

Quelles sont les images cérébrales stratégiques qui pourraient accélérer les progrès de la recherche clinique ?

Par Jean-François MANGIN

Université Paris-Saclay, CEA, CNRS, Baobab, Neurospin

Les pathologies du cerveau constituent un problème de société sans précédent. La neuro-imagerie est une des composantes-clefs d'une médecine du futur personnalisée, fondée sur la stratification des patients en groupes homogènes et caractérisés par une signature incluant des atteintes typiques au cerveau. Les images du cerveau décèlent les prémisses des pathologies à un stade où l'on peut espérer les contrecarrer. Elles peuvent aider la recherche clinique au travers d'un phénotypage profond, des images très détaillées contribuant à la compréhension de la physiopathologie, mais surtout au travers d'un phénotypage large, engendrant des bases de données qui permettront à l'IA de faire émerger ces groupes homogènes. Ces images stratégiques nécessiteront une mutualisation d'infrastructures dédiées et, à terme, la transformation du parc d'imageurs cliniques en un grand instrument permettant de suivre la population. Cette transition s'apparente à celle vécue par le monde de la physique au XX^e siècle.

On considère qu'un tiers de la population sera concerné un jour par une pathologie associée au cerveau. Seule une petite partie de ces pathologies donne lieu à une prise en charge satisfaisante. Trop souvent, nous ne savons pas les guérir et nous en ignorons même l'origine. Les succès récents de l'intelligence artificielle (IA) ont fait miroiter une myriade de révolutions à venir dans le domaine de la santé. Que peut-on en espérer en ce qui concerne les maladies de notre cerveau ? L'image d'Épinal souvent véhiculée dans la presse économique est qu'il suffit de trouver le gisement de données adéquat pour décrocher le jackpot. Existe-t-il de tels gisements en ce qui concerne les pathologies cérébrales ? Comment pourrait-on en générer au plus vite ? L'impact sociétal de ces pathologies est tel qu'il est crucial de tenter de répondre à ces questions, même si le bien-fondé de cette image d'Épinal simpliste peut être discuté.

Les données qui vont ensemencer les bouleversements escomptés ne sont pas uniquement celles classiquement utilisées par le système de santé. Nos téléphones abritent des applications de santé ou de *fitness* qui quantifient les caractéristiques de nos déplacements, voire de nos tremblements, avec des capteurs dédiés (accéléromètres, gyroscopes). Et cette supervision continue ne va pas se limiter à des aspects physiologiques. Des études épidémiologiques insèrent aujourd'hui dans nos téléphones des « mouchards bienveillants », qui décryptent la manière dont nous utilisons nos applications favorites pour déceler

les premiers signes de nos difficultés cognitives. Ces informations multifacettes pourraient permettre de détecter les prémisses d'une pathologie avant même que ses symptômes explicites ne le soient. La médecine du futur qui en découlera, fondée sur la prédiction et la prévention, semble inéluctable.

Néanmoins, quand bien même le monde du *Big Data* viendrait à faire émerger de nouveaux outils de diagnostic permettant d'espérer contrecarrer les maladies à leur tout premier stade, il reste à mettre au point des thérapies adéquates, ce qui reste du ressort des neurosciences. Les technologies d'observation du cerveau sont innombrables. La dernière décennie a d'ailleurs donné lieu à des investissements hors normes pour en augmenter la portée. La Brain Initiative américaine a ainsi engendré des technologies de l'ordre de la science-fiction pour réaliser le *reverse engineering* du cerveau des rongeurs : par exemple, la possibilité de réaliser simultanément l'enregistrement de l'activité d'un million de neurones, ou de cartographier la connectivité synaptique entre un grand nombre de neurones. Ces technologies invasives révolutionnent les neurosciences fondamentales, mais les rongeurs sont malheureusement de piètres modèles pour les maladies cérébrales humaines. Et même si les possibilités d'exploration du cerveau humain sont beaucoup plus restreintes, l'essor de l'imagerie cérébrale suscite, heureusement, beaucoup d'espairs.

Le Human Brain Project (HBP), fleuron de la communauté européenne, développe, quant à lui, une nouvelle génération d'infrastructures informatiques dédiées aux neurosciences, avec, comme point de mire, un boom de l'utilisation des simulations de la dynamique cérébrale, tant à l'échelle des réseaux de neurones qu'à l'échelle macroscopique. Pour nourrir ces simulations, le HBP a lancé un programme de collecte de données stratégiques, dont un pan conséquent est dédié au cerveau humain. Ce programme concerne à la fois la dimension « phénotypage profond », au travers de quelques images hors norme sur le plan de leur richesse, et le « phénotypage large », c'est-à-dire l'accumulation d'images cérébrales concernant le plus grand nombre possible d'individus.

