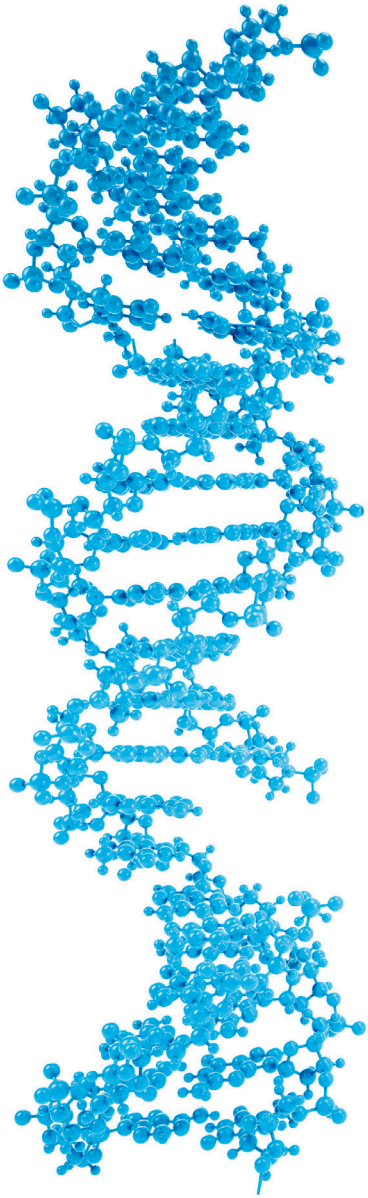
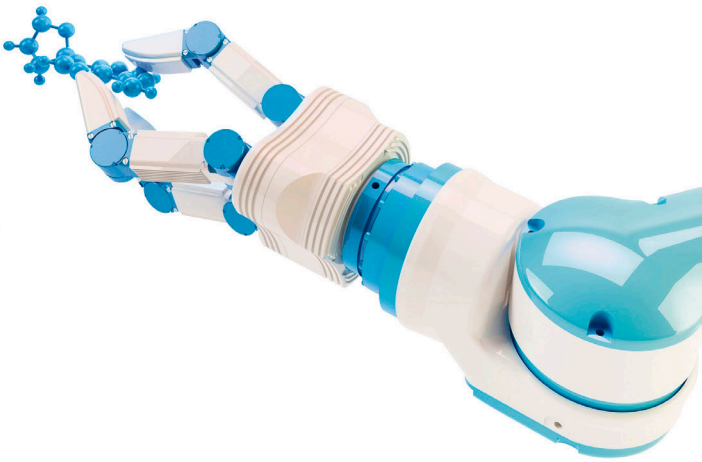
Sous la direction de

BERNARD NORDLINGER
et CÉDRIC VILLANI

Santé et intelligence artificielle





CNRS EDITIONS

Présentation de l'éditeur



Renouveler et perfectionner l'interprétation des images, modéliser la croissance des tumeurs, prédire la réponse aux traitements administrés, augmenter les performances en radiologie, en anatomie pathologique, en dermatologie, tirer parti des données génétiques pour comprendre les maladies, développer la médecine de précision, constituer des recueils de données d'une richesse jusqu'ici inaccessible : toutes choses qui, en médecine, deviennent possibles grâce à l'intelligence artificielle.

Cette révolution des pratiques suppose un renouvellement de la formation des professionnels de santé, et pose un certain nombre de défis : défi technique et technologique posé par la taille des données, défi éthique et légal, pour protéger les données personnelles, défi de la confiance, pour convaincre les acteurs de surmonter leurs réticences.

Mathématiciens modélisateurs et analystes, informaticiens, statisticiens, cancérologues, chirurgiens, onco-généticiens, sociologues, administrateurs hospitaliers, juristes, ont apporté leur contribution pour offrir au lecteur une information claire, objective et raisonnable, loin des fantasmes que le sujet peut susciter.

Cet ouvrage est né de la rencontre d'un mathématicien, Cédric Villani, attaché à montrer les promesses de l'intelligence artificielle, avec un chirurgien chercheur, Bernard Nordlinger, qui cherche à améliorer le traitement des cancers par d'ambitieux essais thérapeutiques. Il reflète les travaux du groupe de travail « Intelligence artificielle et santé » créé par l'Académie nationale de Médecine et l'Académie des Sciences.

© Andrzej Wojcicki /
Science Photo Library / Cosmos.

Maquette : © SYLVAIN COLLET

© CNRS Éditions, Paris, 2018
ISBN : 978-2-271-12291-9

Sommaire

Avant-propos, *Daniel Couturier et Catherine Bréchignac* 7

Introduction, *Bernard Nordlinger et Cédric Villani* 11

L'intelligence artificielle et la santé de demain
Cédric Villani et Bertrand Rondepierre 17

Advancing Healthcare Through Data-Driven Medicine and Artificial Intelligence,
Ran D. Balicer et Chandra Cohen-Stavi 29

AI a Vector for Positive Change in Medicine,
Daniela Rus 39

Partie 1 : Bases de données 45

Apprentissage automatique et données massives en santé
Emmanuel Bacry et Stéphane Gaïffas 47

L'appariement d'une cohorte épidémiologique aux bases de données médico-administratives
Marie Zins et Marcel Goldberg 59

Les données médico-administratives de l'Assurance maladie :
Claude Gissot 75

Comment la base SNIIRAM – PMSI permet l'étude des pratiques en chirurgie
Bertrand Lukacs 85

Les bases de données des hôpitaux : l'entrepôt de données de l'AP-HP
Christel Daniel et Élisa Salamanca 93

Les bases de données des hôpitaux : l'expérience de l'Institut Curie
Alain Livartowski 111

La donnée intelligente, terre rare de l'intelligence artificielle
Michèle Arnoe 121

Anticiper la création de connaissances dès la conception des systèmes d'information
Jean-Yves Robin 131

Partie 2 : Aide au diagnostic et au traitement 139

L'IA pour aider à choisir le bon traitement par le praticien
Christine Duval et Georges Uzbelger 141

L'IA pour aider au choix du traitement et du diagnostic. Commentaire
Jacques Rouessé et Alain Livartowski 145

L'imagerie médicale à l'heure de l'intelligence artificielle
Nicholas Ayache 151

L'intelligence artificielle dans l'imagerie médicale
Johan Brag 155

Vers une tour de contrôle des blocs opératoires ?
Nicolas Padoy 171

Apprentissage statistique en grande dimension et application au diagnostic oncologique par radiomique
Charles Bouveyron 179

La modélisation mathématique des tumeurs
Thierry Colin 191

Vers un radiologue augmenté
Mostafa El Hajjam 209

Imagerie fonctionnelle pour la personnalisation du traitement en oncologie
Dimitris Visvikis et Catherine Cheze Le Rest 215

Image virtuelle et anatomopathologie
Cécile Badoual 225

Partie 3 : Recherche 241

Intelligence artificielle et génomique des cancers
Jean-Philippe Vert 243

Après le génome, l'environnementome ?
Alain-Jacques Valleron et Pierre Bougnères 259

Intelligence artificielle appliquée a l'oncologie
Jean-Yves Blay, Jurgi Camblong et François Sigaux 273

Modélisation physique et mathématique en oncologie : des exemples
Martine Ben Amar 293

L'intelligence artificielle pour la recherche médicale et pharmaceutique
Gilles Wainrib 307

L'Open Science et l'Open Data, accélérateur de la recherche scientifique
Mehdi Benchoufi et Olivier de Fresnoye 319

Partie 4 : Données 331

Les données : protection, conditions juridiques d'accès et de traitement
Jeanne Bossi-Malafosse 333

La CNIL et l'analyse des projets de Big data en santé
Matthieu Grall 347

Vision internationale des Big data
Laurent Degos 353

Partie 5 : Humain 371

Les questions de santé publique sont-elles en danger d'information ?
Gérald Bronner 373

Robots sociaux et affectifs : une intelligence artificielle sans conscience mais utile
Laurence Devillers 383

Augmenter l'homme, et après ?
Guy Vallancien 395

L'intelligence artificielle, vecteur d'évolution des métiers de la santé et de l'organisation de l'hôpital
Jean-Patrick Lajonchère 401

Les auteurs 411

Cet ouvrage a été réalisé avec l'aide précieuse de Nicholas Ayache, Emmanuel Bacry, Jean-Patrick Lajonchère et Jean-Philippe Vert.

