

Comment l'anesthésie et la réanimation seront affectées par Dr Watson, Dr Google and Co ?

Pr Romain Pirracchio

Service d'anesthésie réanimation, Hôpital européen Georges Pompidou

Université Paris 5 Descartes

Groupe ACTEREEA

INSERM UMR-1153, Equipe ECSTRA, 1 rue Claude Vellefaux, 75010 Paris

Auteur correspondant : Pr Romain Pirracchio

Email : romain.pirracchio@aphp.fr

Aucun conflit d'intérêt

Points Essentiels

- Machine Learning et Intelligence Artificielle sont souvent confondus.
- Le Machine Learning est la capacité à décrypter et apprendre la structure intrinsèque des données.
- L'Intelligence Artificielle repose sur des méthodes de Machine Learning.
- Ces méthodes nécessitent de grandes quantités de données. Les « big data » médicale doivent venir du dossier médical informatisé universel.
- L'émergence de ces outils devrait permettre à court terme de prédire des événements (predictive analytics) afin de pouvoir anticiper et même prévenir.
- A moyen terme, ces nouvelles méthodes analytiques permettront d'aider la décision médicale (prescriptive analytics), de manière personnalisée, en apprenant fil de l'eau de l'évolution individuelle de chaque patient.

Introduction

L'intelligence artificielle prend une place de plus en plus importante dans notre quotidien. Alors que celle-ci relevait essentiellement de la science-fiction jusqu'à très récemment, les progrès majeurs en informatique, algorithmique, stockage et traitement des données ont permis à ces techniques de trouver des domaines d'application aussi multiples que variés. En effet, si l'intelligence artificielle a initialement percé dans des domaines professionnels dans lesquels la prise de décision dépend de l'analyse prédictive de séries temporelles complexes, telles que la finance, il a été plus surprenant de la voir intégrée si rapidement dans des domaines aussi variés que la publicité, les loisirs, ou encore la conduite autonome. Mais aussi variés que puissent paraître ces différents domaines d'application, la plupart d'entre eux présentent un point commun crucial qui est la relative homogénéité du format des données générées et analysées. Aussi, si le challenge technologique de l'analyse en temps de données de haut volume comme celles générées par les sites internet est réel, des approches analytiques relativement anciennes (telles que les réseaux neuronaux) ont été suffisantes pour produire des résultats surprenants. Le véritable challenge dans ces conditions a été de collecter les données, de les stocker et de mettre à disposition la puissance de calcul nécessaire pour une analyse en quasi temps réel. L'application à des domaines dans lesquels les sources d'information sont variées, à l'origine de signaux très hétérogènes, parfois difficiles ou impossibles à capturer et sujet à un bruit parasite important constitue un tout autre challenge. Ceci est le cas de la médecine et particulièrement des soins aigus et critiques. En effet, en anesthésie, en réanimation, comme en médecine d'urgence, la prise de décision est un processus complexe faisant intervenir le médecin avec ses propres caractéristiques, le patient et ses capacités d'interaction, les données physiologiques capturées à très haut débit par les moniteurs, l'examen clinique, les éléments paracliniques aussi variés que la biologie ou l'imagerie, des données textuelles plus ou moins systématisées issues de rapports médicaux antérieurs, les circonstances et l'environnement, et une infinité d'autres facteurs dont la plupart ne sont ni enregistrables, ni analysables avec les outils "simples" d'intelligence artificiel mis à disposition du grand public. Ceci explique probablement le "retard" apparent à l'intégration de l'intelligence artificielle en médecine. A ceci s'ajoute bien sûr, la conscience aigüe de l'impact d'une potentielle erreur de prédiction, de diagnostic ou pire de proposition thérapeutique, dont les conséquences peuvent s'avérer singulièrement plus importantes que celle du routage vers la mauvaise page internet ou de l'affichage de la mauvaise publicité ciblée.