Phénotypage profond

Le programme lié au phénotypage profond comporte un premier volet qui consiste à faire passer à l'âge du numérique les technologies de microscopie *post mortem*. C'est à Jülich, en Allemagne, que des milliers de sections 2D issues de microscopes sont transformées en une image 3D, un challenge méthodologique qui nécessite l'appui de grands centres de calcul et des techniques d'IA dédiées. Il est aujourd'hui possible, sur la plateforme EBRAINS issue du HBP, de naviguer en ligne dans une image dont la résolution est de 20 microns, et ce dans les trois dimensions de l'espace, avec un outil inspiré de Google Map. Le prochain objectif est d'obtenir une résolution de 1 micron isotrope, donnant lieu à une image dont la taille défie l'entendement et qui permettra d'observer chacun des 100 milliards de neurones constituant le cerveau. Pour réaliser ce défi, une « ferme » de microscopes dernier cri a été créée. Une stratégie similaire utilisant un microscope en lumière polarisée (PLI) vise à générer une image 3D des axones de la substance blanche d'un cerveau humain, à des niveaux de résolution analogues.

Des équipes françaises localisées à Neurospin, un centre de neuro-imagerie implanté au sein de l'Université Paris-Saclay, contribuent à ce programme en poussant la technologie IRM dans ses retranchements. Plus de 10 000 heures d'acquisitions réalisées avec un champ magnétique de 11,7 Tesla (T) sur un cerveau *post mortem* ont engendré une image de la connectivité mesoscopique de ce cerveau à l'échelle de 200 microns, une première mondiale. Un autre projet mené dans ce même centre, mais *in vivo*, consiste à réaliser un très grand nombre d'expériences d'imagerie fonctionnelle chez un même individu, pour cartographier la totalité des systèmes cognitifs au sein de son cerveau. Ce projet bénéficie de l'implication d'une douzaine de sujets qui auront participé à terme à une cinquantaine de sessions IRM réalisées avec un champ magnétique de 3 T. Ce sont en quelque sorte les « spationautes » du HBP. À Marseille, le phénotypage profond *in vivo* est réalisé en multipliant les modalités d'imagerie. L'IRM à 7 T décrit à une résolution millimétrique l'anatomie, la connectivité et l'organisation fonctionnelle du cerveau, alors que l'imagerie TEP en décrit le métabolisme. D'autres équipes du Sud de la France organisent à l'échelle européenne la collecte d'enregistrements de la dynamique

cérébrale, réalisés à l'aide d'électrodes implantées dans le cerveau de patients épileptiques – avec leur accord, bien sûr – dans le cadre de la procédure chirurgicale visant à les soigner.

Le phénotypage profond *post mortem* est pour l'instant réservé à l'étude de quelques cerveaux seulement et vise à nourrir la recherche fondamentale. Il n'est d'ailleurs pas spécifique au HBP : l'Institut Allen de Seattle a ainsi réalisé à partir de six cerveaux une cartographie de grande valeur de l'expression des protéines. Mais ces preuves de concept pourraient dans le futur susciter l'envie de passer à une autre échelle, sur un échantillon varié de cerveaux normaux mais surtout pathologiques, pour tenter d'éclairer d'un jour nouveau les anomalies qui en perturbent le fonctionnement. De grandes infrastructures dédiées à l'acquisition d'images sont ainsi en train d'émerger dans le monde, à l'instar de ce qui s'est passé dans le domaine de la physique des particules, il y a longtemps. Par exemple, alors que les laboratoires classiques n'utilisent qu'un seul microscope pour étudier le cerveau des souris, le nouvel institut chinois HUST-Suzhou dispose, quant à lui, de cinquante machines automatisées pour prendre des photos haute définition de chaque section et les reconstituer dans une image 3D. Certaines de ces installations émergentes sont trop coûteuses pour que chaque pays puisse en disposer. Elles seront amenées à devenir des plateformes de services auxquelles les scientifiques du monde entier pourront accéder, de la même manière que les astronomes partagent le temps d'utilisation des télescopes. La ferme de microscopes établie à Jülich et dédiée au cerveau humain pourrait ainsi devenir un jour une installation à haut débit mise au service de la communauté.