Avant-propos

Daniel Couturier, Secrétaire perpétuel de l'Académie nationale de Médecine et Catherine Bréchignac, Secrétaire perpétuelle de l'Académie des Sciences

Cet ouvrage est le résultat d'un rencontre, celle d'un mathématicien médaille Fields, Cédric Villani, et d'un brillant chirurgien chercheur, Bernard Nordlinger. *A priori*, rien ne laissait pressentir leur désir de se rapprocher pour travailler ensemble : l'un faisait rayonner les mathématiques à la tête de l'Institut Raymond Poincaré, l'autre s'efforçait d'améliorer le pronostic des cancers par d'ambitieux essais thérapeutiques.

Leur rencontre ne fut cependant pas fortuite. Leurs trajectoires se croisèrent à l'occasion d'une manifestation de soutien à l'innovation et à la recherche. Bernard Nordlinger, dont les préoccupations s'étendaient au-delà de l'acte chirurgical, disposait d'une large expérience des essais thérapeutiques contrôlés. Il n'ignorait rien des exigences méthodologiques et des moyens les plus appropriés d'exploitation des résultats. Cédric Villani, au-delà de l'univers des mathématiques, portait son regard vers les avancées qu'elles allaient permettre dans la plupart des domaines et notamment en santé.

Leur complémentarité était évidente, il fallait prévoir les moyens d'établir des échanges entre ceux qui accumulaient les données et les acteurs de terrain. Un groupe de travail commun à l'Académie des sciences et l'Académie de médecine est constitué, associant les

membres les plus qualifiés des deux académies, entourés de nombreux experts qui se réunissent pour s'informer, écouter les témoignages, échanger les idées.

Une réunion commune des deux académies sur le thème « Mathématiques, mégadonnées et santé, l'exemple du cancer » a eu lieu le 28 novembre 2017. En présentant successivement quelques exemples d'exploitation des bases de données de l'assurance maladie, les résultats déjà obtenus en imagerie et en génétique, ils nous ont apporté une première occasion de faire apparaître les perspectives qu'allaient offrir les mathématiques au progrès des sciences et techniques en santé.

Les caractéristiques des sciences de la santé, leurs « plaies structurelles », selon l'expression de Cédric Villani, et précisément la complexité des phénomènes, l'approximation des conclusions, la lourdeur des financements, le biais interprétatif humain, doivent trouver par la mobilisation des sciences mathématiques une nouvelle dynamique de progrès. En rendant possible l'exploitation des données massives et plus généralement en faisant appel à l'intelligence artificielle, la médecine amplifiera ses moyens d'action. Aucun domaine d'activité ne pourra échapper à cette évolution et on peut entrevoir la résonance toute particulière que l'intelligence artificielle va entraîner dans le domaine de la santé.

Renouveler et perfectionner l'interprétation des images, augmenter les performances en radiologie, en anatomie pathologique, en dermatologie, tirer parti des données génétiques et développer la médecine de précision devient possible grâce à l'intelligence artificielle. Elle permettra un recueil de données d'une richesse inaccessible au plus soigneux des examens cliniques. Elle apportera une contribution irremplaçable au choix diagnostique et thérapeutique.

Une telle évolution suppose un renouvellement de la formation des médecins et plus largement des professionnels de santé. Un tel bouleversement des pratiques oblige aussi à considérer l'évolution des responsabilités et leurs conséquences juridiques. Si le robot et l'algorithme acquièrent une telle place dans la décision, auront-ils une personnalité juridique ? Les espoirs suscités par l'intelligence artificielle provoquent des interrogations voire des inquiétudes. Ces préoccupations sont justifiées. La précaution élémentaire est de se garder

d'aller vers les prévisions fantasmatiques de « l'homme augmenté » qui, sur la base d'extrapolations aventureuses, laissent entrevoir une évolution totalement incontrôlée.

Il faut remercier Bernard Nordlinger et Cédric Villani d'avoir obtenu les participations des meilleurs experts et réuni une suite d'articles solidement argumentés, clairs qui permet aux lecteurs de s'informer posément, sans réserve et sans passion sur les perspectives du développement de l'intelligence artificielle en santé.

Souhaitons que cet ouvrage apporte à ses lecteurs la conviction qu'« il n'est pas question de remplacer le médecin par la machine et que l'enjeu est bien d'organiser les interactions vertueuses entre l'expertise humaine et les apports de l'intelligence artificielle dans l'exercice quotidien de la médecine[1] ».

1. Villani C., *Donner un sens à l'intelligence artificielle*, p. 197.

Introduction

Bernard Nordlinger et Cédric Villani

Révolution pour les uns (de plus en plus nombreux), effet de mode pour les autres (de moins en moins nombreux), l'intelligence artificielle (IA), si mal nommée, entre dans nos vies. Tous les secteurs se préparent à ressentir les bouleversements qu'elle induit.

Au premier rang des secteurs en attente d'impact, se trouve la santé. D'une part, des plaies structurelles grandes ouvertes : erreurs de diagnostic, variabilité des situations, faillibilité des experts, grandes difficultés à transmettre les informations de la recherche vers les praticiens. D'autre part, une IA qui excelle à digérer des monceaux de littératures, à dénicher des corrélations rares, à analyser les images et autres données, toujours plus nombreuses, que produit la médecine. Et bien sûr, un secteur où les enjeux sont littéralement vitaux.

Cet ouvrage s'efforce de faire le tour du sujet de l'IA en médecine, et plus généralement du croisement mathématique/informatique/médecine. Il a l'ambition d'intéresser les experts de l'IA en leur offrant du recul et une vision globale, mais également les non spécialistes curieux de s'informer sur ce sujet d'actualité. Il s'efforce de fournir une information claire, objective et raisonnable sur la question, se gardant des fantasmes que le sujet peut susciter. Il privilégiera le regard kaléidoscopique sur l'approfondissement spécialisé.

Cette compilation s'est construite progressivement, tout au long du groupe de travail « intelligence artificielle et santé » créé par l'Académie nationale de Médecine et l'Académie des Sciences sur notre suggestion. Les séances régulières du groupe de travail étaient l'occasion pour nous d'approfondir les questions que nous nous étions posées lors de notre rencontre, au Conseil éthique du concours Épidemium. Elles étaient surtout un moyen d'inviter toutes sortes de spécialistes reconnus à apporter leur contribution et à collaborer : mathématiciens modélisateurs et analystes, informaticiens, statisticiens, cancérologues, chirurgiens, onco-généticiens, sociologues, administrateurs hospitaliers, juristes…

En effet, c'est dans le partage de compétences que se construit l'IA. Ce dialogue interdisciplinaire que nous avons eu la chance de pouvoir construire sous les auspices des Académies, il faudra, demain, le mettre en place dans nos formations, sur un mode expérimental et sans surseoir.

Le chemin est semé d'embûches, mais le jeu en vaut la chandelle. Car on attend en IA de grandes avancées, et l'on a déjà des témoignages de vies sauvées par ces techniques. Le médecin automatique n'est certainement pas pour demain, et n'est certainement pas souhaitable, mais le médecin « augmenté », mieux outillé pour guérir, prévenir, analyser, décider, et en même temps pourvu d'empathie et doué pour le contact humain, a certainement de beaux jours à venir. Il s'agira d'améliorer les diagnostics, les observations, les choix thérapeutiques.

