

AVIS DE L'EXPERT-E / EXPERT OPINION

Introduction aux enjeux éthiques de l'Intelligence Artificielle en réanimation

Ethical considerations about Artificial intelligence in critical care

Cyril Goulenok^{1*} • Marc Grassin² • Robin Cremer³ • Julien Duvivier⁴ • Caroline Hauw-Berlemont⁵
Mercé Jourdain⁶ • Antoine Lafarge⁷ • Anne-Laure Poujol⁸ • Nicolas de Prost⁹ • Jean-Philippe Rigaud¹⁰
Benjamin Zuber¹¹ • Bénédicte Gaillard-Leroux¹²

Reçu le 9 février 2024 ; accepté le 22 mars 2024.

© SRLF 2024.

Résumé

L'Intelligence Artificielle (IA) va occuper une place grandissante en réanimation dans les années à venir. Les possibilités d'utilisation seront très larges, couvrant notamment les champs des prédictions, de l'aide à la décision, de l'imagerie, de la recherche et la formation des soignants. L'implémentation de l'IA en réanimation pourrait significativement bouleverser la prise en charge des patients, mais aussi l'activité des médecins. Ces changements à venir amènent des questionnements éthiques spécifiques qu'il est nécessaire et urgent d'évoquer. De la mise au point d'un algorithme jusqu'à sa mise sur le marché, il existe plusieurs étapes qui nécessitent une vigilance éthique par une approche habituellement qualifiée d'*ethic by design*. Il est possible de subdiviser ces enjeux en quatre parties distinctes qui seront successivement abordées : les données, le patient, l'algorithme et le soignant. Les questions environnementales mais aussi économiques, dans le contexte de la récente prise de conscience d'une nécessaire diminution de l'empreinte carbone, seront enfin évoquées. Une meilleure appréhension des enjeux éthiques est probablement une première étape vers une intégration efficace de cette nouvelle technologie dans le domaine des soins critiques.

Mots-clés : Intelligence artificielle, USI, Éthique, prise de décision assistée par ordinateur, empreinte carbone

Abstract

Artificial Intelligence (AI) will play an increasingly pivotal role in intensive care over the coming years. The possibilities for its use will be very broad, covering the fields of prediction, decision-making support, imaging, research, and training of healthcare professionals. Implementation of AI in intensive care has the potential to substantially transform patient care approaches and physician practices. These future changes raise specific ethical considerations promptly. From the development of an algorithm to market deployment, there are several stages that require ethical vigilance, embodying an approach commonly known as "ethics by design". These ethical concerns can be subdivided into four distinct parts, which will be addressed in turn: the data, the patient, the algorithm and the healthcare workers. Additionally, environmental and economic issues will be explored, in the context of the recent awareness of the need to reduce our carbon footprint. A comprehensive understanding of the ethical issues involved is probably the first step towards the efficient integration of this innovative technology into critical care setting.

Keywords: Artificial Intelligence, ICU, Ethics, computer-assisted decision making, carbon footprint

Introduction

Si la naissance de l'Intelligence artificielle (IA) peut être datée du début des années 50, il aura fallu attendre

presque 60 ans pour en voir apparaître les applications. Nombreux sont, depuis, les secteurs concernés tels l'industrie, les transports, l'aéronautique. Dans le secteur de la santé, de multiples solutions émergent notamment

*Cyril Goulenok

Service de Médecine Intensive Réanimation, Hôpital Privé Jacques Cartier Ramsay-GDS, Massy. LIPHA.

✉ cyril.goulenok@gmail.com

La liste complète des affiliations des auteurs est disponible à la fin de l'article.



en imagerie, en dermatologie ou en anatomopathologie [1]. Si la réanimation reste encore en retrait, il est probable qu'en raison de son environnement très technologique, elle devienne un lieu propice à son développement. Lorsque les réanimateurs imaginent la réanimation du futur, celle-ci est constamment décrite comme plus technologique, plus connectée, plus personnalisée et moins invasive [2]. Une place importante est alors faite à l'IA qui permettrait de répondre à ces impératifs [3]. L'implémentation de l'IA en réanimation pourrait significativement bouleverser la prise en charge des patients, mais aussi l'activité des médecins. Ces changements à venir amènent des questionnements éthiques spécifiques qu'il est nécessaire et urgent d'évoquer. Il est possible de subdiviser ces enjeux en quatre parties distinctes qui seront successivement abordées : les données, le patient, l'algorithme et le soignant. Les enjeux de l'intégration de l'IA ne se limitent pas à ces seuls éléments et il faut donc s'interroger sur ses externalités et notamment

son impact environnemental à l'ère d'une ambition de décroissance dans le système de santé.

Émergence de l'IA en réanimation, de l'Idée à l'Application

Au sein de l'hôpital, la réanimation est un secteur privilégié pour le développement de l'IA. À court terme, de nombreuses applications pourraient voir le jour (Tableau 1). L'environnement spécifique de la réanimation peut être considéré comme un terreau fertile pour l'IA. En effet, le socle de fonctionnement de ce type d'unité consiste en la surveillance continue de différents paramètres vitaux par de multiples capteurs. Outre cette surveillance continue, s'agrègent différents dispositifs dont l'objectif est de suppléer aux dysfonctionnements d'organes. De nombreuses données sont donc collectées en continu et sont potentiellement exploitables par une IA qui justement y trouve sa matière première. Enfin, le développement

TABLEAU 1 - Exemples d'applications d'IA en réanimation

Systemes de prédiction	Mortalité [29] Apparition d'un sepsis [30] Embolie pulmonaire [31] Insuffisance rénale [32]
Aide à la décision	Remplissage <i>versus</i> catécholamines dans le choc septique [33] Réglage du respirateur artificiel [34] Sevrage de la ventilation mécanique [35]
Imagerie	Automatisation des calculs en échographie cardiaque [36] Analyse des radiographies de thorax (position de la sonde d'intubation par ex.) [37] Analyse TDM / IRM [38]
Recherche	Exploitation des données en NLP (<i>Natural Language Processing</i>) [39] Entrepôts de Données de Santé, Big Data [40] Aide à la rédaction d'articles scientifiques [41]
Formation	Intégration dans la simulation [42]
Logistique - Administratif	Évaluation du risque de réhospitalisation à la sortie de réanimation [43] Optimisation des flux de matériel [44]
Exploitation de nouveaux types de capteurs	Quantification de la mobilisation des patients [45] Évaluation de la douleur par système vidéo d'analyse faciale [46] Évaluation de la qualité du lavage des mains à l'entrée des chambres [47]
Éthique	Optimisation des procédures de LAT [48] Aide à la limitation des biais cognitifs ou implicites [49] Identification des proches à haut risque de PTSD

récent de nouveaux types de capteurs (sensoriels, visuels, sonores) favorise toujours plus la production de ces données et offre une nouvelle forme de surveillance appelée intelligence environnementale [4].