Pourtant, malgré l'ampleur de la tâche, plusieurs équipes se sont attaquées depuis plusieurs à tenter de développer des outils et plateformes d'intelligence artificielle adaptées aux contraintes et aux attentes du monde médical. Pour cela, tout a dû être spécifiquement pensé et développé, depuis l'infrastructure de collection des données, la structuration et l'homogénéisation de leur format, l'anonymisation, le stockage sécurisé, le filtrage du bruit, l'amplification du signal, l'algorithmique d'analyse spécifique, le format de rendu des analyses ou encore l'intégration de ces formations dans la démarche diagnostique et thérapeutique. Chacune de ces étapes constitue en elle-même un domaine d'expertise et de recherche et il est peu probable de voir apparaître une percée significative de ces technologies dans le monde médical sans que l'ensemble de ces domaines n'ait connu un développement suffisamment avancé.

Nous détaillerons dans la suite de cette revue quelques points cruciaux qui font l'objet de recherche active afin de contribuer à l'introduction des technologies d'intelligence artificielle pour l'aide à la décision en anesthésie et en réanimation. Ces domaines couvrent notamment l'acquisition de données médicales de haut volume et leur analyse par des outils innovants d'apprentissage.

Big Data

Les **big data** ou **méga-données** sont des ensembles de données tellement volumineux qu'ils en sont difficiles à traiter avec des outils mathématiques et statistiques classiques. Ils sont généralement définis comme des matrices $n \times p$, avec $p \gg n$.

En médecine, les big data ont longtemps été cantonnées à la génomique. De nos jours il existe une multitude d'autres sources. Les principales sources de big data en anesthésie-réanimation sont :

- Dossier patient informatisé
- Données de monitoring
- Données d'imagerie
- Données de biologie

Ces données restent néanmoins très sous-utilisées. En avril 2016, une recherche sur Pubmed avec les mots clés « big data », « anesthesiology », « anesthesia » permettait de retrouver seulement 26 publications (1^{ère} en 2014).

Ces données restent essentiellement utilisées pour la prédiction du devenir des patients. Mais dans un avenir proche, l'accès à ce type de données et le développement de méthodes d'analyse spécifiques permettra d'avoir accès dans notre discipline à :

- Médecine personnalisée
- Aide à la décision en temps réel
- Prédiction en temps réel des évènements aigus

Les contraintes spécifiques pour l'analyse de ce type de données sont liées :

- Au stockage
 - Super calculateurs
 - Cloud computing
- A la complexité de traitement des données :
 - Données numériques de très haut volume
 - Texte brut (text mining)
 - Images (image mining)
 - Machine Learning

Plusieurs équipes à travers le monde travaillent et ont déjà développé de larges entrepôts de données de santé, visant à collecter de la manière la plus exhaustive possible les informations produites lors du séjour à l'hôpital d'un patient. Le format d'entrepôt de données de type I2B2 (Informatics for Integrating Biology and the Bedside) est celui développé et utilisé depuis le début des années 2000 par une équipe du MIT et a permis de produire la plus grande banque de données publique de réanimation (MIMIC). Ce format est actuellement progressivement

adopté par de nombreuses équipes à travers le monde, dont l'Assistance Publique des Hôpitaux de Paris qui travaille en collaboration avec l'équipe du MIT pour le développement de son entrepôt de données de réanimation.

Mais au-delà du challenge spécifique de la collection et l'intégration des données, de nombreuses autres problématiques existent. Nous n'aborderons que celles concernant la modélisation des données dans la suite de cette revue.

Machine Learning

Comme décrit dans le chapitre précédent, l'enjeu du traitement des données de haut volume est également analytique. En effet, dans la grande majorité des cas, ce type de données ne peut être analysé de manière robuste et efficace en utilisant les méthodes statistiques usuelles. Afin de comprendre les différences fondamentales d'approche entre les méthodes statistiques usuelles et les méthodes plus récentes d'apprentissage, nous opposons les visions paramétrique et non paramétrique.