L'imagerie, et plus généralement l'analyse de données médicales, est sans doute le domaine de la santé où l'apport de l'IA est le plus apparent et où les progrès sont les plus rapides. La santé est un domaine de choix pour la pleine exploitation des images, soit que leurs détails soient trop fins ou subtils pour être bien interprétés, soit que leur résolution médiocre demande un travail d'extrapolation, soit que la bonne interprétation demande une expérience et une connaissance d'une montagne de cas à des degrés inhumains.

Les modèles mathématiques et physiques, en partie phénoménologiques, sont précieux dans un contexte où les ressorts biologiques nous restent bien souvent inaccessibles.

La simple observation de l'évolution d'une maladie, le jugement de l'efficacité d'un traitement sur une tumeur : tout cela est aujourd'hui jugé de façon très approximative, par exemple en mesurant le diamètre de la tumeur. Les procédés de modélisation de la croissance tumorale, de plus en plus performants, permettent de prédire la réponse de la tumeur aux traitements administrés. Prendre en compte les informations de plusieurs niveaux, depuis l'échelle moléculaire jusqu'à l'échelle macroscopique, est un défi à fort enjeu. Non moins importante est la prise en compte, aux côtés de l'imagerie, des données génomiques de la tumeur, ou encore de la modélisation.

La correspondance entre génotype et phénotype reste le plus vexant des mystères de la biologie ; mais les nouvelles méthodes d'apprentissage statistique offrent un contournement. Elles permettent d'automatiser l'analyse de jeux de données moléculaires extraordinaires, pour mieux identifier les régions du génome associées à la progression tumorale et déterminer le traitement le plus approprié. Ce même apprentissage peut aussi aider à surmonter le déséquilibre entre le grand nombre de variables étudiées et la faible taille des échantillons disponibles, en sélectionnant les données les plus pertinentes.

La spécialité de pathologiste est elle aussi renouvelée par l'IA. Les lames virtuelles permettent des analyses à distance, sans microscope. L'analyse multiparamétrique permet de rechercher des informations de plus en plus complexes. De nouvelles technologies d'exploration des organes, extrêmement prometteuses, commencent à s'imposer avec des quantités de données à la clé. Le microbiote, l'environnement, et maintenant des écosystèmes entiers peuvent se recenser par des outils statistiques.

Les essais cliniques sont aussi en train de se renouveler : les plateformes qui permettent d'enregistrer, croiser de grandes données cliniques et moléculaires suggèrent de nouveaux formats d'essais. Les études ciblent les anomalies moléculaires plus que les organes, lorsque des altérations pilotes fortes sont identifiées, l'information moléculaire entre progressivement dans la pratique courante.

Enfin, et peut-être surtout, les données médicales deviennent accessibles, en particulier à de jeunes équipes multidisciplinaires qui travaillent en mode ouvert (*open science*, *open data*).

Ces évolutions ne vont pas sans poser des enjeux majeurs. Et d'abord dans la course aux trois ingrédients majeurs qui sont, partout dans le monde, l'objet d'une compétition effrénée : les cerveaux, le matériel informatique haute performance, les jeux de données.

Les cerveaux d'abord : l'IA se fait surtout avec des experts aguerris, dotés d'esprit de collaboration, prêts à se lancer dans la partie, amateurs de contacts interdisciplinaires. Les profils à double compétence (santé et grandes données, algorithmique et médecine, etc.) sont très prisés. La formation d'une nouvelle génération de scientifiques prêts à occuper ces interfaces est un enjeu majeur. Les formateurs ne sont pas assez nombreux, les programmes trop rares ; il faudra se lancer dans tout cela en mode expérimental.

Le matériel informatique haute performance ensuite : l'Europe a, on le sait, un considérable retard d'investissement par rapport aux autres continents. L'IA en santé du futur aura besoin de grands centres de stockage et de calcul, en partie internes et en partie externalisés.

Les jeux de données enfin et surtout : la recherche de corrélations, de régularités, de signaux faibles, demande de très grands jeux de données multiparamètres, croisant les données d'acteurs variés, privés et publics, spécialisés et banalisés, dont seul l'État peut espérer réaliser le mariage. Les grandes bases de données existantes, imbriquées, avec leurs acronymes abscons (SNIIRAM [Système national d'information

inter-régimes de l'Assurance maladie], SNDS [Système national des données de santé], etc.), les plateformes telles que celle de l'AP-HP (Assistance publique-Hôpitaux de Paris), ont toutes vocation à faire partie d'un ensemble plus vaste.

La construction de ces grands ensembles de données posera un triple défi. Le premier est le défi technique et technologique posé par la taille des données, la nécessité de rapidité d'accès, la variété des formats, les considérables imperfections des données, le besoin de cybersécurité renforcé. Le second est le défi éthique et légal, particulièrement en ce qui concerne la protection des données personnelles. Le troisième, non moins ardu, est celui de la confiance, tant il est difficile en pratique de convaincre les acteurs de surmonter leurs réticences à s'allier. Dans le contexte français, l'État sera amené à jouer un rôle clé sur tous ces sujets.

Insistons enfin sur le rôle important que les patients devront jouer dans ces évolutions : un rôle actif dans le consentement éclairé, mais aussi une participation au développement des outils, une contribution à l'évolution des pratiques et des mentalités. Si les médecins conservent probablement la responsabilité du contact et de l'explication aux patients, un nouvel usage est appelé à se dessiner pour la bonne interaction constructive des humains et des algorithmes dans cette activité – la médecine – particulièrement définissante pour l'humanité.

Aujourd'hui toutes les nations qui ambitionnent de compter sur la scène mondiale de l'innovation se lancent dans des programmes d'IA et santé : la France ne fait pas exception et commence tout juste à se lancer dans la partie. En particulier, c'est l'un des axes qui est développé dans le rapport « Donner un sens à l'intelligence artificielle » que le second auteur a rendu au Gouvernement en mars 2018. Fruit de six mois de travail interdisciplinaire, comportant quantité d'auditions et de nombreux déplacements, ce rapport a l'ambition de tracer la voie pour une politique coordonnée d'IA en France, avec aussi l'Europe en ligne de mire. Le président de la République en a repris l'essentiel des propositions, lors de son discours du 29 mars 2018 au Collège de France.

Cet ouvrage est le fruit d'un autre groupe de réflexion : un groupe de travail commun entre l'Académie des Sciences et l'Académie de Médecine, que nous avons souhaité initier il y a deux ans. Ce groupe de travail est constitué de mathématiciens, informaticiens, statisticiens, cancérologues, chirurgiens, onco-généticiens, sociologues, administrateurs hospitaliers, juristes… tous reconnus dans leur discipline. Ce groupe de travail a nourri une partie de la réflexion prise en compte dans le rapport au Gouvernement ; il a aussi influé sur les détails de la loi Informatique et liberté qui adapte le droit français pour les besoins du Règlement général de protection des données. Il est destiné à continuer à faire progresser les réflexions, aussi bien pour les spécialistes que pour les curieux. Le présent ouvrage rassemble une variété de contributions très succinctes, proposées par des intervenants au groupe de travail. Tous les thèmes exposés précédemment s'y retrouveront.

Pour conclure cet ouvrage, on prendra de la hauteur avec quelques ouvertures mêlant médecine et société : le rôle ambivalent des nouveaux modes d'information et notamment des réseaux sociaux, qui bouleversent la communication et propagent les progrès, mais aussi les fausses nouvelles et les annonces de complots ; le dialogue affectif et parfois trompeur entre humains et machines ; la position de l'humanité par rapport à la tentation de sa propre augmentation sans garde-fous.

