X, Bengtsson L, Holme P. Predictability of population displacement after the 2010 Haiti earthquake. *Proc Natl Acad Sci U S A* 2012 ; 109 : 11576-81.

[6] Celi LA, Moseley E, Moses C, Ryan P, Somai M, Stone D, Tang KO. From Pharmacovigilance to Clinical Care Optimization. *Big Data*. 2014 Sep 1;2(3):134-141.

[7] Celi LA, Zimolzak AJ Stone DJ. Dynamic clinical data mining: search engine-based decision support. *JMIR Med Inform*. 2014 Jun 23; 2(1):e13.

[8] Weill A, Dalichampt M, Raguideau F, Ricordeau Ph, Blotière PO, Rudant J, Alla F, Zureik M. Low dose oestrogen combined oral contraception and risk of pulmonary embolism, stroke, and myocardial infarction in five million French women: cohort study. *BMJ* 2016; 353:i2002

[9] Graham DJ, Campen D, Cheetham Ceal. Risk of acute myocardial infarction and sudden cardiac death associated with use of COX-2 selective and nonselective NSAIDs. Presented at the 20th International Conference on Pharmacoepidemiology and Therapeutic Risk Management, Bordeaux, France, August 22–25, 2004.

[10] Vajda F, O'Brien T. Valproic Acid use in pregnancy and congenital malformations. *N Engl J Med* 2010; 363:1771-1772 [October 28, 2010](#)

[11] Morrow J, Russell A, Guthrie E, et al. Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register. *J Neurol Neurosurg Psychiatry* 2006;77:193-198

[12] Lindhout D, Schmidt D. In-utero exposure to valproate and neural tube defects. *Lancet*. 1986 Jun 14;1(8494):1392-3.

[13] Ioannidis JP. Why most published research findings are false. *PLoS Med* 2005;2:e124.

[14] Bengtsson L, Gaudart J, Lu X, Moore S, Wetter E, Sallah K, Rebaudet S, Piarroux R. Using mobile phone data to predict the spatial spread of cholera. *Sci Rep*. 2015 Mar 9;5:8923. PMID: PMC4352843

[15] Freifeld C.C, Brownstein J.S, Menone C.M, Bao W, Filice R, Kass-Hout T, Dasgupta N Digital Drug Safety Surveillance : Monitoring Pharmaceutical Products in Twitter. *Drug Saf. En ligne* 2014 Apr 29

[16] Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, Upadhaya T, Gonzalez G. Utilizing social media data for pharmacovigilance : A review. *Journal of Biomedical Informatics*. 2015 54;202-212

[17] Levinson D. Hospital incident reporting systems do not capture most patient harm (OEI-06-09-00091) 01-05-2012 <https://oig.hhs.gov/oei/reports/oei-06-09-00091.pdf>

[18] <http://arianeinternet.conseil-etat.fr/arianeinternet/getdoc.asp?id=207192&fonds=DCE>

**Adresse de correspondance**

Professeur Marius Fieschi

OpenHealth Institute

141 rue de Rennes

75006, Paris

E-mail : [marius.fieschi@gmail.com](mailto:marius.fieschi@gmail.com)

Docteur Jean-Yves Robin

OpenHealth Institute

141 rue de Rennes

75006, Paris

E-mail [jy.robin@openhealth.fr](mailto:jy.robin@openhealth.fr)