Statistique paramétrique

La statistique paramétrique est le cadre classique de l'analyse statistique. En pratique, il sous-entend que le modèle statistique utilisé pour décrire la distribution des données peut être décrit par un nombre fini de paramètres. Typiquement, une distribution normale, par exemple, est totalement définie par 2 paramètres, la moyenne et l'écart type. Par extension, la modélisation statistique paramétrique suppose que les variables modélisées suivent une loi de distribution connue. Or, dans le cas de données massive provenant d'un nombre limité d'individus (ce qui est typiquement le cas lorsque l'on souhaite analyser des données médicales de multiples sources à l'échelle d'un seul patient pour prendre une décision individualisée), il est typiquement impossible de connaître la loi de distribution sous-jacente des données. L'utilisation de modèles paramétriques est, de ce fait, rendue impossible et erronée. Il convient donc d'avoir recours à d'autres types d'approches, plus souples, plus adaptatives, capables de d'apprendre des données sans faire d'hypothèse sur leur distribution sous-jacente. Ces méthodes sont typiquement qualifiées de non-paramétriques ou encore méthodes de machine learning et par extension (et abus de langage) d'intelligence artificielle.

Statistique non paramétrique (machine learning)

Par opposition, en statistique non paramétrique, on admet que le modèle ne peut être décrit par un nombre fini de paramètres. On s'autorise donc toutes les distributions possibles, c'est à dire qu'on ne fait aucune hypothèse sur la forme de la distribution des données.

On utilise typiquement ce type d'approche quand i) on n'arrive pas à ajuster correctement les observations avec une distribution paramétrique, ii) on n'a aucune idée du modèle sous-jacent ; iii) quand le nombre de variables est trop grand. Cette dernière situation est typiquement celle de la gestion des *big data*.

Les avantages de ces modèles sont qu'ils permettent de s'amender de tout a priori sur la distribution des données, les rendant ainsi plus robustes. Il existe néanmoins un inconvénient qui est lié à leur vitesse de convergence qui est plus lente que celle des modèles paramétriques. Autrement dit, ils sont moins puissants, ou nécessitent plus d'observations pour converger vers la réalité.

Nous allons voir comment la statistique non-paramétrique, et en particulier des nouvelles méthodes hybrides telles que le Super Learner [1], permet de prendre en charge des données de haut volume.

Predictive et Prescriptive Analytics

Il s'agit de 2 étapes à franchir pour pouvoir envisager l'utilisation des *big data* pour la médecine personnalisée en anesthésie et/ou en réanimation.

Predictive Analytics

Il s'agit de l'utilisation de méga-données et des algorithmes analytiques adaptés afin de pouvoir prédire la survenue d'un événement. Dans notre discipline, l'exemple le plus typique est la prédiction de la mortalité en réanimation.

Nous avons récemment démontré qu'un nouvel algorithme non paramétrique hybride (Super Learner [1]) peut être utilisé pour prédire la mortalité en réanimation de manière beaucoup plus précise et efficace qu'avec les modèles statistiques paramétriques usuels tels que la régression logistique. En effet, il a été largement démontré par le passé, que les scores actuels, tels que la SAPS2 [2] ou même les versions plus récentes discriminent correctement les survivants et les patients qui vont décéder en réanimation mais leur calibration est très mauvaise [3,4]. En effet, tous ces scores ont tendance à largement surestimer la mortalité.

Nous avons récemment démontré que le simple fait d'utiliser des algorithmes non paramétriques optimaux tels que le Super Learner permettait d'améliorer très significativement la performance prédictive [5], à la fois en termes de discrimination que de calibration. Ce type d'algorithmes dits « Stacked Supervised Machine Learning Algorithms » sont les méthodes de choix pour l'analyse des *big data* médicales.

Des versions *online* de ces algorithmes ont récemment été élaborées, permettant à ces outils d'apprendre constamment des nouvelles données générées et donc d'améliorer leurs performances au fil du temps. Ces versions *online* sont donc particulièrement adaptées à la prise en charge de *big data* évolutives dans le temps telles que celles provenant du monitoring continu. Récemment, Caelen et coll. ont rapporté d'excellentes performances prédictives pour des outils de machine learning dans la modélisation du signal de l'index bispectral par exemple.