Chemin faisant, les contributeurs insisteront sur les directions et les idées fortes davantage que sur les détails. Si cet ouvrage parvient à vous faire saisir un panorama d'ensemble de la façon dont se présente l'IA en médecine, alors il aura rempli son but.

L'intelligence artificielle et la santé de demain

Cédric Villani et Bertrand Rondepierre

Le potentiel de l'intelligence artificielle dans le domaine de la santé est immense et a fait l'objet de travaux significatifs dès le début du XXe siècle aux premiers balbutiements de la discipline. Les années 1970 ont vu exploser l'intérêt de la communauté scientifique pour les applications biomédicales, aidé par la mise à disposition des moyens de calculs distribués comme SUMEX-AIM (Stanford University medical experimental computer for Artificial Intelligence in Medicine) en provenance des universités de Stanford et de Rutgers. Parmi les premières applications pratiques pouvant être relevées, le projet MYCIN par exemple, développé dans les années 1960-1970 à l'université de Stanford, permettait d'identifier à l'aide d'un système expert les bactéries à l'origine de graves infections pour ensuite proposer des traitements adéquats.

Outre l'intérêt général qu'il porte, le domaine de la santé a toujours exercé une fascination particulière sur la communauté scientifique en plus d'être un terrain de jeu tout désigné pour l'intelligence artificielle. L'accroissement général des volumes de données disponibles dans des proportions extraordinaires, la complexité, le nombre et la variété des phénomènes en présence font de la santé un sujet d'étude presque infini et extrêmement diversifié. Cette complexité, qui défait *a priori* toute tentative de modélisation complète de la biologie humaine et des

mécanismes à l'œuvre, est justement le lieu d'expression privilégié de l'IA. À l'image d'un médecin, les techniques d'IA s'appuient sur des observations qui, confrontées à une connaissance théorique et empirique, permettent de produire une information exploitable par le praticien. Dans un contexte où il est de plus en plus difficile de reproduire les résultats des études conduites par la communauté et où, comme dans beaucoup d'autres domaines, le passage d'une découverte à un produit ou à une nouvelle pratique se fait rarement sans friction, l'IA fait figure d'exception car elle est perçue comme une capacité directement activable et utilisable sur le terrain.

Le besoin de connaissance théorique est cependant parfois remis en cause par les approches récentes, telles que les réseaux de neurones et l'apprentissage profond, qui parviennent à découvrir par elles-mêmes comment accomplir certaines tâches et cela sans information *a priori* sur le phénomène étudié. À partir de l'observation de nombreux exemples, ces méthodes établissent des liens de corrélation entre motifs observés et capacité à accomplir la tâche demandée. Ces liens découverts peuvent cependant être fallacieux : d'une part par la différence bien connue des mathématiciens entre corrélation et causalité et, d'autre part, parce que les données peuvent véhiculer des biais qui faussent la pertinence du résultat obtenu. De telles faiblesses, sans gravité s'il s'agit de recommander un film, peuvent avoir des conséquences particulièrement désastreuses dans le domaine de la médecine. La découverte autonome par les algorithmes n'est pourtant pas une nécessité car, contrairement à certains domaines où l'être humain est parfaitement incapable d'expliciter les processus cognitifs à l'œuvre pour transcrire une donnée en une information exploitable, des siècles d'études et d'approfondissement de notre connaissance de la médecine nous ont permis de produire un large spectre de modèles applicables à la biologie humaine. Si ces modèles existent, force est de constater que la médecine reste, dans son ensemble, basée sur l'observation, et sa pratique se trouve donc extrêmement dépendante de l'expérience du praticien dont le diagnostic s'appuie sur une connaissance *a priori* ainsi que sur un historique de situations antérieures auxquelles la nouvelle est comparée. Deux praticiens aux expériences différentes auront donc potentiellement, et naturellement, deux diagnostics différents

sur une même observation. À l'image de certaines découvertes thérapeutiques ou épidémiologiques dont l'explication échappe à l'état actuel de nos connaissances, certains phénomènes ne sont appréhendés que par un processus empirique. Ainsi, c'est à l'intersection entre l'approche aveugle et la modélisation formelle complète que l'IA pourra révéler son véritable potentiel et, par exemple, éviter des erreurs grossières, incompréhensibles pour un être humain.

L'importance de l'information empirique amène le constat que la donnée qui en est l'expression, encore plus qu'ailleurs, constitue le cœur de l'exercice de la médecine dans son rapport avec le patient. Il n'est donc pas surprenant dans le cadre d'une numérisation globale de l'ensemble de la pratique médicale et de l'utilisation de capteurs de plus en plus nombreux et variés, d'assister à une augmentation exponentielle de la quantité et de la diversité des données disponibles. Comptes rendus écrits par les professionnels de santé, résultats d'analyses, génomique, imagerie médicale, enregistrements de signaux biologiques, historiques d'utilisation de médicaments (etc.) sont autant d'exemples de sources de données pouvant constituer la matière première de la recherche en IA. Leur diversité est l'expression de la multiplicité des applications possibles pour la santé dont on peut citer quelques exemples récurrents : aide au diagnostic, pharmaco-/matério-/infectio-vigilances, médecine personnalisée, suivi du patient, recherche clinique et usages médico-administratifs… La donnée n'est cependant pas en soi un sésame dont la possession seule permettra le développement de l'IA : une donnée en soi n'a d'intérêt que si elle est bien propre, rangée et bien étiquetée. C'est pour cette raison que toute démarche d'IA passe nécessairement par une étape de travail sur ces données pour parvenir à en extraire l'essence et lui permettre ainsi de les exploiter à leur juste valeur.

L'état de l'art de l'IA pour ces différentes applications est très hétérogène. À l'image d'autres domaines, les réseaux de neurones ont permis de progresser significativement sur l'exploitation des signaux biologiques comme en témoigne la réussite de jeunes entreprises françaises telle que Cardiologs spécialisée dans l'analyse automatique d'électrocardiogrammes. Sur ces problématiques, la combinaison de la disponibilité des données et du raffinement des algorithmes a permis d'égaler, si ce n'est parfois dépasser, le niveau de performance

obtenu par un humain sur la détection de pathologies à partir de ces signaux. De même, grâce à la profondeur des bases de données publiques, les usages liés aux vigilances ont été rendus possibles avec de premières réussites notamment pour la corrélation de pathologies. On peut citer en illustration les travaux parus en février 2018 dans *The Lancet Public Health*, montrant que la consommation excessive d'alcool est associée à un triplement du risque de démence en général et un doublement du risque de développer la maladie d'Alzheimer. Les avancées dans le domaine du traitement du langage naturel ont permis de progresser dans l'extraction d'informations issues notamment des comptes rendus écrits par les médecins ; elles permettent d'espérer pouvoir à terme capitaliser l'information médicale en simplifiant le travail amont de qualification et d'enrichissement des données. La révolution de l'IA pour la santé ne fait que commencer et les années à venir devraient être riches en nouveaux développements.

La clef de la réussite de cette révolution est à la confluence entre trois types de ressources : d'abord humaines, ensuite empiriques au travers des bases de données, et enfin calculatoires puisqu'il s'agit avant tout d'informatique opérée à grande échelle. La ressource humaine est la ressource la plus critique de ce triptyque mais également la plus protéiforme. Critique d'abord pour l'expertise, aussi bien dans le domaine de l'IA que dans le domaine de la santé avec une véritable tension sur les profils arborant cette double compétence. Ces experts bicéphales sont en effet particulièrement prisés pour leur capacité à injecter de la connaissance médicale dans les méthodes d'IA, mais également pour la réciproque visant à rapprocher l'IA des professions médicales en les faisant correspondre aux besoins et usages les plus pertinents. Les ressources humaines sont également critiques quant à l'implication de l'ensemble des professionnels et des patients : la donnée et sa qualification étant une préoccupation majeure, comme matérialisant le lien entre les machines et le monde réel, elle ne peut se passer ni de ceux qui la produisent, ni de ceux qui l'enrichissent pour lui associer une valeur métier, ni de ceux qui l'exploitent dans leur pratique médicale.