À cette complexité croissante s'ajoute la nécessité de prendre un nombre important de décisions pour le médecin avec des enjeux potentiellement majeurs en termes de morbi-mortalité [5]. En moyenne, une dizaine de décisions par heure sont prises dans une unité de soins critiques [6]. Elles nécessitent le plus souvent la prise en compte de très nombreux paramètres cliniques, biologiques, morphologiques. La possibilité d'avoir accès à des systèmes d'aide à la décision générés par des IA pourrait soulager partiellement les praticiens du poids de ces décisions. Si les publications se multiplient sur les systèmes d'aide à la décision [7], il y a plusieurs étapes à franchir avant la mise sur le marché et qu'il est nécessaire de décrire.

La mise en application d'un algorithme se basant sur une IA passe par un processus relativement long pouvant s'apparenter à celui d'une mise sur le marché d'un nouveau médicament avec une temporalité similaire. Il est possible de le segmenter en plusieurs étapes distinctes [8] : définition de la finalité de la solution d'IA, collecte des données, élaboration de l'algorithme à partir d'un set de données, validation de l'algorithme sur un autre set de données, test en vie réelle, évaluation clinique pour mise sur le marché. Une étape supplémentaire peut

être proposée qui consisterait à s'assurer, après cette dernière étape, du bon fonctionnement de l'algorithme, sorte d'audit *a posteriori* (Figure 1). À chacune de ces étapes émergent des questions éthiques qui doivent être traitées en même temps que la conception du produit, justifiant ainsi la mise en œuvre d'une démarche qualifiée *d'éthique by design* [9]. La pluridisciplinarité nécessaire pour l'élaboration de ces solutions numériques impose une collaboration tout au long de sa mise au point entre médecins, *data scientist*, ingénieurs et chargés de projet. Si peu d'études ont franchi le stade de l'évaluation clinique sous la forme d'essais randomisés contrôlés [10], la multiplication très récente de ces publications laisse augurer un changement radical sur la place de l'IA en médecine. Il est donc nécessaire d'évaluer les enjeux éthiques propres à celle-ci, tout au long des étapes de conception de l'algorithme jusqu'à sa mise sur le marché.

Enjeux éthiques spécifiques à l'IA

Données

L'IA ne peut se construire sans données. Ce sont elles qui permettent à l'algorithme de s'élaborer et surtout de s'ajuster progressivement pour arriver au résultat le plus précis possible. Il existe différentes formes de données allant de la *smart data* (données précises, orientées) au *big data* (données massives, non épurées). Ces don-

