

Intelligence artificielle : une révolution annoncée en cancérologie

A la pointe de la lutte contre le cancer, l'Institut Curie mise sur les big data et l'intelligence artificielle, une révolution annoncée pour améliorer la recherche en cancérologie et le traitement des malades. Mieux comprendre la maladie cancéreuse dans toute sa complexité, en prédire l'évolution et mieux la combattre : un espoir pour les quelque 400 000 personnes diagnostiquées chaque année et les 3 millions de personnes en France qui vivent avec un cancer.

Le développement de l'intelligence artificielle est appelé à transformer en profondeur les pratiques en cancérologie : aide au diagnostic, aide à la décision thérapeutique, suivi évolutif du patient... dans le but de traiter plus efficacement les malades et de limiter les effets secondaires.

Algorithmes et calculs permettent à l'IA de traiter des quantités considérables de données de natures très différentes qui mobilisent des équipes multidisciplinaires : des données de génomique (les séquences d'ADN), transcriptomiques (les séquences ARN), des marqueurs sanguins, des signaux physiologiques, des données d'imagerie, des données sur l'histoire médicale du patient et ses modes de vie, parfois enregistrés à l'aide d'objets connectés... Sans compter la masse colossale des données épidémiologiques.

Cela concerne tous les types de cancers et les plus rares en particulier (cancers pédiatriques), et nécessite d'importants moyens techniques/informatiques et des expertises de très haut niveau.

Le recueil de l'ensemble de ces données est réalisé dans le respect de la réglementation en vigueur, notamment le Règlement relatif à la protection des données personnelles (RGPD) et la « Loi Informatique et Libertés ». Pour toutes les recherches impliquant les données personnelles des patients (hors caractéristiques génétiques pour lesquelles un consentement est obligatoire), ceux-ci disposent d'un droit d'opposition au traitement.

L'Institut Curie pionnier et leader en intelligence artificielle

Les grands centres hospitaliers à travers le monde ont un rôle majeur à jouer dans le développement de l'IA car ils concentrent les bases de données les plus riches et les plus proches de la réalité clinique. C'est le cas à l'Institut Curie, riche de 540 000 dossiers et près de 20 millions de documents, et précurseur dans l'informatisation des données cliniques.

L'Institut Curie est aujourd'hui pionnier dans le domaine de l'IA car il a anticipé depuis près de 20 ans sur tous les plans : création du Centre de Ressources Biologiques, informatisation des dossiers patients, investissements en bio-informatique. Depuis 2017, cette thématique a pris un nouvel essor avec la création de la direction des datas, le développement de nombreux partenariats académiques et avec des biotech et le co-pilotage de [projets nationaux](#) (tels que la plateforme ConSoRe ou le groupe OSIRIS).

Les travaux de recherche sur l'IA menés à l'Institut Curie portent notamment sur l'aide au diagnostic avec une meilleure classification des cancers pour proposer des traitements plus personnalisés. Autre axe de recherche : la prédiction de la réponse aux traitements dans le cancer du poumon et du sein, grâce à l'analyse automatique d'images radiologiques. Identifier par exemple les patients qui vont répondre à l'immunothérapie est un enjeu majeur car selon les localisations les bons répondeurs ne représentent que 10% à 40% des patients.

Objectifs de l'IA en cancérologie

- **Mieux prévenir et diagnostiquer** : meilleure classification des tumeurs et la découverte de nouveaux facteurs de risques/susceptibilité
- **Mieux traiter** : nouvelles indications et combinaisons de traitements, biomarqueurs prédictifs de la réponse au traitement, nouvelles molécules, essais cliniques mieux ciblés, suivi des effets secondaires en temps réel
- **Trouver les traitements de demain** : comprendre les résistances aux traitements, identifier de nouvelles cibles pour la médecine de précision, intégrer les données environnementales

« A terme, l'IA sera une aide pour traiter les données générées par la génomique, l'anatomopathologie, l'imagerie... et permettre de développer la médecine de précision. Mais, une chose est sûre, le médecin aura toujours le dernier mot et ce n'est pas demain qu'il sera remplacé par une machine. L'enjeu est bien d'organiser des interactions vertueuses entre l'expertise humaine et les apports de l'IA dans l'exercice quotidien de la médecine », explique le Dr Alain Livartowski, oncologue médical et conseiller du Directeur des data à l'Institut Curie. « C'est parce que l'Institut Curie met en synergie les soins et la recherche que nous pouvons mettre en place des outils bénéfiques pour les patients de l'Institut Curie et au-delà ».