La bataille pour l'IA se joue à plusieurs niveaux, en particulier dans le domaine de la santé où de très nombreux pays essaient de se

positionner comme leaders. En premier lieu, une bataille pour les ressources humaines se joue parallèlement entre États, entre entreprises et État ainsi qu'entre entreprises. En second lieu se joue un combat pour la place de leader en matière d'IA pour la santé. À première vue il semble que les États-Unis sont pour le moment en tête avec une avance certaine, alors que le Canada, le Royaume-Uni et Israël se positionnent comme de très sérieux concurrents. Le cas particulier d'Israël est particulièrement intéressant à ce titre : l'obtention de données propres, obtenues du fait d'une campagne systématique et autoritaire auprès des médecins et de l'ensemble du corps hospitalier a permis de constituer la ressource initiale nécessaire à une démarche d'IA pour la santé à plus grande échelle. On peut ajouter à cela le développement d'équipes de recherche, notamment au sein des quelques grandes caisses d'assurance qui mènent l'ensemble du jeu de la sécurité sociale, capables de mettre sur pied des modélisations habiles en s'appuyant grandement sur une interdisciplinarité et une coopération entre spécialités. Ce terreau fertile est de plus alimenté par le creuset de talents que constitue le service militaire israélien, occasion de recruter de nombreux jeunes dans des programmes stratégiques, combiné à un environnement très favorable aux start-up, à la disponibilité de capitaux, et à une recherche et une économie très ouvertes, tournées à la fois vers les États-Unis d'Amérique et vers l'Europe.

Forces et faiblesses françaises

La France quant à elle jouit d'une situation privilégiée pour tenir une place importante à l'échelle internationale dans le domaine de l'IA pour la santé. L'enseignement supérieur et la recherche français ont toujours occupé une place de choix au niveau international, notamment du fait de l'excellence reconnue de la formation scientifique en France. L'excellence notamment dans le domaine des mathématiques et de l'informatique donne à la France un avantage compétitif, confirmé par la présence de nombreux Français à la tête des plus prestigieuses

équipes travaillant sur l'IA. Le plus célèbre d'entre eux et l'une des figures de proue sur le sujet des réseaux de neurones, Yann Le Cun, travaille actuellement entre la New York University et le Facebook AI Research (FAIR). En élargissant la communauté française à la communauté francophone, on se doit également de citer Yoshua Bengio, autre figure de proue du domaine autour de laquelle se concentre une partie de la stratégie canadienne. Si cette excellence académique est la condition *sine qua non* pour espérer une position de choix dans une compétition internationale rude, elle n'est pas suffisante pour la garantir. La question de l'évolution des formations en médecine subsiste aujourd'hui, et leur rigidité actuelle ne semble pas en mesure de répondre au besoin de multidisciplinarité qui ne pourra aller que croissant. L'intégration de l'IA, et plus largement du numérique, dans les cursus semble être une nécessité pour des professions qui seront sur le front de ces développements technologiques.

En aval de la recherche, et contrairement à certaines idées reçues, l'innovation en France conserve un grand dynamisme malgré un certain manque de publicité et de visibilité. Alors que les géants du numérique voient une part non négligeable de leurs publications scientifiques relayées autant dans la presse spécialisée que dans les quotidiens nationaux voire internationaux, certaines réussites françaises, à l'image de Therapixel, vainqueur d'une compétition mondiale face aux équipes des plus grands laboratoires de la planète, restent confidentielles. Même si la France n'est la patrie d'aucun géant du numérique, elle n'est pas pour autant démunie du point de vue industriel avec des poids lourds du domaine tels que Sanofi ou encore Atos sur un positionnement plus transverse et même OVH pour les hébergeurs.

Les forces académiques, entrepreneuriales et industrielles sont donc bien présentes et peuvent s'appuyer sur une matière première unique au monde. En effet, la tradition jacobine de la France l'a conduite avec le temps à mettre sur pied un système de santé collectant massivement les données nécessaires à l'accomplissement des missions médico-administratives dans un mode essentiellement centralisé et géré par l'État. Ajoutant à cela le fait que plus de 65 millions de Français participent à ce système, lui conférant ainsi la taille critique

nécessaire, on est en droit d'imaginer que l'utilisation des données déjà capitalisées permettra de dégager des réussites à court terme. Ces données sont regroupées au sein du Système national des données de santé (SNDS) qui a aujourd'hui pour objectif premier l'organisation et la mise en œuvre du remboursement des soins.

La sensibilité du domaine et des informations qui sont manipulées pose plus singulièrement la question de la confiance. Dans ce cadre, le fait que la mission de santé publique soit essentiellement exercée par l'État et par des structures qui relèvent directement de lui en fait un intermédiaire de confiance légitime pour développer l'IA en santé et structurer les organisations à cette fin. L'enjeu ici est d'une part de parvenir à faire accepter de tels développements et, d'autre part, de garantir le cadre dans lequel ceux-ci seront réalisés sans avoir recours à des intermédiaires bien que ces derniers ne soient pas non plus exempts d'obligations. En effet, l'existence des réglementations française et européenne avec le Règlement général sur la protection des données (RGPD) comme fer-de-lance, contribue à créer un cadre de confiance favorable au développement de systèmes d'IA respectueux et protecteurs des individus et donc d'autant plus susceptible d'emporter succès et adhésion de la part de nos concitoyens.

Ces avantages comparatifs ne seront pas les seuls déterminants de notre réussite. Le phénomène de tension autour des ressources humaines pourrait se traduire par un assèchement de nos laboratoires de recherche et de nos entreprises qui ne seraient alors plus que des viviers pour les acteurs extra-européens. De ce fait la dynamique de l'innovation s'en trouverait mise en péril, entamant ainsi durablement notre force de frappe en la matière. En raison, de surcroît, de la difficulté de mener des projets de recherche dans le domaine de la santé du fait des complexités administratives et réglementaires, il subsiste une barrière à l'entrée qui ne permet pas en l'état de faire émerger des acteurs majeurs, qu'ils soient académiques ou industriels.

Si le SNDS est jugé comme une richesse immense, il n'en est pas moins limité et insuffisant pour permettre à lui seul le développement d'une véritable force française pour l'IA et la santé. On ne peut pas en effet attendre d'un système qui a été conçu uniquement à des fins médico-administratives que celui-ci soit naturellement équipé pour

supporter et favoriser une démarche de recherche et d'innovation efficiente en matière d'intelligence artificielle. À cette fin manquent de façon très pragmatique les ressources matérielles et logicielles nécessaires à la mise en œuvre d'une telle démarche, mais également la gouvernance et l'organisation nécessaires pour assurer la cohérence à l'échelle nationale. La preuve en est qu'à ce jour les données cliniques disponibles dans les hôpitaux ne sont toujours pas mises en commun avec les données nationales ; même dans le cadre de la mission première du SNDS, il peut se passer des mois voire des années avant qu'une donnée soit capitalisée au sein du système lorsqu'elle n'a pas encore été codée en vue du remboursement. La qualité des données étant primordiale, il faut bien avoir à l'esprit qu'un travail important sera nécessaire pour les rendre directement utilisables par des démarches d'IA, et que leur gestion devra être assortie d'une gouvernance aujourd'hui structurellement mise en défaut.

Enfin, l'excellence de la recherche et de l'innovation en IA et santé ne saurait à elle seule suffire à faire émerger des champions économiques : quelle que soit la qualité des travaux amont, elle doit être associée à l'existence d'un marché de grande ampleur qui n'existe à ce jour ni au sein de l'État, ni auprès des entreprises privées, ni auprès de nos concitoyens. Si l'on convient que le marché est *a minima* européen et même éventuellement international, il reste un gouffre à franchir pour transformer une innovation en réussite économique alors que le capital-risque français reste insuffisant pour faire face à la compétition internationale.

