

Séminaire de statistiques appliquées INED 24/09/2020 – Introduction

Bonjour à toutes et à tous, sur place dans les locaux de l'Ined aussi bien qu'à distance !

Merci à Arnaud et Bénédicte de m'avoir invitée à présider ce séminaire sur un sujet qui m'est cher, celui de l'utilisation par la recherche des données de santé, et avoir tenu bon dans leur projet malgré les difficultés de la période Covid, qui les ont amenés à repousser la séance, puis la changer de format. Cela me permet de découvrir enfin les nouveaux locaux après de nombreuses années passées à préparer, depuis la direction de l'Ined, ce projet de Campus avec tous les partenaires. L'ancien terrain industriel a bien changé...

Données de santé, de quoi parle-t-on ? Elles sont de natures très diverses (données de soins : dossiers, imagerie, données biologiques...), données construites par les chercheurs comme les cohortes...), mais nous nous intéresserons ici principalement à d'autres types de données, les données médico-administratives issues de la gestion du système de soins, regroupées désormais dans le SNDS, ainsi qu'à l'extension de ce système au sein du nouveau Health data Hub.

En réalité, les premières utilisations de ces données par la recherche sont assez anciennes, mais il faut bien dire qu'elles étaient alors limitées à quelques happy few... C'est ainsi par exemple dès 1988 que voit le jour l'enquête ESPS (santé et protection sociale), échantillonnée et appariée avec les fichiers de remboursements de soins de l'assurance maladie. Cette enquête était portée par ce qui était alors le CREDES d'Arié Mizrahi, qui deviendra l'IRDES, institut de recherche et documentation en santé, que j'ai eu aussi l'honneur de diriger pendant quelques années. L'enquête sera par la suite menée chaque année, puis tous les deux ans, jusqu'en 2014 où elle a été jumelée avec l'enquête européenne santé. Enquêtes et données appariées ont fait l'objet de nombreux travaux de recherche et l'IRDES a longtemps porté des compétences fort rares dans le monde de la recherche sur l'exploitation des BDMA.

Les BDMA telles que nous les connaissons se sont constituées peu à peu depuis les années 90.

Ainsi depuis la loi hospitalière de juillet 1991, les hôpitaux doivent décrire leur activité dans le PMSI (programme médicalisé des systèmes d'information) dont la généralisation sera le support de la future tarification à l'activité, mise en place en 2004. Au cours des années 90, plusieurs textes ont régi le codage des actes, des pathologies, des médicaments, ainsi que la dématérialisation des feuilles de soins qui sera permise par le déploiement des cartes Vitale à la fin de la décennie. La LFSS de 1998 crée le Système national d'information inter-régimes (Sniiram), qui sera mis en place par la Cnamts de 2003 à 2009. L'article 193 de la loi santé de janvier 2016 crée le Système national des données de santé (SNDS), qui va adjoindre dans l'immédiat Sniiram, PMSI et causes de décès produite par le CépiDc, et prévoir l'ajout des données sur le handicap et celles d'un échantillon d'assurés complémentaires. Dernier acte en date, la loi Buzyn crée le Health data hub, plateforme des données de santé qui vise à étendre le SNDS aux dossiers hospitaliers, registres de maladies, PMI, GIR etc., et dont l'avancement nous sera présenté lors de ce séminaire.

Depuis la fin des années 2000, suite à la mise en œuvre des premières utilisations du SNIIRAM, plusieurs rapports ont préparé ces textes et visé à promouvoir et faciliter l'accès à ces données, notamment pour la recherche et la conduite des politiques de santé publique. Plusieurs consultations ont été menées auprès des acteurs, incluant les usagers du système de soins, car les questions de protection de la vie privée sont bien évidemment essentielles sur ce sujet. Ma

recension est peut-être incomplète, mais pour celles et ceux que le sujet intéresse, vous pouvez consulter :

- deux rapports du groupe systèmes d'information du HCSP : « Pour une meilleure utilisation des bases de données administratives et médico-administratives nationales pour la santé publique et la recherche », Haut conseil de la santé publique, collection Documents, Mars 2012 et « Les systèmes d'information pour la santé publique », Haut conseil de la santé publique, collection Documents, Novembre 2009.
- en février 2010, la conférence nationale de santé organise un débat public intitulé « santé : quelles données, quels usages ? » ;
- le rapport Bras en 2013, qui vise à répondre aux premières critiques autour de l'utilisation du SNIIRAM et du PMSI à la fois sur les questions de confidentialité et de difficultés d'accès. Le PMSI, pour sa part, était la base du très débattu « palmarès des hôpitaux » publié chaque année par un grand titre de la presse nationale ;
- le rapport de la commission Open data en santé, qui prépare la loi de 2016 créant le SNDS ;
- et le rapport Polton et al., préfigurant le Health Data Hub, nouvelle plateforme d'accès aux données de santé.

Tous ces rapports et consultations vont dans le même sens : il faut faciliter l'utilisation de ces données pour la recherche, dans le respect des droits des personnes concernées.

Le séminaire d'aujourd'hui vise à éclairer l'avancement de la mise à disposition des données de santé ainsi que certaines de leurs utilisations : pour la recherche mais aussi pour la conduite des politiques publiques. L'exhaustivité des données de soins en fait par exemple un élément déterminant pour toutes les analyses territorialisées. Ces données représentent une richesse considérable en termes de contenu et de maille d'informations, mais il faut savoir les apprivoiser, apprendre à les utiliser et en connaître les forces et les limites. C'est un enjeu considérable en termes de connaissance, de méthode mais aussi d'éthique. J'espère que ce séminaire vous permettra de vous en convaincre, de vous faire réfléchir et d'ouvrir peut-être de nouveaux projets de recherche...