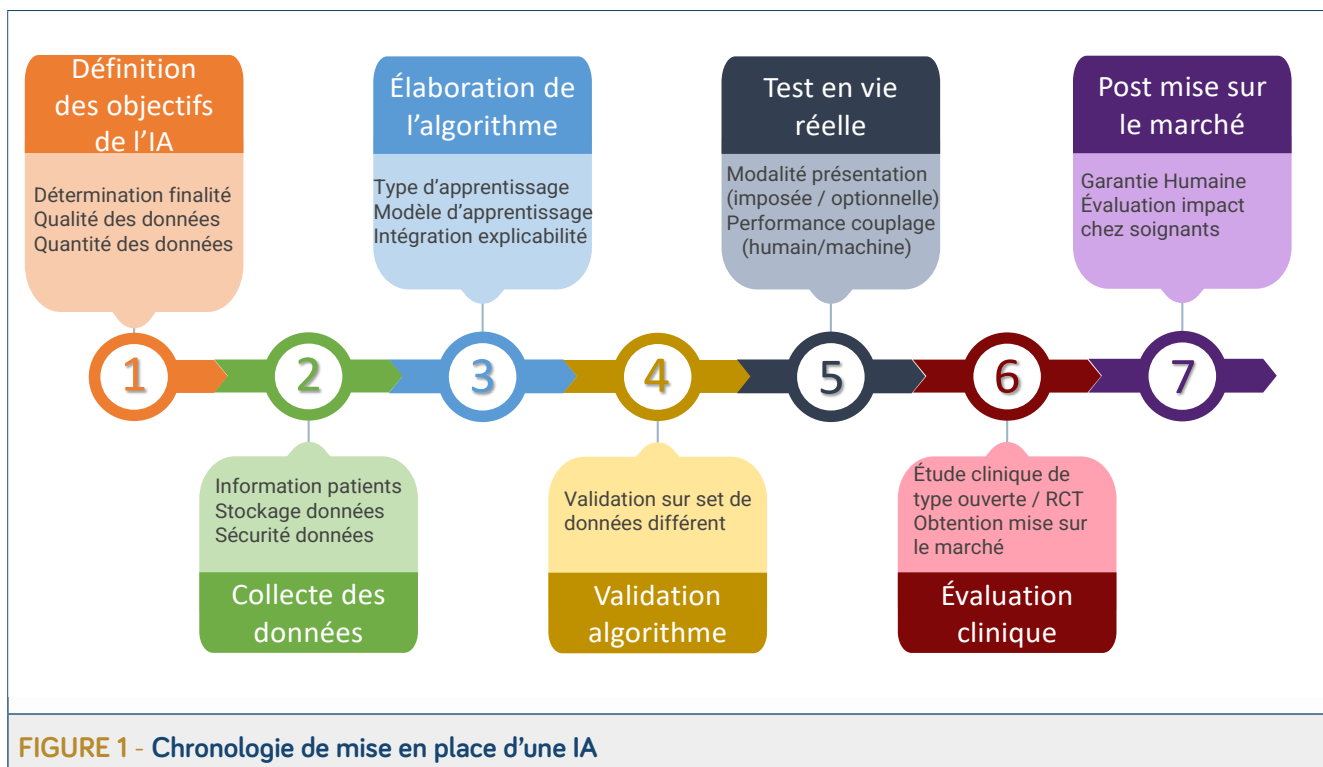


FIGURE 1 - Chronologie de mise en place d'une IA

nées peuvent être des variables numériques pour les données structurées (ex. ensemble des mesures hémodynamiques chez des patients de réanimation) ou des informations contenues dans le texte pour les données non structurées (ex. toutes les données intégrées dans un dossier informatique allant des paramètres biologiques aux résultats d'examens morphologiques, mais aussi aux notes des soignants, aux éléments administratifs, etc.). C'est certainement le premier enjeu éthique en IA. La quantité et la qualité des données sont absolument primordiales.

Quantité des données

Il doit y avoir une proportionnalité des données et une adéquation de ce qui est exploité avec la finalité de l'IA déterminée initialement. Pour illustrer ce propos, il n'est pas nécessaire d'exploiter l'ensemble des données d'une base d'un Entrepôt de Données de Santé (EDS) pour construire un score de prédiction de pneumopathie acquise sous ventilation mécanique (PAVM) en réanimation. Le premier élément de vigilance sur le plan éthique est donc de s'assurer que les données recueillies correspondent bien à l'objectif fixé initialement. En termes de quantité de données, il se trouve une limite mathématique contre-intuitive pour les cliniciens. Pour toute régularité statistique recherchée, il existe un nombre de données à partir desquelles on est certain de la trouver dans une base, y compris si cette base est constituée au hasard [11]. Il s'agit-là d'un problème nouveau ; la puissance de calcul disponible peut maintenant traiter des volumes de données gigantesques, mais un nombre de données trop élevé peut générer des régularités au hasard.

Qualité des données

Les modalités de recueil des données nécessitent une attention particulière puisque c'est à partir d'elles que se construira l'algorithme. Une bonne représentativité de genre, de classes sociales, d'ethnies est primordiale afin de limiter le risque de biais. En effet, il a été largement démontré que les algorithmes pouvaient reproduire des biais implicites d'ethnie ou de genre notamment [12]. Obermeyer *et al.* ont fait la preuve qu'une IA se basant sur les coûts de santé pour évaluer le besoin en soin négligeait les différences ethniques et sociales [13] : selon l'algorithme, les patients les plus pauvres, alors qu'ils sont plus malades, coûtent moins en soins. L'erreur de ce dernier provenait du fait que ces mêmes patients avaient moins recours à des soins coûteux, en raison de leur niveau de pauvreté faussant ainsi l'interprétation de l'algorithme. Un travail en amont est donc indispensable pour s'assurer de la qualité et la diversité des données. En l'absence de cet exercice, le risque serait que des praticiens trop enclins à se fier à l'IA, dans le cadre

d'une aide à la décision, verraient leurs biais renforcés et leurs préjugés majorés.

S'assurer de la diversité des populations incluses dans les bases de données semble donc nécessaire afin de limiter ce type de risque de biais (biais de sélection ou biais d'échantillonnage). La présence du paramètre de l'ethnie a ainsi pu être démontrée comme améliorant la performance d'un algorithme [14]. La législation diffère en fonction des pays qui autorisent, ou non, de relever ce paramètre dans les études cliniques, pouvant alors impacter la qualité de l'algorithme lors de sa conception. En France, il est interdit de relever ce paramètre et donc de l'inclure dans un algorithme ce qui pourrait alors en limiter sa performance. Par ailleurs, il faut souligner que les données socio-économiques ne sont que rarement renseignées dans les dossiers médicaux. Il y a ainsi une possible perte de qualité dans la conception des algorithmes et une impossibilité de corriger d'éventuels biais en amont. De façon très récente, des recommandations ont émergé afin de limiter au mieux les risques inhérents à ces biais ethniques recommandant notamment un accompagnement tout au long de la conception de l'algorithme (formulation du problème, évaluation et gestion des données, développement, formation et validation des algorithmes, déploiement et intégration de ces derniers) [15].

Enfin, la transparence sur l'origine des données est un sujet sensible posant à la fois la question éthique sur la façon dont ont été prélevées les données, leur accessibilité, mais aussi la possible reproductibilité des résultats. Récemment, confrontée à l'impossibilité d'apporter plus de détails sur l'origine de ses données, dans le cadre d'un audit, une équipe avait dû retirer une étude sur l'efficacité d'une thérapeutique dans le COVID [16].

Sécurité des données

L'extraction, l'exploitation puis le stockage de données issues de patients imposent une sécurisation très stricte de ces différentes étapes. La possibilité d'une diffusion sans l'accord du patient, d'un piratage des données ou d'une corruption de celles-ci faussant les résultats de l'IA, nécessite des protocoles élaborés par des spécialistes en sécurité informatique. Le stockage se doit d'être fait dans un serveur sécurisé certifié Hébergeur de données de santé (HDS) [17]. C'est donc un point de vigilance sur le plan éthique très important en raison des possibles conséquences en cas de non-respect de ces règles de sécurité.

Patients

Le patient est un acteur clé dans l'élaboration d'algorithmes puisque ce sont ses données de santé qui serviront à sa

construction. Il est indispensable qu'à toutes les étapes, une réflexion éthique soit entreprise afin de s'assurer du respect de la réglementation européenne sur le RGPD (Règlement Général sur la Protection des Données), mais aussi du consentement du patient et de l'existence d'une stratégie permettant d'éviter toute possibilité de réidentification.