Un projet pour la France

Quel que soit le plan à mettre en œuvre, il faut garder à l'esprit les trois types de défis principaux auxquels toute démarche sera confrontée : 1. techniques 2. légaux et éthiques 3. confiance et organisationnels. Les défis techniques représentent essentiellement la capacité à déployer de nouvelles technologies dans des systèmes adaptés tout en assurant la continuité et la reprise de l'existant. Ces défis existent

quelle que soit la technologie sous-jacente et relèvent d'un travail d'architecte minutieux. Les défis légaux et éthiques soulevés sont en revanche beaucoup plus spécifiques à l'IA et au domaine d'application qu'est la santé du fait de la criticité et de la sensibilité de l'application qui imposent un questionnement particulier. Enfin, les défis relevant de la confiance et de l'organisation révèlent que le premier frein au développement de l'IA n'est pas tant technologique ni réglementaire qu'humain. L'IA ne saurait se développer sans la confiance de ses usagers et le soutien des organisations qui sont souvent mal équipées pour faire face à un sujet aussi large et transverse. Contrairement aux idées reçues, il semblerait que les défis techniques sont souvent plus faciles à relever que ce qui était anticipé, alors qu'à l'opposé les défis relevant de la confiance et de l'organisation sont beaucoup plus difficiles à aborder que prévu. Ceci s'explique certainement par le fait qu'un défi technique appelle la plupart du temps une solution technique, alors qu'un défi humain nécessite un travail de fond, de réflexion et de pédagogie beaucoup plus important.

Afin d'organiser l'action française, le rapport « Donner un sens à l'intelligence artificielle » propose une démarche globale ayant pour objectif de faire émerger un écosystème à la fois académique et industriel de grande envergure dans le domaine de l'IA et de la santé en capitalisant sur nos forces. La première nécessité d'une telle démarche consiste à fixer une direction comportant quelques objectifs ambitieux relevant de l'intérêt général afin d'entretenir et de structurer l'effort de l'écosystème dans la durée. Identifier clairement ces objectifs est également le préalable pour les outiller dans la mesure où les actions ne seront certainement pas les mêmes si l'on s'intéresse à la réduction du risque de maladie nosocomiale ou à la médecine personnalisée. Fixer une direction est également l'occasion de mener en amont les réflexions sur l'éthique et la confiance en les inscrivant dans des projets concrets, évitant ainsi l'écueil des discussions trop théoriques qui pourraient conduire à statuer au détriment du progrès alors même que les dimensions apparemment contradictoires de ces projets étaient conciliables.

Avant que d'autres s'en chargent à notre place, il est indispensable de réorganiser notre système de santé sur le modèle qui a fait le

succès de l'économie numérique, à savoir le modèle de plateforme comme intermédiaire dans l'accès aux informations, contenus, services et biens édités ou fournis par des tiers. Une telle plateforme est l'occasion de déployer des moyens de calcul et de stockage des données conforme à l'état de l'art compatible avec les usages de l'intelligence artificielle. Cet aspect technique représente plus une condition préalable et une mise à jour de l'existant qu'une rupture en soi. En revanche, il s'agit du vecteur qui aura à rassembler l'ensemble des données pertinentes pour les usages de l'IA en santé, à en instrumenter la collecte et la captation en temps réel. Ce faisant, cette plateforme a vocation à être déployée depuis l'échelon central jusqu'à l'échelon local (hôpitaux, professionnels de santé et patients) mettant ainsi en cohérence des systèmes historiquement déconnectés les uns des autres. Du point de vue économique, c'est un moyen pour changer la donne : en permettant à des tiers (chercheurs, entrepreneurs, grandes entreprises mais également puissance publique) de développer et d'expérimenter de nouvelles fonctionnalités qui s'appuient directement sur le système de santé et les données qu'il contient, la plateforme ouvre la voie à une nouvelle façon de créer et de répartir la valeur ajoutée.

L'accent devra être mis sur la recherche et l'expérimentation. Entre obligations réglementaires trop lourdes, délais d'instruction des dossiers trop longs, les plaintes sont récurrentes. Si la situation s'est grandement améliorée ces dernières années, afin de se rendre compatible avec le rythme de l'innovation, la mise en place dans le domaine de la santé de « bacs à sable » ouverts à l'expérimentation, permettrait d'alléger temporairement les contraintes qui pèsent sur les acteurs, de les accompagner dans la prise en compte de leurs obligations, et de réaliser des expérimentations en conditions réelles. La recherche sera également la première bénéficiaire de cette politique : grâce à un accès privilégié aux données et des conditions d'expérimentation assouplies, les laboratoires de recherche publique tels que les instituts interdisciplinaires de recherche en intelligence artificielle (instituts 3IA), proposés pour libérer les forces de notre tissu académique, devraient trouver là un terrain fertile pour le développement de l'IA pour la santé de demain.

Enfin, la difficulté la plus prégnante sera de faire face aux défis organisationnels, en mettant en place une gouvernance assortie d'une politique de la donnée transverse dont l'objectif serait de penser en amont la collecte, la capitalisation des données, la gouvernance de la plateforme de la santé ainsi que l'organisation de l'effort autour du développement de l'intelligence artificielle. Un chantier tel que celui du Dossier Médical Partagé devrait être accompagné d'un axe IA afin que les données ne soient plus seulement capitalisées à des fins documentaires, mais également à des fins de recherche, d'innovation et d'exploitation ultérieures relevant de l'IA. Une telle possibilité nécessite une structuration spécifique avec notamment la possibilité de suivre le parcours du patient, opération rendue aujourd'hui difficile par exemple par le fait que l'utilisation du NIR [1] n'est pas systématique.

Parmi les conditions de réussite, il faut souligner que la montée en compétences des organisations (entreprises, administrations, autorités de régulation, médecins et acteurs du monde de la santé) en matière d'IA est indispensable. Il en est de même pour les patients auprès desquels la nécessité de démystifier l'IA devient de plus en plus prégnante au risque d'un sentiment de défiance, voire d'un refus net de profiter des possibilités ouvertes par ces technologies. Dans la mesure où le développement de l'IA dépend principalement de l'implication des patients, l'acculturation, l'acceptation et la confiance doivent constituer une préoccupation de premier ordre.

Pistes européennes

Il est communément admis que l'échelon français n'est pas suffisant face à des géants tels que les États-Unis ou la Chine, pour atteindre la masse critique nécessaire à cette compétition qu'il faut aller chercher du côté de l'Europe. Dans le domaine de la santé, l'exercice est ardu. Notre voisin allemand par exemple, auprès de qui une coopération aurait été naturelle, dispose d'un système de santé beaucoup plus hétérogène et disparate que le nôtre et n'offre pas de ce fait les

1. Numéro d'inscription au répertoire ou numéro de sécurité sociale.

mêmes opportunités sur le court terme. Il semble y avoir autant de situations différentes dans chacun des États membres et, puisque le temps ne l'a pas permis, il serait utile de mener un travail de recensement et de cartographie des différents systèmes de santé européens pour ensuite établir les potentialités en matière d'IA. En fonction des résultats d'une telle étude, l'opportunité d'une démarche d'élaboration d'un système de santé intelligent et européen pourrait être évaluée afin de déployer les logiques du plan français à l'échelle européenne qui pourrait porter notre politique académique et industrielle de l'IA appliquée à la santé.

Ce n'est donc qu'au prix de grandes transformations en rupture avec l'existant que la France pourra, avec l'Europe, se saisir de la place qu'elle souhaite occuper en matière d'IA et de santé.

