



COMITÉ CONSULTATIF NATIONAL D'ÉTHIQUE
POUR LES SCIENCES DE LA VIE ET DE LA SANTÉ

COMITÉ NATIONAL PILOTE
D'ÉTHIQUE DU NUMÉRIQUE

sous l'égide du
COMITÉ CONSULTATIF NATIONAL D'ÉTHIQUE
POUR LES SCIENCES DE LA VIE ET DE LA SANTÉ

AVIS COMMUN
AVIS 141 CCNE / AVIS 4 CNPEN

**Diagnostic Médical et Intelligence Artificielle :
Enjeux Ethiques**

Avis adopté le 23 novembre 2022 à l'unanimité des membres présents lors de l'assemblée plénière du CNPEN

Avis adopté le 24 novembre 2022 à l'unanimité des membres présents lors de l'assemblée plénière du CCNE

Comment citer cet avis :

Diagnostic Médical et Intelligence Artificielle : Enjeux Éthiques. Avis commun du CCNE et du CNPEN, Avis 141 du CCNE, Avis 4 du CNPEN. Novembre 2022.

TABLE DES MATIERES

TABLE DES MATIERES	3
RESUME	5
INTRODUCTION	7
I. SYSTEMES D'INTELLIGENCE ARTIFICIELLE ET AIDE AU DIAGNOSTIC : DE QUOI PARLE-T-ON ?	11
1.1 Le diagnostic médical assisté par des systèmes d'intelligence artificielle	11
1.2 Systèmes d'intelligence artificielle : quelles définitions actuelles ?	12
1.3 Contextualisation et angle d'approche de l'avis	13
II. EXEMPLES DE SYSTEMES D'INTELLIGENCE ARTIFICIELLE D'AIDE AU DIAGNOSTIC MEDICAL : BENEFICES ET ENJEUX ETHIQUES	19
2.1 Les systèmes apprenants d'analyse d'images	19
2.1.1 L'oncologie	20
2.1.2 Les maladies rétiniennes	21
2.1.3 La dermatologie	22
2.1.4 La cardiologie	22
2.1.5 L'analyse d'images en histopathologie	23
2.1.6 Systèmes d'intelligence artificielle apprenants, incertitudes et performances ...	23
2.1.7 L'interprétation de données génomiques	26
2.1.8 Implications éthiques : incomplétude et incidentalome	26
2.2 Les SIADM dans l'amélioration du parcours de soins	27
2.2.1 L'orientation pré-diagnostique	27
2.2.2 Aide à la gestion de la prise en charge de l'urgence vitale	28
2.2.3 Aide au suivi des patients à domicile	29
2.2.4 Implications éthiques : bénéfices économiques versus bénéfices de santé ...	30
2.3 Dispositifs portables utilisant des SIADM hors du parcours de soins	33
2.3.1 Implications éthiques : risques d'avis non professionnels	34
III. UN ENCADREMENT DES SIADM QUI RESTE A CONSTRUIRE	36
3.1 Pour une évaluation des SIADM qui protège la relation de soin	36
3.1.1 Le rôle de la Haute Autorité de Santé (HAS) dans l'évaluation des SIADM	36
3.1.2 Implications éthiques : pour une meilleure explicabilité et une tempérance des promesses	37
3.2 La certification, étape cruciale pour des SIADM de confiance	39
3.2.1 Une évolution de la réglementation qui peine à suivre le rythme de diffusion des SIADM	39
3.2.2 Une procédure de conformité matérialisée par le marquage « CE »	40
3.2.3 Implications éthiques : un contrôle de conformité en cours d'élaboration	42

3.3	Régime de responsabilité	42
3.4	Pour un contrôle de conformité et une évaluation des SIADM plaçant la qualité du soin avant toute autre considération	43
VI.	ENJEUX ETHIQUES POSES PAR LES SIADM DANS LA TRANSFORMATION DE LA DEMARCHE DIAGNOSTIQUE	
	47	
4.1	Enjeux de la radiomique	47
4.2	Enjeux dans l'interprétation des données génomiques	47
4.2.1	Du diagnostic du symptôme au diagnostic du risque.....	48
4.2.2	Implications éthiques : explicabilité, explication et importance du conseil en génétique	49
4.3	Enjeux concernant l'émergence de nouveaux espaces de démocratie en santé...	50
4.3.1	L'émergence d'un nouveau statut : l'auxiliaire numérique.....	50
4.3.2	Implications éthiques : promouvoir les initiatives de démocratie en santé.....	51
4.4	Les SIADM ne doivent pas dégrader les qualités du cadre diagnostique	52
	CONCLUSION	54
	ANNEXES	55
	Annexe 1 : Membres du groupe de travail.....	55
	Annexe 2 : Liste des personnes auditionnées	56