Consentement éclairé

La protection du patient impose un consentement éclairé de sa part concernant l'utilisation de ses données de santé (RGPD, chapitre II, article 5). La plus récente déclaration de Taipei conforte cet impératif d'une information préalable avant exploitation des données [18]. Son application s'avère difficile chez des patients de réanimation pouvant présenter des troubles de la conscience allant jusqu'au coma. Afin de pallier cette difficulté pratique, une information de l'exploitation éventuelle de ces données doit être notifiée aussi bien dans le livret d'accueil du service de réanimation que dans le compte rendu hospitalier individuel délivré à la sortie du service. Cependant, il faut reconnaître la faiblesse de ces canaux d'information, le patient pouvant décéder durant son séjour ou ne pas forcément être destinataire de son compte-rendu d'hospitalisation. Si les patients doivent avoir la possibilité de savoir à quelles fins leurs données sont exploitées (site internet informant sur les études en cours dans un service ou un hôpital) et avoir la possibilité de s'opposer à leur exploitation, il s'avère que la mise en pratique est bien plus difficile du fait de cette difficulté d'information.

Enfin, Le développement récent du stockage de données massives au sein des Entrepôts de Données de Santé (EDS) interroge sur la possibilité d'informer l'ensemble des patients même s'il faut reconnaître les efforts considérables des établissements de santé sur ce sujet.

Risque de réidentification

Selon la façon dont sont structurées les données, il existe une possibilité de réidentification. Une stratégie d'anonymisation de la base permet de limiter ce risque. Trois stratégies de vérification de l'efficacité de l'anonymisation sont possibles : l'individualisation (impossibilité d'isoler un individu sur un jeu de données), la corrélation (impossibilité de relier entre eux des ensembles de données distincts concernant le même individu), l'inférence (impossibilité de déduire de façon quasi-certaine de nouvelles informations sur un individu) [19]. En recherche médicale, c'est la pseudonymisation qui est le plus souvent privilégiée en introduisant une étape intermédiaire ou séquentielle masquant les données sensibles par un pseudonyme ou un artefact. Il est impératif que les modalités de pseudonymisation

des données soient décrites dans le protocole d'étude. Il peut cependant y avoir un bénéfice à la réidentification dans le cas où, par exemple, un algorithme permet le diagnostic d'un processus tumoral. Cette possibilité reviendrait alors au comité de pilotage de l'EDSH.

Place des proches

Insuffisamment pris en compte, les proches ont une importance considérable dans le contexte spécifique de la réanimation où le patient n'est le plus souvent pas apte à communiquer avec l'équipe soignante. L'intégration de solutions d'IA dans les processus de LAT pourrait justifier auprès d'eux une information spécifique. Par ailleurs, les recherches récentes sur les risques de syndrome de stress post-traumatique chez les proches (PTSD) laissent envisager le développement d'outils d'IA afin de mieux identifier les personnes à risque. Les questions éthiques devront impérativement prendre en compte cette population spécifique.

Algorithme

Un algorithme se construit à partir d'une base de données déjà disponible ou créée à cet effet. Il existe déjà un certain nombre de bases de données pouvant être utilisées, soit pour mettre au point un algorithme, soit pour en tester, dans un second temps, sa validité et son efficacité. Sont à disposition, selon la question à résoudre, différents types d'apprentissage (supervisé, non supervisé, par renforcement, automatique ou profond). S'il existe un réel souci du partage dans la communauté des chercheurs, cela semble un peu moins flagrant dès lors qu'une commercialisation est envisagée.

Algorithme figé ou autoapprenant ouvert

Lors de son élaboration, l'algorithme peut être défini comme fermé ou plutôt figé (donc non évolutif) ou rester au contraire ouvert, c'est-à-dire continuer à se modifier au fur et à mesure de son utilisation. Dans ce dernier cas, il peut s'adapter à l'environnement dans lequel il est utilisé. Cette option comporte des risques, car son efficacité peut évoluer dans le temps soit en s'améliorant, mais aussi possiblement en se dégradant. Cette approche impose donc des évaluations très régulières de l'algorithme dit « autoapprenant » (ou apprentissage continu) afin de s'assurer de l'absence de diminution de performance. Pour cette raison, très majoritairement, les algorithmes soumis à validation dans le domaine de la santé sont de type figé même si une dérive de performance reste envisageable. Des évaluations *a posteriori* doivent être envisagées pour ces derniers justifiant la création d'un label *Garantie humaine* [20] dont l'objectif est d'apporter une assurance éthique après introduction de la solution algorithmique.

Explicabilité

Le rendu d'un résultat issu d'une IA peut être brut, c'est-à-dire sans possibilité d'explications sur les raisons ayant abouti à celui-ci ou, au contraire, accompagné d'une description permettant à l'utilisateur de mieux comprendre les raisons du résultat. C'est la deuxième option qui doit être privilégiée pour les algorithmes développés en santé. Lors de l'utilisation d'une IA générative telle que *ChatGPT 3.5*, un résultat est donné à une interrogation sans explication sur le chemin pris par l'algorithme pour y aboutir. Ce phénomène, appelé plus couramment *blackbox*, prive l'utilisateur de la possibilité de comprendre et donc de critiquer un résultat qui lui est ainsi imposé. Dans les systèmes d'aide à la décision où le praticien doit garder une position de validation finale, il est nécessaire qu'il puisse avoir accès à un minimum d'éléments explicatifs. Une solution d'IA accompagnée d'un système explicatif (plus couramment appelé *X. AI*) existe déjà sous différentes formes tel le système SHAP (*Shapley Additive Explanations*). Dans une étude en pédiatrie, un algorithme d'IA déterminait une série d'examen à réaliser par le patient dès son arrivée aux urgences, et ce, avant qu'il soit vu par le médecin. Ce dernier disposait ensuite, au moment de la consultation, d'un graphique expliquant pourquoi l'algorithme avait décidé de réaliser en amont tel examen et non pas tel autre [21]. Cependant, si un système explicatif paraît de prime abord pertinent, son bénéfice reste encore sujet à controverse [22]. Une IA de confiance passe par l'explicabilité ou à défaut par une transparence sur les finalités des « solutions » proposées et de l'usage afin de garantir et ne pas entraver la relation aux patients. L'IA de confiance est une condition nécessaire pour favoriser son utilisation par le soignant.

