

Colloque Données massives et santé publique

Responsables scientifiques :

Margo BERNELIN et Sonia DESMOULIN-CANSELIER

Date :

25 avril 2019

Lieux :

Nantes : MSH Ange Guépin

Langue :

Français

Accès :

Colloque ouvert à tous les publics intéressés

Argumentaire du colloque

« La médecine de demain ne sera pas la même que celle d'aujourd'hui : prédictive, personnalisée, numérique, elle devra sans cesse s'adapter aux nouveaux enjeux et aux nouvelles technologies »¹. C'est par cette phrase que la Stratégie nationale de santé 2018-2022 présentée par le Ministère des solidarités et de la santé indique vouloir faire de l'introduction du numérique en santé et notamment des *données massives* l'avenir de la santé publique. Ce faisant, elle entend positionner la France dans un mouvement global. Le développement d'outils algorithmiques exploitant les *Big data* au service de la santé des populations constitue, en effet, une innovation phare de ce début de XXI^e siècle. Elle se combine avec la possibilité d'exploiter ensemble des bases de données publiques et des bases de données privées constituées notamment par la collecte de données issues d'objets connectés. Les données alors exploitées peuvent être très diverses : signaux biologiques, dépenses de santé, observance des traitements, facteurs environnementaux, habitudes de vie. À ces capacités nouvelles de traitement s'ajoutent encore la possibilité de stocker un nombre toujours plus important de données. Partant, les discours des acteurs publics sont positifs et énoncent que les recherches menées sur les diverses données collectées devraient permettre de proposer de nouveaux plans de prévention en santé, des traitements mieux ciblés pour les patients, et d'ouvrir des voies de recherche insoupçonnées utiles en santé publique. Par exemple, dans le domaine de la cancérologie, l'utilisation des *Big data* doit permettre d'anticiper la survenue de certains cancers en mettant en lumière de nouveaux facteurs de risques². Les données des hôpitaux sont, elles, utilisées afin d'appréhender et de suivre la dissémination des staphylocoques en milieu hospitalier pour proposer des stratégies de lutte contre la bactérie. Le recueil de données devrait encore pouvoir jouer un rôle en matière de pharmacovigilance, voire de matériovigilance, notamment pour centraliser les événements indésirables, identifier et mesurer les risques d'un médicament ou d'un dispositif médical,

¹ Stratégie nationale de santé 2018-2022, Ministère des solidarités et de la santé, Paris, 2017, p. 63.

² Pour une description voir C. MATHELIN, K. NEUBERGER et I. IBNOUHSEIN, « Cancer du sein et Big Data : évolution ou révolution ? », *Gynécologie Obstétrique et Fertilité*, 2016 (44), pp. 375-376.

conduisant à son retrait du marché³. En somme, l'introduction des données massives en santé publique renouvelerait cette discipline centenaire.

Quelles sont les implications d'une telle évolution ? Quelle reformulation du champ de la santé publique en découle ? L'introduction des données massives induit-elle une rupture ou une continuité pour les pratiques en la matière ? L'évolution décrite dans la Stratégie nationale de santé 2018-2022 renvoie résolument au passé les premières définitions de la santé publique, axées sur l'hygiène, telle que celle d'Eugène Gautrez, qui faisait d'elle « l'application aux intérêts collectifs, et non plus aux seuls individus des mesures édictées par l'hygiène en ce qui concerne ses rapports avec le monde extérieur »⁴. La santé publique est devenue par la suite « la science et l'art de prévenir les maladies, de prolonger la vie, de promouvoir la santé et l'efficacité physique à travers les efforts coordonnés de la communauté pour l'assainissement de l'environnement, le contrôle des infections de la population, l'éducation de l'individu aux principes d'hygiène personnelle, l'organisation des services médicaux et infirmiers pour le diagnostic précoce et le traitement préventif des pathologies, le développement des dispositifs sociaux qui assureront à chacun un niveau de vie adéquat pour le maintien de la santé »⁵. L'utilisation des données massives en santé publique, telle que décrite dans les discours, pourrait offrir de nouveaux outils pour parvenir aux objectifs ainsi fixés. En effet, le traitement des données permettrait, notamment, d'établir des corrélations inattendues ou sous-exploitées, lesquelles nourriraient la recherche en santé publique. On ne procéderait plus alors par confirmation (temporaire) ou infirmation d'une hypothèse déjà formulée. Le changement ainsi décrit aurait aussi pour conséquence de faire intervenir de manière renouvelée les acteurs privés dans le domaine de la santé publique. Le programme *Google Flu* en est un exemple bien connu. En 2008, le géant *Google* a lancé un programme pour anticiper l'évolution de la grippe aux États-Unis. En se fondant sur le traitement des recherches liées à la grippe sur son moteur de recherche, il devait offrir une géolocalisation des foyers d'épidémie de grippe⁶. Aussi efficace que la cartographie de l'épidémie établie par les pouvoirs publics américains, cette exploitation de données massives démontrait qu'elle avait pour ambition de faire du partage des données de chacun, un outil pour assurer la santé de tous.