RÉSUMÉ

Cet avis fait suite à une saisine du Comité national pilote d'éthique du numérique (CNPEN) par le Premier ministre sur les enjeux d'éthique de l'utilisation de l'intelligence artificielle (IA) dans le champ du diagnostic médical, exprimée dans sa lettre de mission du 15 juillet 2019. Eu égard à la dimension transversale de cette thématique relevant tant de la bioéthique que de l'éthique du numérique, le Comité consultatif national d'éthique pour les sciences de la vie et de la santé (CCNE) et le CNPEN ont conduit ensemble la réflexion ayant aboutie à cet avis commun.

Les technologies développées autour de l'IA concernent de nombreux domaines d'application (médecine, transports, cybersécurité, commerce, industrie, etc.) et leur irruption dans nos usages quotidiens s'accélère à un rythme soutenu. Ce contexte encourage de nombreux États et institutions à considérer les enjeux d'éthiques qui accompagnent cette transformation.

Le secteur de la santé et de la médecine apparaît particulièrement concerné par le développement de systèmes d'intelligence artificielle. Leur application au champ médical induit une transformation de la relation entre médecins et patients et ouvre de nombreuses questions sur l'avenir des systèmes de santé. Conformément à la saisine ministérielle, cet avis se concentre sur les enjeux d'éthique des systèmes d'intelligence artificielle appliqués au diagnostic médical (SIADM) visant l'amélioration des performances techniques des praticiens.

Le CCNE et le CNPEN ont construit leur réflexion en déterminant, dans un premier temps, ce à quoi l'utilisation de systèmes d'intelligence artificielle appliqués au diagnostic médical fait référence en dégagant les principaux soubassements de l'application de ces nouvelles techniques (partie 1) et leur portée opérationnelle (partie 2). L'avis questionne ensuite le processus de régulation des SIADM et ses caractéristiques actuelles (partie 3). Enfin, une analyse des impacts des SIADM sur la démarche diagnostique décrit les questions en cours et les évolutions qu'elles impliquent (partie 4).

Au fil des seize recommandations et des sept points de vigilance identifiés par le CCNE et le CNPEN, plusieurs axes de tensions éthiques apparaissent. Il nous a d'abord paru important de dresser un panorama de ce que les SIADM sont effectivement capables de faire aujourd'hui. Les promesses sont nombreuses, et on peine parfois à les distinguer des faits. Cette distinction nous a semblé être une première tâche d'éthique. Ensuite, nous rappelons que les SIADM produisent des résultats basés d'une part sur des approches probabilistes et d'autre part qu'ils peuvent être entachés d'erreurs. Nous insistons sur le fait que **les équipes soignantes et les patients ne doivent pas se priver des avantages apportés par ces outils, tout en se donnant constamment les moyens de prendre de la distance avec le résultat fourni**. Nous présentons les principales ressources qui permettent cette prise de recul. Tout SIADM doit être soumis à un contrôle humain. Ses résultats doivent être explicables. Le contrôle de conformité du SIADM, qui assure qu'il n'est pas dangereux et ainsi autorise sa mise sur le marché, doit être amélioré, et surtout doit à l'avenir être accompagné d'une évaluation de son efficacité clinique montrant non seulement son absence de nocivité mais aussi qu'il contribue efficacement au principe de bienfaisance.

Les systèmes d'intelligence artificielle appliqués au diagnostic médical doivent donc toujours être utilisés **en priorité dans une optique d'amélioration du soin**, avant les intérêts organisationnels, économiques, ou managériaux.