Soignants

S'il est impératif de démontrer qu'il y a un bénéfice pour le patient lors de la mise en place d'un système algorithmique, il est tout aussi important de ne pas négliger l'impact que cela peut avoir chez le soignant en charge de celui-ci. Positionné le plus souvent comme un système d'aide à la décision, il laisse à l'humain le statut de décideur final. Malgré ce processus en apparence vertueux, il est nécessaire d'éclairer sur les conséquences éthiques de cette nouvelle approche. Le mode de présentation du résultat, l'absence d'explication, une temporalité bouleversée, la performance du couplage humain/machine, sont quelques-uns des sujets d'attention à ne pas négliger afin de s'assurer de la préservation de l'autonomie du soignant.

Modalité de délivrance des résultats de l'IA

L'influence que peut avoir une proposition d'aide à la décision qu'elle soit d'emblée imposée au soignant ou

qu'elle soit au contraire optionnelle, impose une réflexion dès sa mise en place. Lorsque le résultat de l'IA apparaît sans sollicitation, ce qui pourrait être qualifié d'approche par imposition, l'humain a plus tendance à suivre la recommandation. *A contrario*, si l'aide est optionnelle, il a tendance à élaborer une critique plus structurée du résultat proposé par l'IA [23]. Dans cette deuxième approche, le temps nécessaire à la réflexion est respecté, évitant un possible conflit de temporalité au cours duquel l'extrême rapidité de réponse d'une IA se heurterait à un temps décisionnel bien plus long pour un humain. Il est donc nécessaire qu'en amont de l'intégration de la solution d'IA, soit abordé le sujet de la modalité d'apparition du résultat sous une forme imposée d'emblée ou secondairement après accord du praticien, voire sous une forme optionnelle. Cette dernière approche respecterait au mieux une autonomie décisionnelle, socle de l'exercice médical.

Relation humain/machine

L'apport d'une aide est classiquement considéré comme une plus-value dans un processus décisionnel. Une nécessaire vigilance est requise afin d'être sûr que le médecin puisse être en mesure de formuler une critique du résultat produit par l'IA. En effet, si la question posée s'avère trop complexe pour lui ou s'il n'a pas les outils nécessaires pour formuler cette critique, il se retrouvera dans une position de soumission à la technologie. L'autre risque, non négligeable, serait celui de l'installation d'une trop grande confiance envers la machine aboutissant à un biais d'automatisation. Dans ce cas, le médecin ne chercherait plus à développer une quelconque critique du résultat proposé et suivrait systématiquement la recommandation de l'IA. À terme, il perdrait la maîtrise du sujet pour lequel l'IA l'aurait finalement suppléé. De façon très récente, il a pu être démontré qu'une IA intégrant un système d'explicabilité pouvait augmenter les performances des médecins lors de l'interprétation de radiographies, mais aussi possiblement induire en erreur le praticien en cas d'introduction intentionnelle de biais [24].

Responsabilités

La possibilité d'une IA autonome sans intervention humaine semble à ce jour peu probable dans l'environnement de la réanimation. Il y aura donc nécessairement une intervention humaine dont l'un des objectifs sera la validation, ou non, du résultat produit par la technologie. Cette liberté laissée au soignant conduit cependant à des remarques. En effet, par sa position de décideur final, c'est l'entièreté de sa responsabilité qui reste engagée, quel que soit le résultat. S'il valide une proposition erronée, il reste responsable en tant qu'acteur de cette validation. *A contrario*, dans la mesure où le soignant

décide de ne pas suivre la recommandation d'une IA, cette liberté qui lui a été offerte pourrait se retourner contre lui. Pour illustration se développent des systèmes de prédiction de sepsis [25] permettant de détecter avec précision, chez des patients hospitalisés, ceux à fort risque de présenter une infection à court terme. Il y a une réduction de mortalité lors de l'utilisation de ces systèmes prédictifs. La question de la responsabilité d'un clinicien décidant de ne pas suivre l'alerte formulée par l'IA amène à la réflexion plus générale de la réelle autonomie décisionnelle du médecin.

Le développement de l'IA face à la question environnementale

Il est difficile de ne pas évoquer le sujet de "l'empreinte carbone" de l'IA à l'heure où justement, une prise de conscience amène les acteurs du système de santé à avoir pour objectif de réduire significativement les GES (Gaz à Effets de Serre). Selon le Shift Project [26], il est estimé que 8 % des GES émis en France sont attribuables au système de santé dont 37 % reviennent au système hospitalier (public et privé confondus). Un objectif de réduction de 80 % de ces émissions a été fixé pour 2050, imposant donc une urgente cartographie des secteurs participants à ces GES afin de proposer des actions en regard. Le numérique représente une part non négligeable au sein de l'hôpital et l'implémentation d'une nouvelle technologie comme l'IA risque de se heurter à cette ambition de décroissance. Il n'existe que très peu de données sur l'empreinte carbone des solutions d'IA en santé. On peut s'essayer à une estimation en individualisant deux catégories : empreinte directe et indirecte. Dans l'empreinte directe, on intègre la conception de l'algorithme, son entraînement et ensuite son utilisation courante en système hospitalier. La dernière est très peu énergivore, la plupart des algorithmes étant de type fermé avec ainsi une consommation énergétique qui reste faible. C'est la première phase qui s'avère dispendieuse sur le plan énergétique, même si cela n'égale pas les IA génératives telles que *ChatGPT* par exemple. Le plus intéressant est certainement l'empreinte indirecte dans laquelle s'intègre la numérisation des données de santé. Si actuellement, il est estimé que seulement 1 % de ces données sont numérisées, la croissance exponentielle de la numérisation fait comprendre dans quelle mesure une majoration de l'empreinte carbone pourra être observée. La numérisation, mais aussi et surtout, le stockage de ces données risque de participer de façon toujours plus significative à l'augmentation des GES imputables au système de santé. Les externalités d'une technologie englobent toutes les conséquences autres que celles pour lesquelles elle a été conçue. Pour une IA, l'empreinte carbone qu'elle génère pour

sa conception puis son utilisation est partie intégrante de ses externalités. Dans une réflexion plus globale, à l'échelle planétaire, il devient difficile de négliger ces conséquences indirectes. Face à cette injonction contradictoire, il est urgent de définir les limites de ce qui est acceptable en matière de bénéfice pour une nouvelle technologie telle que l'IA en regard de ses conséquences écologiques.