Le traitement des données massives en santé publique pourrait être davantage qu'une voie méthodologique supplémentaire. En s'imposant, elle pourrait entraîner un changement de cadre de pensée et provoquer ou exprimer une modification profonde de l'objet même de la santé publique.

Un tel bouleversement – qu'il soit purement méthodologique ou plus fondamental – appelle une étude critique pluridisciplinaire. Il suscite, en effet, des interrogations sur les opportunités à saisir et des craintes sur les risques induits, notamment sur la manière d'assurer

³ Voir le cas du système I2b2 (*informatics for integrating biology and the bedside*.) J. BROWNSTEIN, S. MURPHY, A. Goldfine, R. GRANT, M. SORDO, V. GAINER, J. COLECCHI, A. DUBEY, D. NATHAN, J. GLASER, I. KOHANE, « Rapid identification of myocardial infarction risk associated with diabetic medications using electronic medical records », *Diabetes Care*, 2010 Mar;33(3):526-31.

⁴ E. GAUTREZ, « De la protection de la santé publique », Conférence donnée à Clermont-Ferrand, le 27 mars 1896. Cette définition était fort répandue à cette période. Elle donna lieu à l'adoption de la célèbre loi du 15 février 1902 qui contraint chaque municipalité à se doter d'un règlement sanitaire.

⁵ Propos de Ch-E. WINSLOX repris par D. FASSIN, « Santé Publique », dans D. Lecourt (dir.) *Dictionnaire de la pensée médicale*, Paris, PUF, 2004, pp.1014-1018. La santé publique englobe aujourd'hui des questions sociales diverses qui semblaient jusque-là étrangères à son champ tels que la sexualité, l'échec scolaire, l'obésité, les addictions, la politique de la ville, l'immigration, la vie carcérale. Ch. NOUVILLE et F. BELLIVIER, « Santé publique et nouveaux rôles du médecin en France », *Publication en droit de la santé de McGill*, 2007, vol. 1.1, pp. 45-52, p.47.

⁶ V. MAYER-SCHÖNBERGER et K. CUKIER, *Big Data : A Revolution That Will Transform How We Live, Work and Think*, Londres, John Murray Publishers, 2014, p.1-2.

le respect de la vie privée des individus dont les données sont collectées, sur la fiabilité des données exploitées par des opérateurs privés et publics, sur le contrôle social des choix méthodologiques qui auront des conséquences sur les politiques de santé publique futures. **A quelles transformations faisons-nous face, pour quels risques et pour quels bénéfices?** Telle est la question qui sera l'objet de ce colloque.

Pour en traiter, nous proposons de revenir sur des thèmes clés de la santé publique, pour interroger les continuités et les ruptures à l'œuvre.

I- Détection et prévention des maladies

*Président de session : [Pierre-Antoine GOURRAUD](#) (Immunologie et génétique, PhD MPH)

Comme le met en lumière l'exemple du programme *Google flu*, l'utilisation d'outils algorithmiques dans le traitement de données massives en santé publique, doit permettre la détection de virus et l'anticipation des épidémies. Cependant, le système ne semble pas infaillible et des craintes quant aux biais induits par les données utilisées ont conduit à la remise en question du programme et à de nouvelles propositions d'algorithmes. L'utilisation des données massives en épidémiologie permet d'interroger les conditions de collecte et de traitement des données ainsi que les implications du recours à ces nouveaux outils sur les systèmes de gouvernance et de veille sanitaire, mais aussi de renouveler la réflexion sur les rôles et fonctions des secteurs privé et public en la matière.

Du côté de la génétique, la multiplication des tests et de diagnostics précoces mène essentiellement à la découverte de prédispositions avec une grande variabilité individuelle. *In fine*, elle permet, entre autres, d'inciter les individus à mener une vie plus équilibrée selon ces variabilités individuelles. Une question émerge : quel gain pour la santé publique au regard du coût que représentent ces tests ? De plus quelles sont les incidences de ces résultats sur les stratégies commerciales et industrielles qui président au développement des tests et diagnostics précoces notamment face aux impératifs de prescription à l'œuvre (ex : les tests compagnons) ?

- Lutte contre les maladies virales et infectieuses

*Intervenante invitée (extérieur au programme DataSanté): [Lara KHOURY](#) (juriste, PR Université McGill)

*Intervenant du programme DataSanté: [Paul VERON](#) (juriste, MCF Université de Nantes)

- Prévention des maladies génétiques :

*Intervenant invité (extérieur au programme DataSanté): [Ma'n H. ZAWATI](#) (Directeur exécutif au Centre de génomique et de politiques de McGill)

*Intervenants du programme DataSanté:

- [Sandra MERCIER](#) (Génétique Médicale, CHU Nantes)
- [Emmanuelle RIAL-SEBBAG](#) (juriste, DR INSERM).