Les grandes infrastructures de recherche mutualisées dédiées aux images stratégiques du cerveau ne sont pas cantonnées à l'imagerie *post mortem*. Le caractère stratégique peut d'ailleurs reposer principalement sur des espoirs de l'ordre de la sérendipité. Les grands instruments scientifiques, lorsqu'ils sont uniques en termes de sensibilité ou de résolution, sont en effet construits pour ouvrir de nouveaux « espaces de découverte ». Et les découvertes les plus importantes réalisées à l'aide de ces instruments sont souvent celles qui n'avaient pas été prévues dans le dossier scientifique initial. Ainsi, le CEA a décidé de s'appuyer sur le savoir-faire de ses physiciens à l'origine des aimants du CERN pour concevoir une nouvelle génération d'IRM. L'aimant de 11,7 Tesla (voir la Figure 1 de la page précédente) localisé au Neurospin devrait, par exemple, permettre de zoomer pour accéder à de petites structures cérébrales hors d'atteinte actuellement, comme de petits noyaux de neurones impliqués dans la maladie de Parkinson que leur richesse en fer rend visibles à haut champ magnétique. Cet imageur hors norme va également permettre de faire « résonner » d'autres espèces chimiques que les protons de l'eau actuellement à l'origine de 99,99 % des images IRM, pour, par exemple, cartographier le lithium, un des plus anciens neuroleptiques, lequel est efficace dans le traitement du trouble bipolaire, mais dont on ignore l'action biologique.



Photo © CEA-Patrick Dumas

Figure 1 : L'aimant Iseult 11,7 T de 130 tonnes lors de son arrivée à l'Institut Neurospin.

Phénotypage large

Le programme du HBP lié au phénotypage large doit nourrir la recherche en IA évoquée au début de cet article. Il ne vise pas en premier lieu la conception d'outils d'aide au diagnostic, mais plutôt la possibilité de redéfinir le paysage des pathologies cérébrales, en s'appuyant notamment sur la neuro-imagerie. Il est fondé sur un constat d'échec en ce qui concerne la recherche clinique actuelle, qu'il attribue à l'hétérogénéité des syndromes psychiatriques ou neurologiques. Souvent, deux tableaux cliniques similaires peuvent avoir des causes très différentes, car ces syndromes agrègent probablement plusieurs maladies, d'où une grande confusion dans les essais thérapeutiques. Pour mener à bien un tel programme, il faut réunir un très grand nombre d'images, ce que le HBP compte réaliser en agrégeant les bases existantes au sein d'une centaine d'hôpitaux européens de premier plan. Pour ce

faire, ses informaticiens ont créé une infrastructure distribuée dédiée aux techniques d'apprentissage fédéré qui permettent d'accéder aux images *via* des silos sécurisés localisés dans chacun de ces hôpitaux.

Bien que l'infrastructure informatique du HBP soit désormais opérationnelle, celle-ci n'est pas encore nourrie par de grandes quantités d'images. Le diable est dans les détails. Les données acquises en soins courants sont rarement curées et organisées de manière adéquate pour pouvoir être intégrées simplement dans ce type d'organisation. En outre, mettre en place dans chacun de ces hôpitaux l'équivalent de ce que l'on appelle aujourd'hui en France un entrepôt de données de santé, nécessite du temps sur le plan administratif et des investissements financiers bien supérieurs à ce qui avait été naïvement provisionné. Mettre les masses de données de soins courants acquises par le passé à la disposition de la recherche n'est pas loin d'être