IA et profit

La question du bénéfice économique dans un système de santé aux ressources contraintes et limitées se pose. Le développement de l'IA et du secteur des données de santé est un marché qui nécessite des investissements importants en termes de recherche et développement. Les entreprises qui développent ces solutions et les vendent ne sont pas philanthropiques et recherchent un profit. L'implémentation de l'IA et son usage méritent d'être évalués, non seulement en termes d'efficacité et d'effet d'impact pour les soignants et les patients, mais aussi en termes d'efficience économique et d'allocations au regard du coût des ressources disponibles et ce de manière écosystémique. Les profits légitimes réalisés par les entreprises privées, par la mise à disposition des solutions d'IA de leur sécurité et de leur amélioration, doivent être évalués en regard de ce qu'ils obligent en termes de choix dans un système de santé financé par la solidarité nationale. L'IA de santé n'échappe pas à l'enjeu d'une éthique économique-sociale dans laquelle des arbitrages financiers de la santé sont toujours à faire.

Perspectives dans l'urgence

La mise au point d'algorithmes basés sur de l'IA en réanimation s'inscrit dans une courbe de croissance, dont il est difficile de dire si elle atteindra un point d'inflexion ou au contraire prendra une allure exponentielle. Les utilisateurs futurs sont encore très mal préparés à cette transformation qui leur imposera l'utilisation d'outils avec lesquels ils ne sont pas familiarisés. La première urgence est donc représentée par la formation des futurs utilisateurs sur le fonctionnement des IA, les façons d'appréhender l'outil et les questions éthiques qu'elles soulèvent. Il faut préparer les cliniciens à cette transformation inéluctable qu'ils devront maîtriser [27]. Il est urgent de développer la recherche sur les questions éthiques spécifiques à l'IA afin qu'une approche critique soit envisageable sur l'ensemble des étapes de création des algorithmes. Une prise de conscience de la complexité éthique des systèmes d'aide à la décision amène à la question du couplage entre humain et machine. Il doit être un sujet d'attention spécifique afin d'évaluer les performances et déterminer au mieux les modes de collaboration de ces

deux formes d'intelligence. Enfin, l'inévitable confrontation avec le sujet de l'urgence écologique impose, outre une prise de conscience, une cartographie des différents éléments participants à l'empreinte carbone spécifique aux unités de soins critiques [28]. Une sensibilisation associée à cette analyse préliminaire permettra de mieux déterminer les attentes d'une nouvelle technologie en la confrontant à ses externalités.

Conclusion

La réanimation est un terreau propice au développement et à l'utilisation de l'Intelligence Artificielle. Cette nouvelle technologie pourrait considérablement améliorer la prise en charge des patients de soins critiques. Il n'est pas envisageable, à ce jour, qu'elle fonctionne en pleine autonomie, car elle doit être supervisée par ses utilisateurs, les médecins de réanimation. Des sujets éthiques existent à tous les niveaux, de la mise au point d'algorithmes d'IA à l'implémentation finale au sein des unités de réanimation. Il est indispensable de les appréhender. Ces sujets ne peuvent s'envisager *a posteriori* et doivent être explorés tout au long des étapes de création dans une démarche d'éthique *by design*. Les quatre piliers de la réflexion éthique sur l'IA sont les données, l'algorithme, les patients et les soignants. Une attention particulière doit être portée aux processus d'aide à la décision. Enfin, la question des externalités d'une nouvelle technologie et notamment son empreinte carbone ne peut plus être négligée à l'aune d'un crucial réveil écologique. Leur prise en compte s'avèrera indispensable afin de respecter les objectifs de décroissance et estimer au mieux les réels bénéfices de l'IA en réanimation.

Conflits d'intérêts

Les auteurs déclarent ne pas avoir de conflit d'intérêts.

Affiliations

¹Service de Médecine Intensive Réanimation, Hôpital Privé Jacques Cartier Ramsay-GDS, Massy. LIPHA.

²Faculté de philosophie de l'ICP.

³Réanimation et soins intensifs pédiatriques
Hôpital Jeanne de Flandre – CHU de Lille.

⁴Service de réanimation polyvalente / Usv, CH de Draguignan

⁵Service de Médecine Intensive Réanimation
Hôpital européen Georges Pompidou, AP-HP, Paris.

⁶Service de Réanimation C - Pôle de MIR
Hôpital R Salengro – CHU de Lille.

⁷Service de Médecine Intensive Réanimation
Hôpital Saint Louis, AP-HP, Paris

⁸Équipe VCR, École de Psychologues Praticiens de l'Institut Catholique de Paris – Service de réanimation chirurgicale polyvalente
Groupe Hospitalier Pitié Salpêtrière, AP-HP, Paris.

⁹Service de Médecine Intensive Réanimation
GHU Henri Mondor, DMU Médecine, AP-HP, Créteil

¹⁰Service de Médecine Intensive Réanimation, CH de Dieppe

¹¹Service de Réanimation Polyvalente
Hôpital Foch, Suresnes.

¹²Service de Réanimation Pédiatrique
Hôpital Femme-Mère-Enfant, CHU de Nantes.