INTRODUCTION

Dans sa lettre du 15 juillet 2019, le Premier ministre a souhaité que le Comité national pilote d'éthique du numérique (CNPEN) émette un avis sur l'utilisation de systèmes d'intelligence artificielle dans le diagnostic médical. Eu égard à la dimension transversale de cette thématique relevant tant de la bioéthique que de l'éthique du numérique, le Comité consultatif national d'éthique pour les sciences de la vie et de la santé (CCNE) et le CNPEN ont conduit une réflexion commune pour formuler cet avis.

La notion d'intelligence artificielle (IA) recouvre une pluralité de techniques reposant sur des approches distinctes et poursuivant des objectifs différents. Les débats autour de la définition même de l'IA révèlent les enjeux qu'elle mobilise, notamment sur les plans juridiques et industriels¹. La définition d'un système d'intelligence artificielle (SIA) retenue dans la rédaction de cet avis correspond au cadre proposé par la Commission européenne, comme nous le verrons en première partie.

Les technologies développées autour de l'IA concernent de nombreux domaines d'application (médecine, transports, cybersécurité, commerce, industrie etc.) et leur irruption dans nos usages quotidiens s'accélère à un rythme soutenu. Des pans entiers de l'activité humaine sont en passe d'être transformés par l'essor des technologies d'intelligence artificielle ouvrant des horizons qui semblaient jusqu'ici inatteignables. Ces bouleversements poussent les États à considérer l'IA comme un nouvel outil stratégique alors même qu'aucun cadre juridique adapté ne permet de répondre aux enjeux complexes de gouvernance soulevés par les technologies d'intelligence artificielle. Ce contexte encourage de nombreux États et institutions à proposer une meilleure régulation en la matière et à considérer les questions éthiques qui accompagnent cette transformation.

En 2018, lors de sa présentation du programme *AI for Humanity*² déclinant la stratégie de la France en matière d'intelligence artificielle, le Président de la République, Emmanuel Macron, annonçait ainsi vouloir créer « un environnement réglementaire et financier favorable à l'émergence de champions de l'IA »³. À cette occasion, il désignait l'IA comme un outil stratégique de développement économique et industriel national et européen appelé à se démarquer face aux grandes puissances et en particulier les États-Unis et la Chine⁴. Parmi les engagements du Gouvernement français apparaît la mise en place d'une « réflexion sur les enjeux éthiques liés au développement des technologies d'IA » indiquant notamment l'**explicabilité** des technologies d'IA comme l'un des facteurs de leur acceptabilité sociale.

¹ Il n'existe pas aujourd'hui de définition communément admise de l'IA, et de nombreuses institutions ont développé leur propre vision. Voir : Samoili S. et al., (2021), *AI watch defining artificial intelligence 2.0: towards an operational definition and taxonomy for the AI landscape*, JRC technical report, 125 p.

² Voir le site dédié : <https://www.aiforhumanity.fr/>.

³ « Intelligence artificielle : faire de la France un champion », 29 mars 2018, voir le site du gouvernement français : <https://www.gouvernement.fr/argumentaire/intelligence-artificielle-faire-de-la-france-un-champion>.

⁴ *Ibid*, « Cette stratégie s'inscrit aussi dans l'Europe, l'Union européenne ayant vocation à devenir elle-même un champion mondial de l'IA et à porter en ce domaine une voix singulière, distincte de celle des États-Unis ou de la Chine. »

La volonté de développer les conditions d'une réflexion éthique sur l'intelligence artificielle prend place dans un contexte international de multiplication des débats sur le sujet. Ainsi, à ce jour, près d'une centaine d'institutions ont formulé des recommandations et lignes directrices pour nourrir le débat public et guider les décideurs politiques en matière de gouvernance de l'IA. On retrouve parmi les positions les plus marquantes les « 23 principes d'Asilomar », signés en 2017 par des centaines de scientifiques spécialisés dans l'IA et la robotique, et faisant écho à la Conférence d'Asilomar organisée en 1975 par Paul Berg appelant à un moratoire sur les manipulations génétiques⁵.