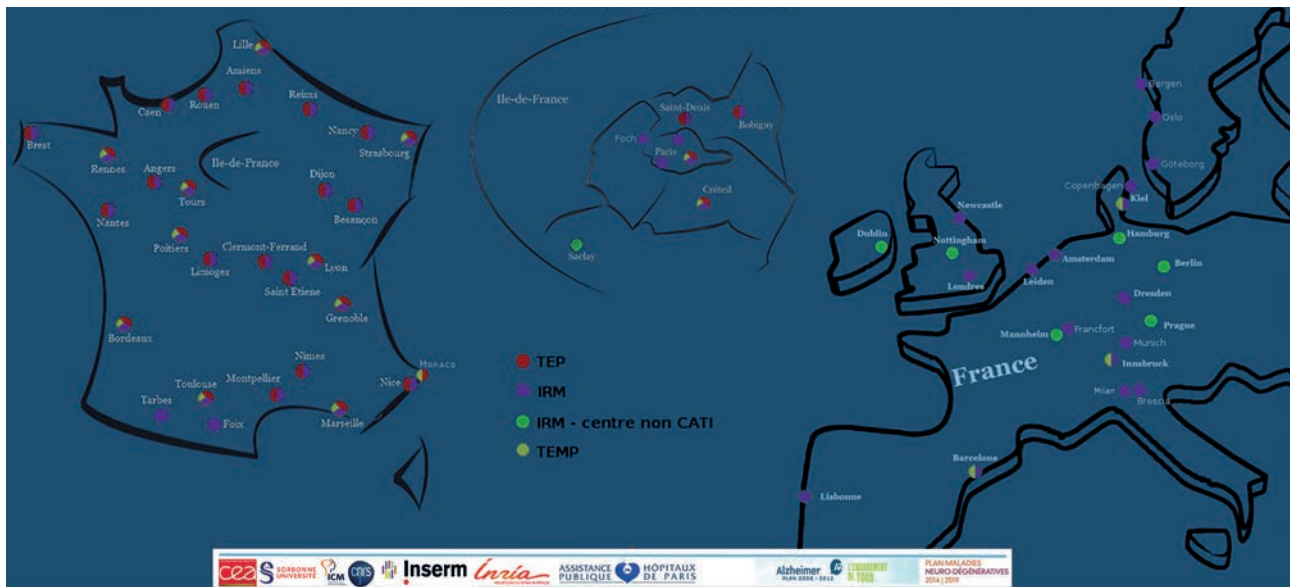


Figure 2 : Un réseau de scanners harmonisé pour alimenter l'apprentissage automatique (©CATI).

une gageure. Espérons que le Health Data Hub français, qui n'a pas à gérer la complexité liée au côté international de la fédération des données, sera plus efficace.

À court terme, le projet de stratification des populations de patients, qui est aujourd'hui consensuel au sein de la communauté clinique, passera plus probablement par l'exploitation de grandes bases de données acquises dans un contexte de recherche, avec les moyens humains nécessaires pour en maximiser la qualité. Un des handicaps des images issues des soins courants est qu'elles proviennent d'une grande variété de matériels d'imagerie et sont acquises sans standardisation, ce qui crée beaucoup d'hétérogénéité. Il est donc très difficile d'y distinguer des groupes de patients homogènes. En revanche, les images issues de grandes cohortes épidémiologiques bénéficient d'une standardisation des acquisitions. La plus grande d'entre elles, UKBiobank, imagera à terme 100 000 volontaires, en recourant à trois imageurs dédiés, pour une étude sur le vieillissement. Elle est déjà à mi-chemin et met ses données à disposition de la communauté. Avec le temps, grâce au suivi des volontaires, elle permettra de découvrir des biomarqueurs très précoces des pathologies du vieillissement et de travailler sur la problématique de la stratification de la population.

Il est aujourd'hui clair que la stratégie qui consiste à associer les meilleurs experts et le meilleur matériel pour générer de grands jeux de données stratégiques ouverts à la communauté scientifique est primordiale. Mais elle ne couvre pas tous les besoins. Des études cliniques plus ciblées resteront nécessaires, et il serait dommage qu'elles ne soient pas conçues de manière à maximiser leur réutilisation par cette communauté. En outre, les biomarqueurs mis au point grâce aux données de grande qualité de la recherche doivent faire l'objet d'une transposition pour servir au parc hétérogène d'imageurs du système de santé. C'est pour répondre à ces besoins qu'une plateforme de

collecte dédiée aux études multicentriques a été créée en France (<http://cati-neuroimaging.com>). Cette plateforme appelée CATI s'est appuyée sur des experts disséminés sur tout le territoire pour construire de manière itérative ce que les physiciens appelleraient une grande infrastructure dédiée à l'imagerie des cohortes, c'est-à-dire un réseau constitué d'une centaine de scanners (voir la Figure 2 ci-dessus) dont les paramètres d'acquisition sont réglés de manière à minimiser les effets de site. Cet instrument distribué s'apparente aux réseaux de capteurs utilisés en sismologie ou pour prédire le climat. Aujourd'hui, ce réseau est surveillé en permanence pour réagir à toute mise à jour matérielle ou logicielle. Les images sont collectées par un service Web sécurisé et leur qualité est contrôlée dans un centre centralisé. Plus de quarante études ont fait appel à ses services, sur un large spectre de pathologies. La base de données multi-pathologies harmonisée qui en résulte, forte de 10 000 patients, est sans équivalent dans le monde. Le CATI espère la mettre progressivement à disposition des chercheurs en IA.

S'il semble aujourd'hui plus facile de nourrir l'IA avec les images issues de la recherche clinique, l'infrastructure ultime sera probablement le réseau national des scanners cliniques. Ce réseau devra être harmonisé et surveillé pour permettre une utilisation optimale des biomarqueurs, mais aussi pour alimenter la communauté des chercheurs avec des bases de données représentatives de la population et des dernières versions des imageurs. Le savoir-faire accumulé par la plateforme CATI pourrait permettre de réaliser, en synergie avec les radiologues et les médecins nucléaires, une preuve de concept de standardisation de l'imagerie clinique, qui pourrait d'ailleurs associer un contrôle des paramètres d'acquisition et des corrections *a posteriori* issues de techniques d'IA en train d'émerger. Cette standardisation créerait le terreau adéquat pour que des acteurs économiques proposent des outils d'aide à la décision dans un contexte maîtrisé.