

Partie 1 — La donnée de santé comme levier d'amélioration des systèmes actuels

L'exploitation des données de santé pour renforcer l'efficacité des systèmes de soins

Introduction

Jamais les systèmes de santé n'ont généré autant de données qu'aujourd'hui. Chaque interaction avec le système médical, qu'il s'agisse d'une consultation, d'un examen, d'une prescription ou d'une hospitalisation, produit une information numérique qui vient s'ajouter à un immense ensemble de données en croissance constante. Selon les estimations du cabinet IDC Health Insights, le volume mondial de données de santé dépassait les 2 300 exaoctets en 2020 et continue de croître à un rythme supérieur à 40 % par an [1].

Et pourtant, un paradoxe frappe les observateurs. Malgré cette accumulation exponentielle de données, plus de 80 % d'entre elles restent aujourd'hui sous-exploitées lorsqu'il s'agit de les mobiliser à des fins analytiques avancées [2]. Une très large partie de ces informations demeure confinée dans des systèmes d'information isolés, éparpillés entre établissements, administrations et acteurs privés, souvent non interopérables et difficilement exploitables dans une logique de santé publique intégrée.

Ce constat fait écho à la réflexion menée par Gilles Babinet dans *Big Data, penser l'homme et le monde autrement*, où il souligne que les données, bien au-delà de leur valeur technique, sont appelées à devenir des leviers de transformation de nos institutions et de nos modèles sociaux. La santé, en tant que bien collectif, n'échappe pas à cette dynamique.

Derrière cette accumulation de données se dessine donc un enjeu fondamental. Nous disposons d'une ressource précieuse, inédite par sa profondeur et sa diversité, mais que nous peinons encore à transformer en véritable levier d'amélioration des systèmes de soins. Comment alors faire de ces gisements de données un moteur de transformation concrète des organisations sanitaires, des parcours de soins et des politiques de santé publique ? C'est tout l'objet de cette première analyse.

Une richesse de données aujourd'hui sous-exploitée

Les sources de données disponibles dans le champ de la sante sont multiples et variées. Les dossiers médicaux électroniques rassemblent les informations cliniques fondamentales des patients : diagnostics posés, antécédents médicaux, prescriptions médicamenteuses, résultats d'examens biologiques ou d'imagerie médicale. À ces données cliniques s'ajoutent les informations administratives issues des systèmes de facturation et de remboursement, qui permettent de reconstituer avec précision les trajectoires de soins, les durées d'hospitalisation ou encore les coûts associés à chaque épisode de prise en charge.

L'explosion récente des objets connectés et des dispositifs de télésurveillance médicale a également enrichi ces bases de données de mesures physiologique collectées en temps réel. Les montres connectées, les tensiomètres intelligents, les glucomètres portables et d'autres dispositifs fournissent des données continues sur l'état de santé des patients, y compris à domicile. Par ailleurs, les progrès de la génomique et des biobanques ont permis de constituer des bases de données biologiques et génétiques d'une ampleur sans précédent. Enfin, les données comportementales et sociales issues des applications de santé, des enquêtes populationnelles ou même des données environnementales viennent compléter ce tableau extrêmement riche et multidimensionnel.

Cependant, cette abondance d'informations est aujourd'hui largement fragmentée. Les bases de données restent cloisonnées entre les hôpitaux, les professionnels de ville, les organismes d'assurance maladie, les laboratoires et les industriels de santé. L'absence d'interopérabilité technique et organisationnelle freine leur exploitation globale. Pour tenter de surmonter ces obstacles, plusieurs pays ont lancé des projets structurants visant à fédérer ces sources disparates. En France, le Health Data Hub incarne cette volonté de mutualiser et centraliser les données de santé a des fins de recherche et d'amélioration des soins [3]. Le Royaume-Uni développe de son côté la Federated Data Platform du NHS [4], tandis que les États-Unis s'appuient sur les initiatives du National Coordinator for Health IT [5].

Les premiers bénéfices tangibles de l'exploitation des données de santé

Malgré les difficultés techniques et institutionnelles, plusieurs expériences concrètes démontrent déjà les apports considérables qu'une exploitation intelligente des données de santé peut produire dans la gestion des systèmes de soins.

Sur le plan de la prise en charge individuelle des patients, l'analyse des parcours de soins permet par exemple d'identifier les ruptures de suivi, les retards de diagnostic ou encore les patients les plus vulnérables face à un risque de complication ou de réhospitalisation. Une étude menée au Danemark a ainsi démontré que l'analyse des historiques hospitaliers permettait de prédire avec une précision de 75 % le risque de réhospitalisation dans les trente jours suivant une sortie [6].

L'optimisation de l'organisation hospitalière constitue un autre champ d'application déjà largement documenté. Plusieurs établissements, comme le centre médical de l'Université Johns Hopkins aux États-Unis, ont mis en place des plateformes numériques de pilotage opérationnel en temps réel. En agrégeant les données de flux patients, de disponibilité des lits, du personnel et des équipements, ces outils permettent de fluidifier l'accueil aux urgences, de réduire significativement les délais d'attente et d'améliorer la rotation des lits hospitaliers [7].