Le secteur de la santé et de la médecine est particulièrement concerné par le développement de systèmes d'intelligence artificielle. Leur application au champ médical suggère une transformation de la relation entre médecins et patients et ouvre de nombreuses questions sur l'avenir des systèmes de santé. Les SIA en santé s'accompagnent en effet d'enjeux hétérogènes voire parfois contradictoires, entre intérêts économiques et industriels (conception des outils de SIA), promesses thérapeutiques pour les patients (bénéfices sanitaires attendus, dépistage, prévention et traitement), modification de la pratique professionnelle des personnels de santé concernés (activité des soignants pouvant être amenés à les utiliser, aide au diagnostic, éventuellement même des suppressions ou modifications de postes) et objectifs de régulation pour les autorités publiques. Ce champ nécessite une attention particulière du fait des valeurs qu'il mobilise, au premier rang desquelles le respect de la vie humaine et la garantie de conditions équitables d'accès aux soins.

L'une des principales tensions soulevées par la mise en œuvre des technologies d'IA en santé résulte du déséquilibre entre les connaissances, le débat public et la rapidité potentielle de la transformation des pratiques induites par les techniques innovantes reposant sur l'IA. L'écart est souvent important entre les promesses suscitées par les technologies d'IA, les représentations que le public se fait de l'intelligence artificielle en santé et la réalité de ce que *peuvent* et *ne peuvent pas* faire les SIA appliqués à la médecine. L'échec du programme d'IA Watson Health, dont la société IBM annonçait lors de sa création en 2015 la promesse d'une aide au diagnostic inédite grâce à l'intelligence artificielle, invite ainsi à accueillir avec précaution les annonces agitant le spectre d'une substitution de l'homme, et en l'occurrence des professionnels médicaux, par la machine⁶. À l'inverse, certains succès, particulièrement en imagerie médicale que nous développons plus bas, montrent que ces techniques peuvent être utiles et expliquent l'engouement qu'elles peuvent inspirer.

Cette temporalité propre aux SIA a deux conséquences importantes qui s'inscrivent dans un contexte et un état de la science particuliers. Étant donnée la rapidité des transformations observées dans le champ qui nous occupe ici, cette contextualisation qui va d'habitude sans le dire, mérite d'être rappelée. D'autre part, le très grand nombre de

⁵ Voir par exemple : Orsini A., *Les « 23 principes d'Asilomar » veulent encadrer le développement de l'intelligence artificielle*, Numerama.com, 01/02/2017. On peut aussi mentionner la recommandation de l'UNESCO sur l'éthique de l'IA https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre et le texte du *High-Level Expert Group on AI* de l'UE <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

⁶ Cette division de la société informatique IBM, issue du système d'intelligence artificielle Watson ayant remporté le jeu télévisé américain *Jeopardy* en 2011, promettait notamment de pouvoir établir, à partir de l'analyse de grandes masses de données de santé, des diagnostics à propos de certains cancers de manière plus fiable que les médecins. Le programme n'a finalement pas atteint les ambitieux objectifs annoncés et sera finalement vendu en 2022 à une société de capital-investissement.

promesses plus ou moins réalistes qui marquent le champ de l'IA (sur lesquelles nous reviendrons) nous oblige, ici encore plus qu'ailleurs, à faire un état des lieux de la science précis, documenté et à froid. Celui-ci fait partie intégrante de la réflexion éthique dans la mesure où ce qu'est et ce que fait réellement l'IA aujourd'hui est loin d'être clair aux yeux de nombreux acteurs. Nous avons donc décidé de donner une place particulièrement importante à l'établissement des faits qui est en tant que telle une clarification éthique.

D'autre part, l'utilisation de l'IA en santé est un domaine si vaste qu'il serait périlleux de tenter de l'embrasser entièrement. Conformément à la saisine ministérielle, nous avons restreint notre analyse au cas des systèmes d'intelligence artificielle appliqués au diagnostic médical (SIADM) visant l'amélioration des performances techniques des praticiens. Autrement dit, nous n'avons pas abordé spécialement la question des pronostics, de la prévention ou de l'établissement des traitements, sauf s'ils entrent dans la construction d'un diagnostic.

