

Covid-19: l'AP-HP s'appuie sur son entrepôt de données de santé pour suivre l'épidémie et aider la recherche

L'Assistance publique-Hôpitaux de Paris (AP-HP) a mis en place une "task force Covid-19" au sein de son entrepôt de données de santé pour analyser et suivre l'épidémie de coronavirus mais aussi pour aider la recherche grâce à "la masse de données" produite, ont expliqué vendredi plusieurs professionnels de l'AP-HP lors d'un point presse.

Dès le début de la crise épidémique, en mars, alors que le nombre de cas de patients infectés ou suspectés de Covid-19 augmentait, l'AP-HP s'est retrouvée "confrontée à un besoin urgent d'exploitation des données", a expliqué Elisa Salamanca, directrice du département web innovation et données de l'AP-HP.

"Nous avons donc monté une *task force* data, qui était d'abord constituée de personnels de l'AP-HP: des médecins, des *data scientists*, l'équipe 'données' de la direction des systèmes d'information (DSI) et de la direction de la stratégie et de la transformation (DST), des personnels de l'Institut Imagine [l'institut des maladies génétiques de l'hôpital Necker, AP-HP] pour aider à consolider et à structurer ces données Covid-19", a-t-elle détaillé.

"Assez vite, nous avons eu des propositions d'appui par l'Institut national de recherche en informatique et en automatique (Inria) et des entreprises variées: Dataiku, OVH, BCG, Philips, Arkhn et l'appui d'une vingtaine de volontaires hors AP-HP", a-t-elle complété.

Interrogée par TICsanté, Elisa Salamanca a détaillé le profil de ces "volontaires": "il s'agit d'ingénieurs de données, de développeurs informatiques, de *data scientists* et d'ingénieurs de recherche, tous extérieurs à l'AP-HP".

Au total, cette *task force* installée au sein de l'entrepôt de données de santé (EDS) de l'AP-HP compte aujourd'hui 120 recrues ("mobilisées 7 jours sur 7 et jusqu'à 15 heures par jour") qui se concentrent aujourd'hui sur "trois grandes séries de tâches":

- l'ingénierie des données: l'accélération des données, leur enrichissement et leur consolidation
- le renforcement des infrastructures de données (qui s'appuient sur l'EDS de l'AP-HP)
- le traitement des données et la réalisation d'indicateurs pour appuyer les cellules de crise.

"La coordination de ces 120 personnes a été un défi et l'équilibre entre les personnels de l'AP-HP et les personnels extérieurs, qui n'étaient pas forcément habitués à travailler dans un environnement EDS, s'est finalement assez bien fait", a salué la directrice du département web innovation et données de l'AP-HP.

De façon générale, l'EDS de l'AP-HP compile les données de soins, d'imagerie et de biologie médicale "de 15 millions de patients" soignés dans les 39 hôpitaux du CHU francilien et "s'appuie sur le dossier patient informatisé (DPI) Orbis* (Agfa) de l'AP-HP", a rappelé Elisa Salamanca.

65.000 patients inclus dans la base de données Covid-19

"Le travail a commencé début mars par l'intégration et la réunion dans une base unique des résultats de virologie des établissements de l'AP-HP", a relaté le Dr Mélodie Bernaux, médecin de santé publique au sein de la DST de l'AP-HP.

"Le critère d'inclusion dans la base EDS-Covid-19 est la réalisation du test PCR coronavirus et les actions menées depuis nous ont permis d'associer ces résultats aux autres données des patients: données administratives, les historiques de ses venues à l'AP-HP, les données cliniques (comptes rendus et observations médicales) et données d'imagerie et de biologie médicale, les données de prescriptions médicamenteuses, les données des codages des actes médicaux et de bloc opératoire notamment...", a-t-elle détaillé.

Les données des urgences ou celles issues de la base Covidom de suivi à domicile des patients suspects de Covid-19 sont également intégrées. La plateforme compte aujourd'hui plus de 40.000 patients, [rappelle-t-on](#).

La base "EDS-Covid" inclut à date 65.000 patients dont "près de 23.000" ont été détectés positifs au Covid-19 après la réalisation d'un test PCR, a-t-on également appris. Parmi ces patients, "50% sont ou ont été hospitalisés à l'AP-HP et un quart a fait ou fait actuellement un passage en réanimation", a souligné le Dr Bernaux.

"La base a immédiatement été standardisée pour correspondre aux standards internationaux et faciliter l'exploitation des données et leur interopérabilité", a-t-elle également fait savoir.