Références

1. Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 25:44–56. DOI : 10.1038/s41591-018-0300-7
2. Vincent J-L, Creteur J (2019) Critical care medicine in 2050: less invasive, more connected, and personalized. *J Thorac Dis* 11:335–338. DOI : 10.21037/jtd.2018.11.66
3. Benbenishty JS, Bülow H-H (2018) Intensive care medicine in 2050: multidisciplinary communication in-/outside ICU. *Intensive Care Med* 44:636–638. DOI : 10.1007/s00134-017-4915-9
4. Davoudi A, Malhotra KR, Shickel B, et al (2019) Intelligent ICU for Autonomous Patient Monitoring Using Pervasive Sensing and Deep Learning. *Sci Rep* 9:8020. DOI : 10.1038/s41598-019-44004-w
5. Goulenok C, Cohen-Solal Z, David R, et al (2022) La décision médicale en réanimation, un indispensable fardeau. *Médecine Intensive Réanimation*, 31(2). DOI : 10.37051/mir-00101
6. Lundgrén-Laine H, Kontio E, Perttilä J, et al (2011) Managing daily intensive care activities: An observational study concerning ad hoc decision making of charge nurses and intensivists. *Crit Care* 15:R188. DOI : 10.1186/cc10341
7. Hong N, Liu C, Gao J, et al (2022) State of the Art of Machine Learning-Enabled Clinical Decision Support in Intensive Care Units: Literature Review. *JMIR Med Inform* 10:e28781. DOI : 10.2196/28781
8. van de Sande D, van Genderen ME, Huisken J, et al (2021) Moving from bytes to bedside: a systematic review on the use of artificial intelligence in the intensive care unit. *Intensive Care Med* 47:750–760. DOI : 10.1007/s00134-021-06446-7
9. Fischer F (2019) L'éthique by design du numérique : généalogie d'un concept. *Sciences du Design* n° 10:61–67. DOI : 10.3917/sdd.010.0061
10. Lång K, Josefsson V, Larsson A-M, et al (2023) Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. *The Lancet Oncology* 24:936–944. DOI : 10.1016/S1470-2045(23)00298-X
11. Longo G (2023) Le cauchemar de Prométhée: Les sciences et leurs limites. PUF, Paris, p206.
12. Courtland R (2018) As machine learning infiltrates society, scientists grapple with how to make algorithms fair. *Nature Jun*;558(7710):357–360. DOI : 10.1038/d41586-018-05469-3.
13. Obermeyer Z, Powers B, Vogeli C, Mullainathan S (2019) Dissecting racial bias in an algorithm used to manage the health of populations. *Science Oct* 25;366(6464):447–453. DOI : 10.1126/science.aax2342.
14. Segar MW, Hall JL, Jhund PS, et al (2022) Machine Learning-Based Models Incorporating Social Determinants of Health vs Traditional Models for Predicting In-Hospital Mortality in Patients With Heart