Les SIADM à eux seuls amènent de nombreux enjeux éthiques. Dès les premières étapes de son développement (collecte de données, apprentissage à partir de ces données, construction puis évaluation de l'algorithme), la conception d'un SIADM implique qu'une vigilance soit portée aux finalités du système et à l'utilisation vers laquelle il guidera le médecin. Le développement de la notion d'**éthique by design** faisant écho aux tensions éthiques mobilisées dès la conception d'un système reposant sur des technologies numériques est tout à fait adaptée aux SIADM, dont le champ d'application médical implique des sujets particulièrement sensibles. En conjuguant une éthique de l'intention centrée sur le sujet, une éthique par conception centrée sur l'objet, et une éthique de la médiation centrée sur la relation sujet-objet⁷, l'approche éthique de la conception d'un SIADM pose une pluralité de questions sur la chaîne de responsabilités (du concepteur, de l'évaluateur, du certificateur, de l'utilisateur final et plus globalement du décideur politique). L'*ethics by design*, par l'idée de responsabilisation des concepteurs de systèmes qu'elle développe, s'étend à la finalité d'un système impliquant par exemple le fait de favoriser des algorithmes privilégiant l'intérêt du plus grand nombre face aux intérêts minoritaires, ou encore la stricte **protection des données** à caractère personnel.

Les effets d'un SIADM, tant pour le praticien qui va être amené à l'utiliser, que pour le patient qui va en bénéficier, et la répercussion de l'utilisation de ces nouvelles technologies sur l'ensemble de la société, doivent être interrogés à l'aune des méthodes qui visent à garantir la relation de **confiance** entre patients et médecins⁸ et un **accès équitable** à des soins de qualité pour tous. De plus, il convient de développer des méthodes pour garantir l'explicabilité, qui reste souvent à préciser, des résultats fournis par un SIADM à chaque étape de sa diffusion.

L'approche des enjeux éthiques introduits par le développement de SIADM doit ainsi prendre en compte la pluralité des acteurs impliqués, les différentes étapes du processus de diffusion de ces systèmes et les implications sociales qu'ils induisent. D'autre part, notre réflexion se nourrit des travaux de l'Union Européenne sur la question, qui eux-mêmes peuvent être influencés par les comités d'éthique nationaux.

⁷ Voir : Fischer F., (2019), « L'éthique by design du numérique : généalogie d'un concept », *Sciences du Design*, 2019/2 (n° 10), p. 61-67.

⁸ B. Mittelstadt, L'impact de l'intelligence artificielle sur les relations patients médecins, Rapport remis au Conseil de l'Europe. 2022. <https://rm.coe.int/inf-2022-5-report-impact-of-ai-on-doctor-patient-relations-f/1680a6885a>

Le CCNE et le CNPEN ont construit leur réflexion en déterminant, dans un premier temps, ce à quoi l'utilisation de systèmes d'intelligence artificielle appliqués au diagnostic médical fait référence en dégagant les principaux soubassements de l'application de ces nouvelles techniques (partie 1) et leur portée opérationnelle (partie 2). L'avis questionne ensuite le processus de régulation des SIADM et ses caractéristiques actuelles (partie 3). Enfin, une analyse des impacts des SIADM sur la démarche diagnostique décrit les questions en cours et les évolutions qu'elles impliquent (partie 4).

Cet avis décline au fil de son cheminement des **implications éthiques** motivant des **recommandations** (16) et préconisations pour un usage des SIADM respectueux des patients, de la relation entre médecins et usagers du système de soins et de notre société démocratique. Ces réflexions sont enrichies de **points de vigilance** (7) destinés à alerter sur les principales tensions éthiques qui peuvent apparaître au cours de la diffusion d'un SIADM.

I. SYSTEMES D'INTELLIGENCE ARTIFICIELLE ET AIDE AU DIAGNOSTIC : DE QUOI PARLE-T-ON ?

1.1 Le diagnostic médical assisté par des systèmes d'intelligence artificielle

Le **diagnostic médical** tel que défini par l'Académie nationale de médecine « consiste à reconnaître les maladies par leurs symptômes et leurs signes, et à les distinguer les unes des autres. L'élaboration du diagnostic a comme premier objectif la prise en charge appropriée du malade. Il est un élément essentiel de la décision médicale »⁹.

L'élaboration d'un diagnostic est un processus qui requiert un ensemble d'étapes, chacune apportant des informations (cliniques, biologiques, radiologiques, etc.), aboutissant à une appréciation de l'état de santé d'un patient. La précision de l'identification et de la caractérisation d'une pathologie ou d'un état donné est un élément décisif de l'approche diagnostique puisqu'elle détermine la démarche thérapeutique et implique donc des conséquences cliniques.