Prescriptive Analytics

Il s'agit de la phase ultérieure grâce à laquelle l'analyse adaptée des méga-données permettra de suggérer la prescription au clinicien. Cette phase est également souvent décrite comme celle de l'aide à la décision. Grâce aux méthodes de online learning décrites ci-dessus, cette aide à la décision peut, dans un avenir proche, se concevoir en temps réel en fonction de l'évolution des paramètres vitaux monitorés.

Nous décrivons ci-dessous les fondements d'une approche statistiques dérivées des méthodes d'inférence causale, afin de pouvoir estimer le bénéfice d'un traitement à l'échelle de l'individu et non plus de la population. La définition et l'estimation de telles paramètres, permettra bien, si cela est couplé à des méga-données, de proposer au clinicien pour un patient donné le choix thérapeutique optimal. Il s'agit donc d'un changement de paradigme majeur par rapport à l'ère de l'evidence-based médecine au cours de laquelle les décisions thérapeutiques étaient essentiellement élaborées à partir des résultats d'essais cliniques qui eux-mêmes concluait à l'échelle populationnelle et non individuelle.

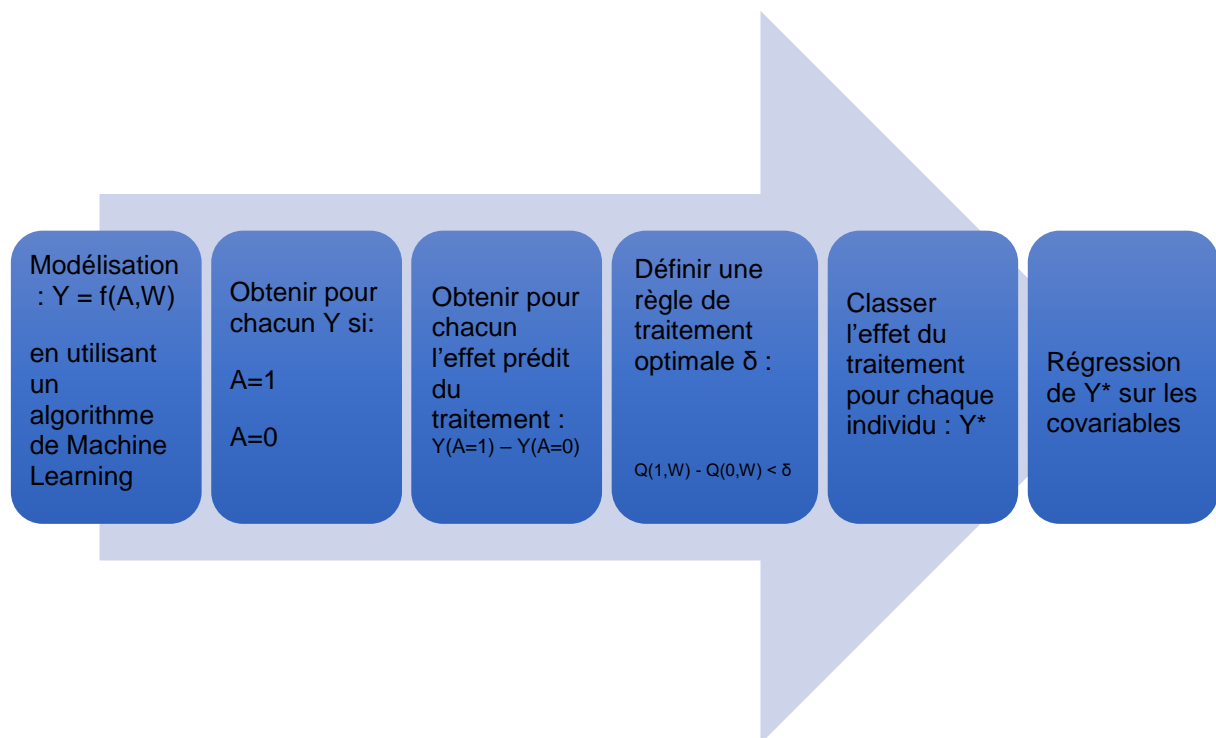


FIGURE 1. Exemple de « roadmap » d'analyse prescriptive basée sur le paradigme de l'inférence causale.