2) Les frontières de la santé publique

*Présidente de session: [Patricia LEMARCHAND](#) (Pneumologie, PU-PH Université Nantes/CHU)

La collecte des données massives et leur traitement doivent mettre en lumière de nouveaux déterminants de la santé humaine. Socle d'études comportementales en santé publique, les résultats des recherches menées sur des données massives invitent - ou contraignent - les individus à adopter des comportements vertueux pour leur santé. Cependant les corrélations entre des habitudes de vies et la santé soulèvent des difficultés : la protection de la vie privée, le contrôle privé (assurantiel notamment) ou public de ces habitudes, l'utilisation à des fins commerciales instrumentalisantes des données (par exemple en matière de nutrition)⁷, les craintes d'une *ubérisation* de la santé par l'emploi d'objets pour la collecte des données et la prévention en santé⁸. De plus, les études menées sur les données massives en santé peuvent conduire à élargir le champ des questions traitées par cette discipline. Par exemple, le taux d'endettement personnel est déjà rattaché au risque de pression artérielle élevée⁹. Le champ des pratiques sportives de haut niveau, individuelles ou en équipe, constitue également un observatoire intéressant. L'utilisation des *Big data* en la matière vise principalement à l'amélioration des performances des joueurs ou des équipes, mais impacte également et de manière incidente la gestion de leur santé (calcul du rythme cardiaque et évaluation du niveau de stress, des temps de repos nécessaires, prévention des risques de blessures, voire détection du dopage). Dès lors, tout ce qui concerne l'individu est susceptible d'être utile à la collectivité. Quels sont les angles morts ou les impensés de ce phénomène ? Quelles peuvent être leurs implications juridiques ?

- **Renouvellement des champs classiques : hygiène et éducation**

*Intervenant invité (extérieurs au programme DataSanté): **David BUCKERIDGE** (Informatique appliquée à la santé publique, PR Université McGill)

*Intervenante du programme DataSanté: **Marine FRIANT-PERROT** (juriste, MCF Université de Nantes)

- **Extension à de nouveaux champs de la santé publique ?**

*Intervenant invité (extérieur au programme DataSanté) *pressenti* : **Jean-François TOUSSAINT** (Physiologie, PR Université Paris-Descartes, Directeur de l'Institut de recherche biomédicale et d'épidémiologie du sport (Irmes)).

*Intervenante du programme DataSanté: **Margo BERNELIN** (juriste, post-doctorante DataSanté/DCS)

⁷ Par exemple, l'étude canadienne sur les liens entre les stratégies marketing de réduction des prix sur les sodas et leur influence sur les individus en fonction de leur niveau social, Hiroshi MAMIYA, Erica E. M. MOODIE, Yu MA, David L BUCKERIDGE, « Susceptibility to price discounting of soda by neighbourhood educational status: an ecological analysis of disparities in soda consumption using point-of-purchase transaction data in Montreal, Canada », *International Journal of Epidemiology*, 22 June 2018, <https://doi.org/10.1093/ije/dyy108>.

⁸ Nicolas POSTEL-VINAY fait référence à une *ubimédecine* laquelle se fonde sur « des données de santé recueillies à l'initiative de l'utilisateur en des moments et des lieux multiples », dans « De l'automesure à l'ubimédecine » dans Conseil National de l'Ordre des Médecins, Santé Connectée, de la e-santé à la santé connectée, Livre blanc, Paris, Janvier 2015, p.12. Pour une approche plus large de l'ubérisation voir C. LEQUILLERIER, « L'« ubérisation » de la santé », *Dalloz IP/IT*, 2017, p. 155 et s.

⁹ Risque utilisé par l'assureur Aviva au Royaume-Uni pour calculer les primes d'assurance santé de ses clients, V. MAYER-SCHÖNBERGER et K. CUKIER, *Big Data : A Revolution That Will Transform How We Live*, précit., p. 52.

Synthèse : [Sonia DESMOULIN-CANSELIER](#) (juriste, CR CNRS, DCS)

Déroulé de la journée : 09h- 17h00

Matin:

- 09h-09h30 : accueil-café
- 09h30: Allocution d'accueil : [Stéphane TIRARD](#) (Histoire des sciences, PR Université de Nantes, Responsable du Programme DataSanté)
- 09h40-12h30 Session 1 Détection et prévention des maladies
 - *09h40-11h40 : 5 interventions de 20 mn environ
 - *11h40-12h40 : discussion avec la salle animée par le président de session.

12h40-13h40 : pause déjeuner

Après-midi:

- 13h40-17h00 : Session 2: Les frontières de la santé publique
 - * 13h40-15h40: 4 interventions de 20 mn environ.
 - * 15h40- 15h50: pause-café
 - * 15h50- 16h40 : discussion avec la salle animée par le président de session

- * 16h40: Synthèse