Comme le souligne l'Académie nationale de médecine, si de nouvelles technologies et méthodes sont régulièrement proposées de manière à affiner le diagnostic final et à le rendre plus sûr et plus précis, « aucun des procédés nouveaux ne doit mettre en cause la primauté de la clinique. La diffusion d'une nouvelle technique diagnostique ne doit jamais être préconisée sans qu'aient été bien vérifiées la reproductibilité, la valeur diagnostique comparée aux moyens déjà existants, la sécurité du patient et les conséquences économiques »¹⁰.

Selon Bonnet et Atlan¹¹, dans le cadre d'un examen médical, deux modes de pensée se conjuguent pour élaborer un diagnostic. D'une part, une pensée que l'on peut qualifier de « **globale** », qui repose sur l'art médical, l'expertise scientifique (notamment l'*Evidence-Based Medicine*), l'expérience, l'intuition, la prudence, la discussion collégiale et l'observation des spécificités d'une situation. D'autre part, une pensée que l'on qualifiera ici d'**algorithmique**, reposant sur une méthode « pas à pas » ou probabiliste. Cette pensée passe néanmoins sous silence les spécificités contextuelles du cas examiné, que seule la prudence, au sens aristotélicien du mot, permet d'identifier.

Les premiers systèmes d'intelligence artificielle appliqués au diagnostic médical ont été développés dans les années 1980. Leur fonctionnement par modélisation systématique et aboutissant à une conclusion à l'issue d'une série d'étapes d'inférence logique, ou grâce à un apprentissage utilisant des réseaux de neurones artificiels, leur permet alors de reproduire une partie de ce raisonnement algorithmique humain.

Les nouvelles technologies d'IA renouvellent profondément les potentiels de cette pensée « pas à pas ». Appliquer cette pensée au service du diagnostic final implique donc de

⁹ Académie nationale de médecine, « Le diagnostic en médecine : histoire, mise en œuvre présente, perspectives », Séance du 20 juin 2006 [En ligne : <https://www.academie-medecine.fr/06-12-le-diagnostic-en-medecine-histoire-mise-en-uvre-presente-perspectives/>]

¹⁰Académie nationale de médecine, *Ibid.*

¹¹ Bonnet J.-L., Atlan G., (1980), « L'informatique médicale », In *Médecine Pathologie générale, 1 : Présentation des sciences de base*, édité par Pierre de Graciansky et Henri Péquignot, 578-600. Encyclopédie de la Pléiade 45. Paris: Gallimard.

parvenir à trouver de nouvelles conjugaisons, positives et productives, et non destructrices, entre ces deux modes de raisonnement (pensée dite « globale » et pensée algorithmique) qui s'articulent tout au long de la démarche diagnostique. **Un grand nombre de tensions éthiques surviennent ainsi non pas des limites de ces deux modes de raisonnement mais des difficultés rencontrées pour les articuler pertinemment via un système d'IA.**

Cette tentative de modélisation de la pensée algorithmique peut en retour permettre de mieux comprendre quelles sont les spécificités et la puissance propre de la pensée globale. Se fier exagérément à l'IA dans la décision finale induit un risque de manquer les spécificités propres à chaque cas médical ; à l'inverse, se passer de l'IA revient à un refus d'exploiter une source d'information supplémentaire, peut-être décisive, offerte par les nouvelles technologies. Mais quelle que soit la manière dont cela s'articule, l'utilisation des technologies d'IA peut et doit enrichir, et non appauvrir, la manière dont les soignants vont utiliser leurs compétences professionnelles dans la démarche diagnostique.

1.2 Systèmes d'intelligence artificielle : quelles définitions actuelles ?

L'expression « intelligence artificielle », omniprésente dans le débat public, est porteuse de représentations, de promesses et d'inquiétudes, en particulier dans le domaine de la santé. Elle est couramment employée pour désigner les méthodes d'apprentissage machine (*machine learning*), c'est-à-dire des approches basées sur des algorithmes qui, se fondant sur des données, construisent des modèles (arbres de décisions, réseaux de neurones, ou autres) qui seront ensuite appliqués à de nouvelles données pour calculer une décision à partir de celles-ci. *Stricto sensu*, l'apprentissage machine constitue un sous-ensemble de méthodes parmi le vaste champ d'étude que recouvre l'IA et plus généralement le numérique¹². Ces systèmes, qui ont connu un développement très rapide ces dernières années, notamment grâce à une capacité accrue de collecte, de stockage et d'analyse de masses de données considérables, peuvent être conçus de manière à évoluer dans le temps, en fonction des données qui leur sont fournies lors de leurs usages.