Comment ces technologies vont-elles changer nos pratiques ?

Les exemples existants

Certaines applications de technologies basées sur l'utilisation d'algorithmes d'intelligence artificielle ont déjà franchi le domaine de l'investigation et de la recherche pour mises être à disposition du clinicien.

Par exemple, le Acumen Hypotension Prediction Index (Edwards Lifesciences, Irvine, CA, USA) est outil de prédiction des évènements hypotensifs basé sur un algorithme d'apprentissage entraîné à prédire l'hypotension à partir d'une large base de données de patients. Cet outil est d'ores et déjà commercialisé même si son bénéfice en termes d'optimisation de la prise en charge reste à prouver.

Watson est un système informatique évolutif de réponse au questionnement utilisant une connaissance encyclopédique comme source d'information et un processus d'apprentissage du questionnement basé sur le *natural language processing* et le *machine learning* pour fournir une réponse optimale à une question ouverte. Si l'accès à ce système n'est pas encore du domaine de la pratique clinique quotidienne, Watson a déjà montré des performances intéressantes dans le domaine de l'analyse de cas cliniques complexes. Mais le domaine d'application de ce type d'outil basé sur l'apprentissage académique est plus le diagnostic que la prédiction ou l'aide à la décision à partir des données issues des patients.

Les risques et enjeux

Étant donné l'impact de l'intelligence artificielle sur notre quotidien, les attentes sont grandes quant à son utilisation dans le domaine de la santé. Les enjeux financiers sont également majeurs. Dans ce contexte, le risque est probablement de précipiter le transfert vers la pratique d'algorithmes encore imparfaits. Or en santé, l'erreur peut avoir des conséquences fatales pour le patient et de ce fait pour tout un axe de recherche. Il est donc urgent de prendre le temps de la recherche et du développement.

En outre, les enjeux en termes médico-légaux et d'anonymisation sont majeurs. La capture de données en temps réel est essentielle pour envisager d'utiliser ces outils en pratique en anesthésie ou en réanimation. Mais le volume de ce type de données oblige à les stocker de manière dématérialisée, hors de l'enceinte sécurisée de l'hôpital. Les enjeux de sécurité sont donc essentiels. De plus, la priorité de ces données est un sujet médico-légal qu'il convient de clarifier au plus vite. Enfin, si, dans un avenir proche, les algorithmes d'intelligence artificielle sont capables de proposer une attitude thérapeutique, à qui revient la responsabilité de la prescription ? Ou celle de suivre ou de ne pas suivre le « conseil de la machine » ? Autant de questions philosophiques et bioéthiques qui émergent dans le sillon de ces nouvelles technologies.

Conclusion

Les méthodes succinctement décrites ci-dessus permettent d'envisager, au cours des années à venir, un changement radical dans notre relation avec les outils de monitoring et les résultats biologiques et d'imagerie obtenus chez nos patients. En effet, ces données de très haut volume, dont la densité et la rapidité d'évolution sont impossibles à appréhender par le cerveau humain, seront analysées en temps réel par des algorithmes qui pourront i) prédire le risque de survenue de tel ou tel événement au cours des heures à venir et ii) fournir une aide précieuse à la décision thérapeutique en temps réel. Bien loin de la fiction, il ne nous faut pas appréhender cette mutation qui est salutaire. La machine ne sera pas là pour décider à la place du médecin (elle est loin d'être assez « intelligente » pour cela) mais fournira en revanche un nouveau type d'information résumée qu'il nous faudra apprendre à intégrer dans notre démarche médicale. Le schéma ci-dessous décrit un exemple d'architecture intelligente permettant d'utiliser les méga-données au lit du patient.