Par ailleurs, pour répondre aux "trois grandes séries de tâches" identifiées par l'AP-HP afin de structurer et d'enrichir ces données, plusieurs groupes de travail se sont créés au sein de la *task force*.

Le Dr Bernaux a ainsi évoqué: "un groupe de traitement automatique des langues qui utilise l'intelligence artificielle (IA) pour isoler les caractéristiques des patients au moment de leur prise en charge", "un groupe de biologie qui précise les résultats des laboratoires" ou encore, "un groupe d'épidémiologie qui travaille spécifiquement à l'analyse spatio-temporelle des survenues de cas Covid-19 et à l'identification de *clusters*".

"En parallèle, un groupe de travail sur la qualité des données est également mobilisé pour qualifier et fiabiliser les données de la base", a-t-elle poursuivi.

Cette base de l'AP-HP doit ainsi servir au pilotage du CHU francilien grâce au recueil et à l'analyse de ces données en temps réel (via un tableau de bord dédié) mais aussi à "alimenter quotidiennement la base de Santé publique France".

"Ce *dashboard* (tableau de bord) permet aussi de restituer les premières explorations des travaux de recherche sur le Covid-19", a rapporté le Dr Bernaux.

Une "base Covid-19" pour accompagner la recherche

Antoine Neuraz, docteur en médecine et *data health scientist* à l'hôpital Necker (Paris, AP-HP) et le Pr Etienne Audureau, épidémiologiste à l'hôpital Henri-Mondor à Créteil (AP-HP) sont, eux, revenus sur "la masse de données précieuses pour la recherche".

"L'observation des données structurées permet de détailler le phénotypage des patients et les comorbidités", a expliqué le Dr Neuraz. "Dans le cadre du Covid-19, nous sommes dans une situation inédite pour laquelle nous avons peu de connaissances et peu de données structurées", a-t-il déploré.

Ainsi, en collaboration avec l'Inria et l'Institut Imagine, l'AP-HP a mis en place des tuyaux d'extraction de l'information clinique disponible notamment basés sur de l'IA sur plusieurs sujets, dont les prescriptions médicamenteuses, ou sur certaines données cliniques ciblées comme celles relatives à la consommation de tabac ou l'obésité. L'idée est de répondre aux questions de recherche "qui émergent au jour le jour avec ce virus", a précisé le Dr Neuraz.

Par ce biais, les chercheurs envisagent de faire "de la détection de signaux" et ils ont pu intégrer dans leur processus de recherche "plus d'un million de textes cliniques dans 20.000 points d'informations" et ces données concernent "plus de 35.000 patients".

"Plus de 30 projets de recherche ont été soumis au comité éthique de l'EDS de l'AP-HP et 90% d'entre eux reposent sur les données extraites des textes cliniques", a fait savoir Antoine Neuraz.

Dans le même ton, le Pr Audureau a salué la "plus-value" apportée par ces données et les impacts positifs pour la recherche de ces "cohortes dynamiques".

"Grâce à ces données en temps réel, nous pouvons faire de la recherche sur le temps long mais aussi sur un temps plus court pour répondre à des questions de santé publique et faire des approches ciblées", a-t-il commenté.

Parmi les thèmes abordés dans ces travaux de recherche, l'AP-HP a mis l'accent sur les outils de prédiction pour déterminer les facteurs de risques du Covid-19 et se préparer au déconfinement.

Ces travaux permettent également de caractériser les patients atteints de Covid-19 via "une fouille de données pour identifier des profils de patients auxquels les chercheurs n'auraient pas forcément pensé", a expliqué le Pr Audureau. "Il s'agit notamment de sous-groupes de patients avec des associations de comorbidités ou des trajectoires de soins particulières", a-t-il complété.

Le troisième thème de recherche abordé concerne, lui, l'évaluation de l'efficacité ou de la dangerosité thérapeutique en vie réelle. "Le juge arbitre absolu, ce sont les essais thérapeutiques randomisés de grande échelle", a souligné le Pr Audureau, rappelant "que c'est la seule façon de répondre aux questions sur l'efficacité d'un traitement, par exemple, l'hydroxychloroquine".

Les premiers résultats des travaux de recherche menés par l'AP-HP devraient être publiés "dans les prochains jours à semaines", ont précisé les deux chercheurs.

Par ailleurs, l'AP-HP participe, tout comme le CHU de Bordeaux, à un réseau international d'entrepôts de données créé à l'initiative de l'université américaine de Harvard, baptisé *Covid Clinical*, en vue de partager les données agrégées entre plusieurs pays pour faire avancer la recherche sur le coronavirus.

Wassinia Zirar

Wassinia.Zirar@apmnews.com