Combiner séquençage nouvelle génération et IA pour des applications cliniques concrètes

L'équipe de Sarah Watson, médecin chercheuse au sein de l'équipe Génétique et Biologie des Tumeurs Pédiatriques de l'Institut Curie est parvenue à établir **une cartographie des séquences ARN de plus de 20 000 tumeurs et de tissus normaux issues de près de 40 types de cancers et de 50 tissus différents**. Chacun de ces échantillons exprimant en moyenne 20 000 gènes, la quantité d'informations obtenues est gigantesque. Pour rendre tant de données intelligibles, Sarah Watson et ses collègues ont utilisé des algorithmes et des méthodes d'IA qui ont permis d'abord de réduire les dimensions d'analyse, puis de classer les données en fonction de leur tissu d'origine. Ils disposent ainsi d'un système capable d'analyser un nouvel échantillon, de le comparer et de proposer un diagnostic.

Ces travaux montrent que l'IA peut être utilisée pour analyser des données génomiques et transcriptomiques et les interpréter cliniquement pour orienter les choix thérapeutiques, dans des délais qui sont tout à fait compatibles avec la prise en charge des patients. « Nos outils d'IA ne remplaceront pas l'anatomopathologie mais permettront de compléter, de faciliter le diagnostic et de guider la thérapeutique dans des indications qui sont complexes », précise Sarah Watson.

Gérer des volumes de données en croissance exponentielle

Il est désormais indispensable de travailler sur des grandes masses de données : recenser et faire communiquer les données dispersées dans différentes bases est un enjeu majeur, dans l'hôpital, entre la recherche et l'hôpital, entre différents établissements... C'est aujourd'hui possible grâce à des outils, des protocoles informatiques qui permettent de créer des passerelles entre les différents systèmes et de formuler les informations de part et d'autres de façon structurée.

Parallèlement aux données cliniques, l'évolution de la biologie et des sciences du vivant en général s'accompagne de volumes de données en croissance exponentielle. Le séquençage d'un génome représente 1 Tera octet de données et nécessite donc une capacité de calcul importante.

L'Institut Curie, qui a créé la plateforme francilienne SeqOIA aux côtés de l'AP-HP et de Gustave-Roussy se dote de capacités croissantes en bioinformatique. De même, l'Institut Curie participe depuis cette année à un projet européen ambitieux : [PerMedCoE](#) qui consiste à mettre en synergie les supercalculateurs européens pour optimiser l'analyse des données biologiques et médicales des patients.

Emmanuel Barillot, directeur de l'unité de bio-informatique « Cancer et Génome » de l'Institut Curie¹, qui est passée de 5 à 100 personnes en 15 ans explique : « SeqOIA permettra d'entraîner des algorithmes qui pourront améliorer en continu la prédiction des effets d'une thérapie pour le patient. Aujourd'hui, nous avons besoin d'une chaîne d'analyse des données pour en retirer les informations utiles à la prise de décision clinique dans le domaine des cancers et des maladies rares, assure Emmanuel Barillot. C'est ici que se construit la médecine de demain ».

¹ Unité « Cancer et génome : bioinformatique, biostatistiques et épidémiologie » (Institut Curie, Mines ParisTech, Inserm)

Champollion déchiffre les données médicales de l'Institut Curie

Développé par une équipe de la Direction des datas de l'Institut Curie, le **projet Champollion est un nouvel outil de machine learning** destiné à structurer des observations cliniques, à des fins de recherche, à partir des 16 millions de comptes rendus médicaux contenus dans le système d'information de l'Institut Curie. Le projet Champollion constitue ainsi une démarche originale et un véritable atout pour l'élaboration de bases de données solides et utiles en recherche clinique et translationnelle.

Interdisciplinarité et diversité au cœur des recherches en IA

Modélisations mathématiques, deep-learning, machine learning, algorithmes innovants, analyse automatique, statistiques... les équipes de l'Institut Curie s'investissent pleinement dans des projets d'IA aux applications très variées depuis la bio-imagerie, l'identification de biomarqueurs ou l'étude de cellules uniques. Depuis 2019, plusieurs équipes se fédèrent autour du projet parisien d'Institut Interdisciplinaire d'Intelligence Artificielle PRAIRIE (*PaRis Artificial Intelligence Research Institute*) porté par le CNRS, l'Inria, l'Institut Pasteur, l'Université de Paris et l'Université PSL (dont l'Institut Curie est membre associé). Grâce à la formidable quantité de données issues à la fois des soins et de la recherche dont ils disposent, les chercheurs de l'unité de bio-informatique « Cancer et génome »¹ sont fortement impliqués dans le volet « santé » de ce projet PRAIRIE. C'est le cas de Thomas Walter qui s'intéresse à la bio-imagerie. Il s'attache notamment à décrypter les processus fondamentaux en biologie (comme l'expression des gènes), à générer des outils pour étudier, comprendre et optimiser des traitements anti-cancéreux, mais aussi à apporter une aide au diagnostic en analysant des biopsies, afin de mieux prédire la réponse aux traitements médicaux.