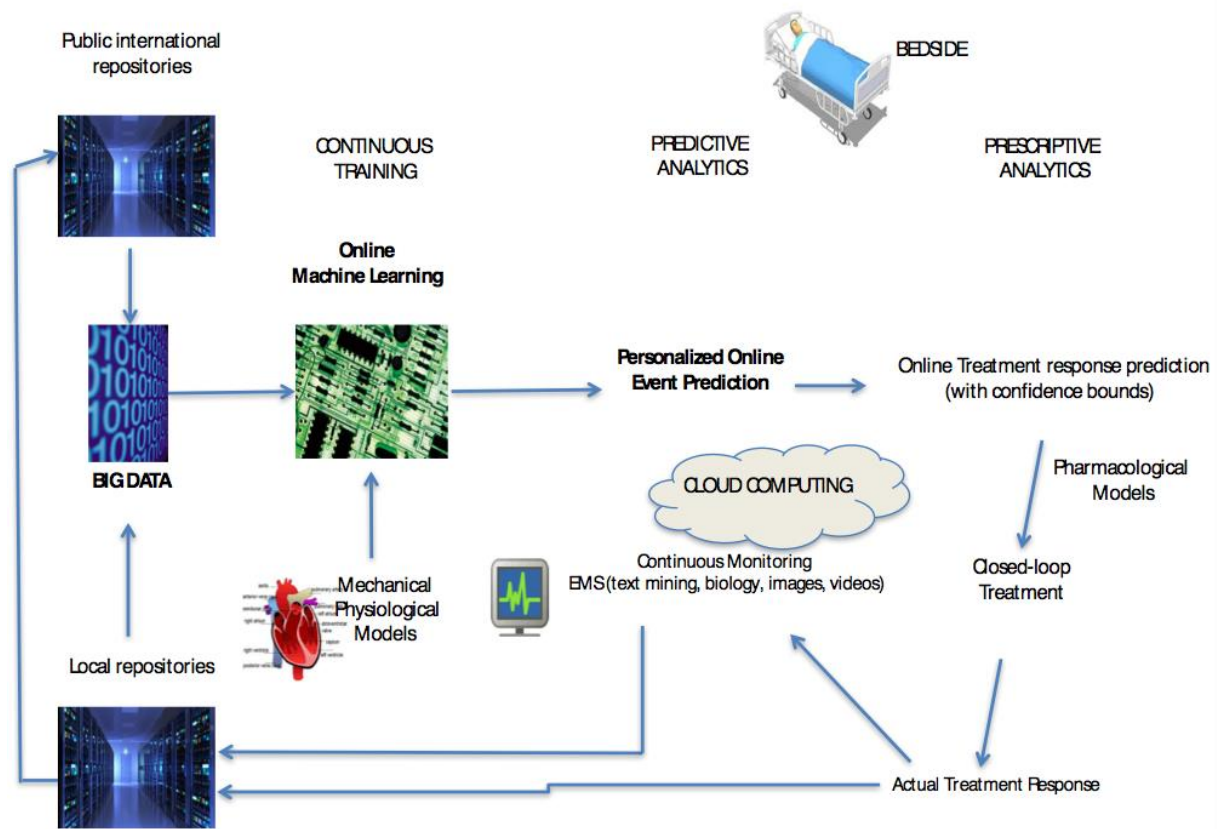


FIGURE 2. Proposition d'architecture d'acquisition et de traitement en temps réel des données d'anesthésie et/ou de réanimation.

Références

- [1] van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol*. 2007;6(1):1–21.
- [2] Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA J Am Med Assoc*. 1993 Dec 22;270(24):2957–63.
- [3] Capuzzo M, Moreno RP, Le Gall J-R. Outcome prediction in critical care: the Simplified Acute Physiology Score models. *Curr Opin Crit Care*. 2008 Oct;14(5):485–90.
- [4] Poole D, Rossi C, Latronico N, Rossi G, Finazzi S, Bertolini G. Comparison between SAPS II and SAPS 3 in predicting hospital mortality in a cohort of 103 Italian ICUs. Is new always better? *Intensive Care Med*. 2012 Aug;38(8):1280–8.
- [5] Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med*. 2015 Jan;3(1).