- Failure. *JAMA Cardiol*. DOI : [10.1001/jamacardio.2022.1900](https://doi.org/10.1001/jamacardio.2022.1900)
15. Chin MH, Afsar-Manesh N, Bierman AS, et al (2023) Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care. *JAMA Netw Open* 6:e2345050. DOI : [10.1001/jamanetworkopen.2023.45050](https://doi.org/10.1001/jamanetworkopen.2023.45050)
 16. Mehra MR, Desai SS, Ruschitzka F, Patel AN (2020) Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *The Lancet* S0140673620311806. DOI : [10.1016/S0140-6736\(20\)31180-6](https://doi.org/10.1016/S0140-6736(20)31180-6)
 17. HDS. In: Agence du Numérique en Santé. <https://esante.gouv.fr/produits-services/hds>. Accessed 11 Sep 2023
 18. WMA - The World Medical Association-Déclaration de l'AMM sur les considérations éthiques concernant les bases de données de santé et les biobanques. <https://www.wma.net/fr/policies-post/declaration-de-lamm-sur-les-considerations-ethiques-concernant-les-bases-de-donnees-de-sante-et-les-biobanques/>. Accessed 10 Jan 2024
 19. L'anonymisation de données personnelles. <https://www.cnil.fr/fr/lanonymisation-de-donnees-personnelles>. Accessed 6 Feb 2024
 20. Gruson D, Kirchner C (2019) La « garantie humaine », levier majeur de régulation des enjeux éthiques associés au numérique et à l'intelligence artificielle en santé. *ADPS (Actualité et Dossier en Santé Publique)* 108, Sept :41-42.
 21. Singh D, Nagaraj S, Mashouri P, et al (2022) Assessment of Machine Learning-Based Medical Directives to Expedite Care in Pediatric Emergency Medicine. *JAMA Netw Open* 5:e222599. DOI : [10.1001/jamanetworkopen.2022.2599](https://doi.org/10.1001/jamanetworkopen.2022.2599)
 22. Nagendran M, Festor P, Komorowski M, et al (2023) Quantifying the impact of AI recommendations with explanations on prescription decision making. *npj Digit Med* 6:206. DOI : [10.1038/s41746-023-00955-z](https://doi.org/10.1038/s41746-023-00955-z)
 23. Desbiolles J-P, Colombet G (2023) Humain ou IA ? Qui décidera le futur ? : Défis et opportunités d'un monde où l'IA nous dépasse. Dunod
 24. Jabbour S, Fouhey D, Shepard S, et al (2023) Measuring the Impact of AI in the Diagnosis of Hospitalized Patients. *JAMA* Dec 19;330(23):2275-2284. DOI : [10.1001/jama.2023.22295](https://doi.org/10.1001/jama.2023.22295).
 25. Adams R, Henry KE, Sridharan A, et al (2022) Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med* 28:1455-1460. DOI : [10.1038/s41591-022-01894-0](https://doi.org/10.1038/s41591-022-01894-0)
 26. Notre ambition - The Shift Project, think tank de la transition carbone. In: The Shift Project. <https://theshiftproject.org/ambition/>. Accessed 17 Nov 2023
 27. James CA, Wachter RM, Woolliscroft JO (2022) Preparing Clinicians for a Clinical World Influenced by Artificial Intelligence. *JAMA*. DOI : [10.1001/jama.2022.3580](https://doi.org/10.1001/jama.2022.3580)
 28. See KC (2023) Improving environmental sustainability of intensive care units: A mini-review. *World J Crit Care Med* Sep 9;12(4):217-225. doi: [10.5492/wjccm.v12.i4.217](https://doi.org/10.5492/wjccm.v12.i4.217).
 29. Chen Y, Li Y, Deng P, et al (2022) Learning to predict in-hospital mortality risk in the intensive care unit with attention-based temporal convolution network. *BMC Anesthesiol* 22:119. DOI : [10.1186/s12871-022-01625-5](https://doi.org/10.1186/s12871-022-01625-5)
 30. Giannini HM, Ginestra JC, Chivers C, et al (2019) A Machine Learning Algorithm to Predict Severe Sepsis and Septic Shock: Development, Implementation, and Impact on Clinical Practice*. *Critical Care Medicine* 47:1485-1492. DOI : [10.1097/CCM.0000000000003891](https://doi.org/10.1097/CCM.0000000000003891)
 31. Banerjee I, Sofela M, Yang J, et al (2019) Development and Performance of the Pulmonary Embolism Result Forecast Model (PERFORM) for Computed Tomography Clinical Decision Support. *JAMA Netw Open* 2:e198719. DOI : [10.1001/jamanetworkopen.2019.8719](https://doi.org/10.1001/jamanetworkopen.2019.8719)
 32. Flechet M, Falini S, Bonetti C, et al (2019) Machine learning versus physicians' prediction of acute kidney injury in critically ill adults: a prospective evaluation of the AKIpredictor. *Crit Care* 23:282. DOI : [10.1186/s13054-019-2563-x](https://doi.org/10.1186/s13054-019-2563-x)
 33. Komorowski M, Celi LA, Badawi O, et al (2018) The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 24:1716-1720. DOI : [10.1038/s41591-018-0213-5](https://doi.org/10.1038/s41591-018-0213-5)
 34. Peine A, Hallawa A, Bickenbach J, et al (2021) Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care. *npj Digit Med* 4:32. DOI : [10.1038/s41746-021-00388-6](https://doi.org/10.1038/s41746-021-00388-6)
 35. Parreco J, Hidalgo A, Parks JJ, et al (2018) Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement. *Journal of Surgical Research* 228:179-187. DOI : [10.1016/j.jss.2018.03.028](https://doi.org/10.1016/j.jss.2018.03.028)
 36. Madani A, Arnaout R, Mofrad M, Arnaout R (2018) Fast and accurate view classification of echocardiograms using deep learning. *npj Digital Med* 1:6. DOI : [10.1038/s41746-017-0013-1](https://doi.org/10.1038/s41746-017-0013-1)
 37. Brown MS, Wong K-P, Shrestha L, et al (2023) Automated Endotracheal Tube Placement Check Using Semantically Embedded Deep Neural Networks. *Academic Radiology* 30:412-420. DOI : [10.1016/j.acra.2022.04.022](https://doi.org/10.1016/j.acra.2022.04.022)
 38. Zhang K, Liu X, Shen J, et al (2020) Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. *Cell* 182:1360. DOI : [10.1016/j.cell.2020.08.029](https://doi.org/10.1016/j.cell.2020.08.029)
 39. Clapp MA, Kim E, James KE, et al (2022) Comparison of Natural Language Processing of Clinical Notes With a Validated Risk-Stratification Tool to Predict Severe Maternal Morbidity. *JAMA Netw Open* 5:e2234924. DOI : [10.1001/jamanetworkopen.2022.34924](https://doi.org/10.1001/jamanetworkopen.2022.34924)
 40. Boss JM, Narula G, Straessle C, et al (2022) ICU Cockpit: a platform for collecting multimodal waveform data, AI-based computational disease modeling and real-time decision support in the intensive care unit. *Journal of the American Medical Informatics Association* 29:1286-1291. DOI : [10.1093/jamia/ocac064](https://doi.org/10.1093/jamia/ocac064)
 41. Hutson M (2022) Could AI help you to write your next paper? *Nature* 611:192-193. DOI : [10.1038/d41586-022-03479-w](https://doi.org/10.1038/d41586-022-03479-w)
 42. Mathur P, Burns ML (2019) Artificial Intelligence in Critical Care. *International Anesthesiology Clinics* 57:89-102. DOI : [10.1097/AIA.0000000000000221](https://doi.org/10.1097/AIA.0000000000000221)
 43. Lin Y-W, Zhou Y, Faghri F, et al (2019) Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. *PLoS ONE* 14:e0218942. DOI : [10.1371/journal.pone.0218942](https://doi.org/10.1371/journal.pone.0218942)
 44. Zhang J, Pathak HS, Snowdon A, Greiner R (2022) Learning models for forecasting hospital resource utilization for COVID-19 patients in Canada. *Sci Rep* 12:8751. DOI : [10.1038/s41598-022-12491-z](https://doi.org/10.1038/s41598-022-12491-z)

45. Yeung S, Rinaldo F, Jopling J, et al (2019) A computer vision system for deep learning-based detection of patient mobilization activities in the ICU. *npj Digit Med* 2:11. DOI : [10.1038/s41746-019-0087-z](https://doi.org/10.1038/s41746-019-0087-z)
46. Fontaine D, Vielzeuf V, Genestier P, et al (2022) Artificial intelligence to evaluate postoperative pain based on facial expression recognition. *European Journal of Pain* 26:1282–1291. DOI : [10.1002/ejp.1948](https://doi.org/10.1002/ejp.1948)
47. Haque A, Guo M, Alahi A, et al (2017) Towards Vision-Based Smart Hospitals: A System for Tracking and Monitoring Hand Hygiene Compliance. Proceedings of the 2nd Machine Learning for healthcare conference, 68, 75-87 (PMLR, 2017).
48. Biller-Andorno N, Ferrario A, Joebges S, et al (2022) AI support for ethical decision-making around resuscitation: proceed with care. *J Med Ethics* 48:175–183. DOI : [10.1136/medethics-2020-106786](https://doi.org/10.1136/medethics-2020-106786)
49. Chen Y, Clayton EW, Novak LL, et al (2023) Human-Centered Design to Address Biases in Artificial Intelligence. *J Med Internet Res* 25:e43251. DOI : [10.2196/43251](https://doi.org/10.2196/43251)