Advancing Healthcare Through Data-Driven Medicine and Artificial Intelligence

Ran D. Balicer et Chandra Cohen-Stavi

« If it were not for the great variability among individuals, medicine might as well be a science, not an art. »

Sir William Osler, 1892.

Introduction – The status quo and its toll

Health systems are facing escalating challenges in recent decades, as the increase in demand and costs cannot be met by a similar rate of increase in human resource and invested capital. The rise in demand is the result of several key trends: The population is growing older, and the relative proportion of the population over the age of 65 is rapidly increasing. As the average healthcare needs of the elderly population are far greater than the younger population segments, this trend is straining healthcare systems. In parallel, the chronic disease morbidity in all age groups is becoming increasingly more prevalent. A majority of people above the age of 50 years have chronic disease comorbidity or multi-morbidity – 2 diseases or more, and this multi-morbidity has a synergistic impact on their healthcare needs.

These strained systems and individual providers are working in an ineffective and error-prone traditional model of reactive care. In the current state of affairs, the intuitive and experience-based art of medicine is failing us, and allows for sub-standard care to be the norm, even in the most advanced healthcare systems. It has been estimated that 45% of the evidence-based interventions required in routine care are missed, and many deaths are the result of medical errors or inappropriate care[1], despite clinical professionals doing their best to deliver the right care. The increased complexity in the options of care models, medical treatments and medical technologies, bring about intensifying challenges in disentangling the evidence for effectiveness, with almost 2.5 million English-language articles per year published in scientific journals[2]. The providers themselves bear severe professional and personal cost for this continuous increase in workload and inappropriately designed work setting, with recent surveys suggesting that up to 50% of physicians reporting severe burnout signs and symptoms[3].

The promise of Artificial Intelligence in transforming healthcare for the better

It has been suggested that in view of these basic inadequacies and escalating challenges, healthcare systems must adapt and profoundly

1. Makary M.A., Daniel M., "Medical error – the third leading cause of death in the US", 2016. BMJ 2016;353:i2139; Hogan H, Healey F, Neale G, *et al.*, "Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study", *BMJ Quality and Safety*, 2012.

2. Ware M., Mabe M., *The STM report: An overview of scientific and scholarly journal publishing*, 2015, The Hague: International Association of Scientific, Technical and Medical Publishers. Retrieved 7 August 2017, from: https://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf

3. Shanafelt T.D., Hasan O., Dyrbye L., Sinsky C., Satele D., Sloan J., *et al.*, "Changes in burnout and satisfaction with work-life balance in physicians and the general US working population between 2011 and 2014", *Mayo Clinic Proceedings*, 2015, 90 (12), 1600-1613.

Medscape survey of 15,000 physicians. Report accessed at: https://www.medscape.com/slideshow/2018-lifestyle-burnout-depression-6009235?faf=1

change in their operating model to allow sustainable effective care to their members. One way to navigate rising healthcare demand and exponentially increasing complexity in care and treatment options is to proactively anticipate future demand, risks and health needs, and use this knowledge for better care allocation and patient-centered treatment decisions. The ability to understand the dynamic factors affecting health through data analytics, pattern detection, and prediction is paramount to creating more intelligent healthcare.

Effective and efficient models of care must strive to be proactive and prevention-focused rather than reactive and treatment-focused; integrated rather than divided into detached care silos; precise and tailored to the specific patient's unique set of characteristics and needs rather than based on average population observations. It is a common wisdom that the smart use of data in healthcare is both a key requisite and driving force to allow such changes to be implemented on a large scale.

The expansion of embedded information technology (IT) systems across healthcare settings provides an increasing wealth of data that can be harnessed through advanced analytics, in a systematic transformation to narrow key gaps for the healthcare systems of the future. These include the widespread implementation of electronic health record systems, the availability of health-related data from personal and healthcare devices, and the increasing use of genetic testing which produce massive patient-specific datasets. Continuous improvement in computing power and widespread implementation of AI in other daily domains of life increase the drive to implement similar systems in healthcare as well.

Among many opportunities that data and advanced analytics offer for healthcare advancement, there is the potential for preventing complications and improving prognosis (to preempt), supporting patient-focused treatment decision-making (to individualize), integrating patient involvement in navigating their care (to align expectations between patients and providers), and the automation of clinical tasks and service allocation to improve efficiency and safety of health institutions (to make repetitive technical tasks redundant, reduce errors and reduce clinical variability).

Multiple advanced data analytics techniques are being increasingly used in academic and commercial innovation platforms to develop data-driven tools for the healthcare setting, from basic algorithms based on straightforward regression statistics to the most complex *machine learning* tools based on multiple layers of artificial neural networks. Indeed, most of the computerized solutions in current clinical practice do not rely on independent computer intelligence, but rather on human-generated algorithms based on expertise and research. Still, in this text, we will loosely use the term Artificial Intelligence (AI) to denote the wide scope of tools that allow computers to perform tasks just as well as, if not better than, humans.

The National Health Service in the UK has outlined the ways AI can and is expected to improve and help transform healthcare, allowing for: the prediction of individuals who are at risk of illness to target treatment more effectively towards them; the development of AI diagnostic and treatment decision tools tailored to individual need supporting health professionals and patients; and the automation of clinical tasks and service allocation to improve efficiency and safety of health institutions[4].

The potential for improvement is profound. It has been shown, for example, that through the combination of *deep learning* AI techniques with pathologists'diagnosis for identifying metastatic breast cancer, an 85% reduction in the human error rate resulted[5]. Such automated analysis has also been shown to find and match specific lung nodules on chest CT nearly twice as fast than a panel of radiologists. The power of AI algorithms to diagnose through imaging results has been demonstrated in a study in which the performance

4. Harwich E., Laycock K., *Thinking on its own: AI in the NHS*, 2018, Reform Research Trust, London, UK.

5. Wang D., Khosla A., Gargeya R., Irshad H., and Beck A. H., https://scholar.harvard.edu/humayun/publications/deep-learning-identifying-metastatic-breast-cancer" Deep Learning for Identifying Metastatic Breast Cancer, arXiv preprint arXiv:1606.05718, 2016.

of algorithms to identify skin cancers was on par with a panel of 21 experts[6].

AI applications have been ranked based on how likely adoption is, and what potential exists for annual savings[7]. AI-assisted robotic surgery was ranked at the top, as it was estimated it can result in fewer complications and errors that would generate a 21% reduction in patients'length of stay, and 40$ Billion in annual savings in 8 years. The following highest ranked applications were: Virtual nursing assistants, administrative workflow, fraud detection and dosage error reduction. Yet very few clinical applications of AI have been realized in practice, and research backed implementations are scarce.

AI implemented in clinical practice

In view of the scarcity of clinical AI tool implementation, we review two examples of utilizing AI in the clinical setting, both from the Israeli healthcare system, where data and AI have been used in practice for improving clinical care, and supporting the transformation discussed above toward more proactive, preventive, precise and patient-centered care.

Israel, like France, has universal healthcare coverage for all its citizens, and provision of care is offered by four health funds. These health funds are public non-profit payer/provider systems that focus on extensive networks of community-based primary and secondary care clinics, but that also pay for all inpatient care. As the switching rate between health funds is very low, and generally people stay the majority of their life covered under one health fund, the organizations have financial interests to maintain their members'health and provide preventive services rather than to increase medical care volume. Israel

6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S., "Dermatologist-level classification of skin cancer with deep neural network", *Nature*, 2017, 542(7639), 115-118.

7. Kalis B., Collier M., Fu R., "10 promising AI Applications in Health Care", *Harvard Business Review*, may 10, 2018.

also benefits from a unique combination of attributes that support healthcare data-driven innovation, with two particularly notable attributes. There is an extensive information technology infrastructure of data repositories based on two decades of interoperable electronic health records across inpatient and outpatient settings. These data serve as a platform for data-driven innovation, rapidly translated into practice. The second is the "start-up nation" culture that exists, with the highest rate globally of start-ups, venture capital investment, and of medical equipment patents per person[8].