Elle aussi chercheuse au sein de l'unité de bio-informatique « Cancer et génome »¹, Chloé-Agathe Azencott porte un autre axe de recherche et d'enseignement basé sur l'analyse d'une grande quantité de données de toutes sortes qui sont mesurées sur le génome. Son objectif : développer des méthodes pour identifier des biomarqueurs fiables, capables de détecter l'existence d'une maladie ou sa gravité, ou encore de prédire la tolérance ou la réponse à un traitement. Andrei Zinovyev, leur collègue au sein de la même unité, étudie les mécanismes de l'oncogénèse en intégrant au sein de modèles d'apprentissage toutes sortes de données moléculaires : ADN, ARN, protéines. De surcroît, il exploite les données générées sur cellule unique, qu'on sait maintenant caractériser à grande échelle grâce à la combinaison de la microfluidie et du séquençage. Les tumeurs sont alors décrites très précisément à l'échelle de la cellule individuelle, multipliant par mille ou plus le volume d'information : une mine pour l'apprentissage statistique et la caractérisation précise des différents clones de la tumeur et de son environnement comme les infiltrats immunitaires.

Focus cancers du sein : des projets IA prometteurs pour les patientes

Chaque année en France, 59 000 femmes sont touchées par un cancer du sein. Avec près de 7 500 patientes en cours de traitement en 2019, l'Institut Curie est leader dans la prise en charge de cette pathologie et est très dynamique en matière de recherche, notamment dans le domaine de l'IA.

« Nous sommes extrêmement enthousiastes de mener à bien nos projets de recherche en IA et nous sommes convaincus que les algorithmes modifieront profondément nos pratiques et notre quotidien dans les années à venir » déclare **Anne Vincent-Salomon, cheffe du Pôle de médecine diagnostique et théranostique de l'Institut Curie**. « La digitalisation des lames de biopsies mammaires dont nous disposons à l'Institut Curie est une source de données inouïe pour des analyses complexes ». En collaboration avec les chercheurs de l'unité de bio-informatique « Cancer et génome » (†) (Thomas Walter et Tristan Lazard), l'équipe d'Anne Vincent-Salomon, grâce à l'IA, décrypte certaines anomalies génétiques dans ces coupes numérisées. Plus concrètement, des travaux sont en cours tentent de prédire la présence ou l'absence d'un marqueur essentiel dans les cancers du sein triple négatif : l'HRD (pour Homologous recombination deficiency).

D'autres travaux d'IA portent sur l'analyse des images médicales, issues notamment des IRM ou des TEP-scans. **Irène Buvat est directrice du Laboratoire d'Imagerie Translationnelle en Oncologie (Inserm) à l'Institut Curie** ; elle explique : « en IRM, par exemple, nous sommes en train d'étudier comment l'IA peut permettre de prédire la réponse ou non à une chimiothérapie néo-adjuvante (administrée avant la chirurgie) dans les cancers du sein ». Les chercheurs développent des algorithmes pour scruter les caractéristiques des tumeurs primaires (texture, bords, hétérogénéité...), pour les corrélés à d'autres paramètres pathologiques et in-fine, prescrire - ou non - cette chimiothérapie qui demeure un traitement lourd. En utilisant les TEP-scans, un autre projet d'IA démarre dans les cancers du sein. Son objectif : prévoir l'atteinte ganglionnaire des patientes à partir des caractéristiques de la tumeur primaire. A terme, il s'agirait d'éviter à certaines patientes une chirurgie invasive (curage ganglionnaire). « Toutes les méthodes d'IA consistent à apprendre à partir de données. En croisant des informations cliniques, d'imagerie, anatomopathologiques, moléculaires, génomiques, protéomiques... nous essayons de construire des modèles qui décrivent mieux la complexité de la maladie, qui permettent de mieux prédire son évolution et de proposer une stratégie thérapeutique la plus adaptée et efficace pour chaque patient » conclut Irène Buvat.

A propos de l'Institut Curie

L'Institut Curie, 1er centre français de lutte contre le cancer, associe un centre de recherche de renommée internationale et un ensemble hospitalier de pointe qui prend en charge tous les cancers y compris les plus rares. Fondé en 1909 par Marie Curie, l'Institut Curie rassemble sur 3 sites (Paris, Saint-Cloud et Orsay) 3 600 chercheurs, médecins et soignants autour de ses 3 missions : soins, recherche et enseignement. Fondation privée reconnue d'utilité publique habilitée à recevoir des dons et des legs, l'Institut Curie peut, grâce au soutien de ses donateurs, accélérer les découvertes et ainsi améliorer les traitements et la qualité de vie des malades. *Pour en savoir plus : curie.fr*

CONTACTS PRESSE :

Elsa Champion – 07 64 43 09 28 - elsa.champion@curie.fr

Myriam Hamza – 06 45 87 46 51 - myriam.hamza@havas.com