L'exploitation des données offre également de nouvelles perspectives en matière de prévention. En croisant les bases médicales avec des données socio-économiques et territoriales, il devient possible d'identifier les zones géographiques cumulant plusieurs facteurs de vulnérabilité sanitaire, et de déployer des programmes de prévention ciblés en amont des pathologies chroniques majeures telles que le diabète, l'obésité ou les maladies cardiovasculaires. Aux États-Unis, le programme All of Us vise ainsi à rassembler les données médicales, génétiques et sociales de millions de participants afin de développer des actions de prévention et de recherche personnalisées à grande échelle [8].

En matière de veille sanitaire, la centralisation des données permet une réactivité accrue face aux menaces épidémiques émergentes. L'exemple sud-coréen, au début de la pandémie de COVID-19, illustre parfaitement cette capacité à croiser rapidement les données médicales, de mobilité et de contacts pour reconstituer les chaînes de transmission virale et déployer des actions ciblées de contrôle épidémiologique [9].

Enfin, le potentiel pour la recherche clinique et pharmaceutique est immense. Le recours aux grandes bases de données cliniques permet de simuler des essais cliniques virtuels, de valider rapidement l'apparition d'effets secondaires médicamenteux et d'identifier de nouveaux biomarqueurs d'évolution des pathologies. En Europe, le consortium EHDEN s'appuie sur des centaines de bases de données cliniques pour faciliter l'évaluation accélérée de traitements innovants dans des contextes variés [10].

Les obstacles structurels à dépasser

Pour généraliser ces usages et passer à une véritable stratégie nationale d'exploitation des données de santé, plusieurs obstacles restent toutefois à franchir.

Sur le plan technique, l'hétérogénéité des formats de données, la coexistence de systèmes non interopérables et la présence massive de données non structurées, telles que les comptes rendus médicaux rédigés en texte libre, compliquent l'analyse automatisée de ces informations. La qualité des données, marquée par des erreurs de saisie ou des champs manquants, reste également un enjeu central. Des standards internationaux, comme le format HL7-FHIR, visent justement à harmoniser les échanges de données médicales entre systèmes distincts [11].

Les difficultés organisationnelles sont tout aussi déterminantes. La multiplicité des acteurs, la fragmentation des responsabilités, les réticences au partage des données entre établissements concurrents ou entre secteurs public et privé freinent encore largement les initiatives de mutualisation à grande échelle.

L'enjeu humain n'est pas moindre. La formation des professionnels de santé aux outils d'analyse de données et aux systèmes d'aide à la décision est encore largement insuffisante. Il ne s'agit pas simplement de doter les établissements de nouveaux logiciels, mais de transformer en profondeur les compétences et les pratiques médicales quotidiennes.

Enfin, les questions éthiques, juridiques et sociales doivent impérativement accompagner ces évolutions. Le respect du consentement éclairé des patients, la protection de la vie privée et la gouvernance transparente des données collectives restent des conditions indispensables à la légitimité et à l'acceptabilité de ces transformations.

Conclusion

L'exploitation des données de santé ne relève plus aujourd'hui d'une simple promesse théorique. Les expériences pionnières, menées dans différents pays et à différentes échelles, démontrent que ces données constituent un levier très concret d'amélioration des soins, d'optimisation des ressources médicales et de personnalisation des prises en charge. Elles permettent d'imaginer un système de santé plus réactif, mieux coordonné, capable d'anticiper plutôt que de subir.

Mais comme le suggère Gilles Babin, les données n'ont pas seulement une portée opérationnelle : elles ouvrent une nouvelle façon de penser notre rapport au réel, au

collectif, à l'action publique. Dans le domaine de la santé, leur bon usage peut contribuer à inventer un système plus intelligent, mais aussi plus juste, plus accessible et plus humain. Cette transformation ne pourra aboutir que si elle repose sur des choix technologiques ambitieux, des réformes organisationnelles profondes et une réflexion éthique permanente. Ces dimensions feront l'objet des prochaines parties de cette série.

Références

Babinet, G. (2015). Big Data : penser l'homme et le monde autrement. Éditions Le Passeur.

[1] IDC Health Insights (2020). Worldwide Global Datasphere Forecast.

[2] HIMSS Analytics (2021). The State of Healthcare Data Analytics.

[3] Health Data Hub (France). Présentation de la plateforme française des données de santé.

[4] NHS Digital (2021). Data saves lives: reshaping health and care with data.

[5] Office of the National Coordinator for Health IT (ONC), U.S. Dept. of Health.

[6] Schmidt M. et al. (2021). Predicting readmission risk after hospital discharge. BMJ Open.

[7] GE Healthcare & Johns Hopkins Hospital (2020). Command Center Experience.

[8] NIH (2021). All of Us Research Program.

[9] KCDC (Korea Centers for Disease Control, 2020). COVID-19 Response Report.

[10] European Health Data & Evidence Network (EHDEN), IMI Consortium.

[11] Health Level Seven International (2020). HL7 FHIR Standard.