The first example presented, reviews an AI implementation toward proactive, predictive prevention. It is well accepted that attempts to address chronic disease in the later phases of overt organ damage and painful dysfunction, are limited in their effectiveness and are usually both costly and associated with serious adverse events. If the precursors of the pathological processes could accurately be identified at earlier stages, milder interventions would be more effective in delaying or completely averting the deterioration. Patients'health data and advanced analytics can be used to predict which patients are expected to undergo acute or chronic slow deterioration towards a preventable outcome or complication, allowing proactive preventive intervention.

In aiming to improve clinical prognosis and prevent deterioration to end stage renal disease (ESRD) – a major costly chronic condition that is a large burden to patients, their families and the healthcare system – Clalit Health Services, Israel's largest health fund, initiated the planning of a data analytics – driven targeted prevention program in 2010 to reduce the number of ESRD cases through data analytics-driven targeted prevention. It has been shown, that patients with early chronic kidney disease can change their expected deterioration path to ESRD through a set of simple preventive interventions. Yet early chronic kidney disease is fairly prevalent, and it is usually not feasible in most health systems to reach out to all of these seemingly

8. Balicer R.D., Afek A., "Digital health nation: Israel's global Big data innovation hub", *Lancet*, 2017 Jun 24, 389(10088), 2451-2453.

healthy patients to prevent future illness. Thus, at the Clalit Research Institute, a predictive model was developed and continuously improved, to identify which Clalit members were at high-risk to develop ESDR within a period of five years. Given resource considerations, the model targeted 8% of the population in question at highest risk, and that group included nearly 70% of those identified as high-risk for future deterioration to ESRD. Clinical action and tailored interventions for this target group of patients were implemented, and primary care teams were trained to follow pre-defined steps once a high-risk patient is among their patient list. Preliminary results of several years of implementation suggest a clinically meaningful drop in the incidence of ESRD among the patients that were targeted for intervention by the model.

This example shows the utility of predictive algorithms to target treatment more effectively toward high-risk patient groups for the prevention of major chronic disease complications. This approach is being applied in other relevant domains to allow for the identification of populations at-risk and early identification of impending complications of multiple acute and chronic illnesses[9].

A second example, reviews the development and design of a precise and individualized treatment decision tool. There is great variability among individuals in health-related characteristics, health behaviors and clinical response to treatment, and healthcare systems are ill equipped to deliver care for such variability[10]. Generally, clinical guidance and action is based on limited evidence of average effects in popula-

9. Balicer RD, Cohen CJ, Leibowitz M, Feldman BS, Brufman I, Roberts C, Hoshen M. "Pneumococcal vaccine targeting strategy for older adults: customized risk profiling", *Vaccine*, 2014 Feb 12, 32(8), 990-995; Shadmi E., Flaks-Manov N., Hoshen M., Goldman O., Bitterman H., Balicer R.D., "Predicting 30-day readmissions with preadmission electronic health record dat", *Med Care*, 2015 Mar, 53(3), 283-289; Dagan N., Cohen-Stavi C., Leventer-Roberts M., Balicer R.D., "External validation and comparison of three prediction tools for risk of osteoporotic fractures using data from population based electronic health records: retrospective cohort study", *BMJ*, 2017 Jan 19, 356, i6755.

10. Wallace E., Smith S.M., Fahey T., "Variation in medical practice: getting the balance right", *Family Practice*, 2012.

tions and not based on individualized, patient-specific evidence[11]. This results in patients who need care not getting that care, and patients who don't necessarily need care getting it, or in the selection of the wrong empirical first line treatment for serious illnesses.

In recent years, increasing use of patient-specific data analytics in clinical practice to drive clinical decisions tailored to the person and the affected tissue. These data are mined from medical records and lab tests that generate massive amounts of data (i.e. genetic sequencing). As many of the frequently discussed examples of precision care focus on use of personal genetic data, we will try to focus on using clinical ('phenomic') data to achieve care precision and personalization. We have demonstrated the potential of such data to drive treatment selection recently by analyzing the hypertension treatment SPRINT trial[12]. In this exercise, we used data on several thousand participants of clinical trial called the SPRINT trial. The data was made available to us as part of a global competition (the SPRINT Challenge) run by the New England Journal of Medicine[13].

We have been able to show that while the SPRINT trial summary suggested that on average, high risk patients with elevated blood pressure should have the blood pressure treated to a threshold below 120 mmHg, this was not true for a large subgroup of patients in this study. When multiple iterative individual-level modeling (thousands per individual) was performed for each of these trial subjects separately, it was possible to tease out those individuals that would likely suffer the harms of this suggested intensive treatment without

11. Yeh R.W., Kramer D.B., "Decision Tools to Improve Personalized Care in Cardiovascular Disease: Moving the Art of Medicine Toward Science". Circulation 2017;135(12):1097-100; McCartney M., Treadwell J., Maskrey N., Lehman R., "Making evidence based medicine work for individual patients", BMJ 2016, 353:i2452.

12. Wright J.T. *et al.* (the SPRINT Research Group), "A Randomized Trial of Intensive versus Standard Blood-Pressure Control", *N Engl J Med*, 2017 Dec 21, 377(25), 2506.

13. Burns N.S., Miller P.W., "Learning What We Didn't Know – The SPRINT Data Analysis Challenge", *N Engl J Med*, 2017 Jun 8, 376(23), 2205-2207.

experiencing its benefits, from those cases that were expected to mainly benefit from this treatment by preventing the complications of uncontrolled blood pressure. This demonstrates the marked impact that individual-level AI-driven decision support may have on every clinical decision faced by physicians, in view of their limited capabilities to compile and integrate all the patient's characteristics, the often fragmented and vast amount of clinical data, and extensive body of clinical evidence to reconcile.

Aspirations and Challenges

While the prospects are vast, they are tempered by the limitations and challenges in achieving intelligent health systems, which include the establishment of adaptable and interoperable infrastructure for sustainable solutions, the issue of the quality and scope of available data, and the multifaceted ethical issues that arise from the use of sensitive data and liabilities related to algorithms used for medical decisions[14].

Ultimately, the science of data analytics and AI holds much promise, but clinical impact assessment must become more rigorous and the purpose of such scientific efforts must be defined and driven by the unmet clinical needs. Uses of AI in practice for clinical decision making still remains in its infancy, and we must wait for the early adopters to provide sound evidence on the continuous and sustainable beneficial effects for patient outcomes in a sustained meaningful way. Yet it is our expectation, as well as that of many of the current opinion leaders in this field, that these applications will realize the promise and make a significant impact to transform healthcare for the better towards more sustainable and effective healthcare systems.

14. Nuffield Council on Bioethics 2018.

AI, a Vector for Positive Change in Medicine

Daniela Rus

Our world has been changing rapidly. Today, telepresence enables students to meet with tutors, and doctors to treat patients thousands of miles away. Robots help with packing on factory floors.

Networked sensors enable monitoring of facilities, and 3D printing creates customized goods. We are surrounded by a world of possibilities. These possibilities will only get larger as we start to imagine what we can do with advances in artificial intelligence and robotics.

On a global scale, AI will help us generate better insights into addressing some of our biggest challenges: curing diseases by better diagnosing, treating, and monitoring patients, understanding climate change by collecting and analyzing data from vast wireless sensor networks that monitor the oceans, the greenhouse climate, and the plant condition; improving governance by data-driven decision making; eliminating hunger by monitoring, matching and re-routing supply and demand, and predicting and responding to natural disasters using cyber-physical sensors. It will help us democratize education through MOOC offerings that are adaptive to student progress, and ensure that every child gets access to the skills needed to get a good job and build a great life. It may even help those kids turn their childhood dreams into reality, as Iron