La polysémie de la notion d'intelligence artificielle, due en partie à l'évolution constante des techniques algorithmiques, en complique les tentatives de définition, particulièrement en matière juridique.

En France, le texte de référence en matière de SIADM est l'article 17 de la **loi n° 2021-1017 du 2 août 2021 relative à la bioéthique**¹³. Cette dernière introduit ainsi un cadre légal à l'utilisation de dispositifs médicaux embarquant un système particulier d'intelligence artificielle en insérant dans la quatrième partie du code de la santé publique (consacrée aux « Professions de santé ») l'article L. 4001-3 visant l'utilisation, par un professionnel de santé et dans le cadre d'« un acte de prévention, de diagnostic ou de soin » d'un « dispositif médical comportant un traitement de données algorithmique dont l'apprentissage a été réalisé à partir de données massives ». L'article mentionne également des points cruciaux tels que l'**information** des professionnels de santé quant au recours au traitement de données, l'**accessibilité** des données pour les patients et les **résultats** qui en sont issus ou encore l'**assurance de l'explicabilité**¹⁴, par les concepteurs, du fonctionnement du traitement algorithmique aux utilisateurs.

¹² CNPEN, *Manifeste pour une éthique du numérique*, Paris, avril 2021, 5 p.

¹³ Consultable au lien suivant : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000043884384/>

¹⁴ Ce terme sera explicité en 3.1.2.

À l'échelle communautaire, le projet de réglementation commune sur l'IA (ou IA Act) publié le 21 avril 2021 par la Commission européenne¹⁵, mentionné dans l'introduction, évoque la nécessité d'une « définition claire des applications d'IA à haut risque » face « aux nouvelles possibilités qu'offre l'intelligence artificielle et aux nouveaux défis qui en découlent » et propose une définition plus précise de ce qu'est un « système d'intelligence artificielle » appliqué au diagnostic médical (SIADM). Dans son article 3, la Commission définit ainsi un SIADM comme un logiciel « qui peut, pour un ensemble donné d'objectifs définis par l'homme, générer des résultats tels que des contenus, des prédictions, des recommandations ou des décisions influençant les environnements avec lesquels il interagit »¹⁶ et relève des techniques et approches suivantes :

- (a) Approches d'apprentissage automatique, y compris d'apprentissage supervisé, non supervisé et par renforcement, utilisant une grande variété de méthodes, y compris l'apprentissage profond.
- (b) Approches fondées sur la logique et les connaissances, y compris la représentation des connaissances, la programmation inductive (logique), les bases de connaissances, les moteurs d'inférence et de déduction, le raisonnement (symbolique) et les systèmes experts.
- (c) Approches statistiques, estimation bayésienne, méthodes de recherche et d'optimisation.

L'expression « systèmes d'intelligence artificielle » appliqués au diagnostic médical, ou SIADM, sera ici employée au sens déterminé par cette version du règlement européen.

1.3 Contextualisation et angle d'approche de l'avis

Des contributions antérieures, sur lesquelles le CCNE¹⁷ et le CNPEN¹⁸ s'appuient dans ce travail, ont notamment permis de nourrir la réflexion sur la construction d'un cadre relatif à l'utilisation de systèmes d'IA dans les domaines du soin. Le concept de « garantie humaine » qui y est développé a ainsi été récemment repris par l'Organisation mondiale de la santé (OMS)¹⁹. La terminologie de « garantie humaine » peut aussi se trouver précisée dans la notion de « **contrôle humain** », au sens de vigilance humaine. Cette notion a été privilégiée dans la rédaction de cet avis car elle correspond aux avancées du droit européen. Le projet de réglementation de la Commission européenne sur l'intelligence artificielle (IA Act mentionné plus haut), dans son article 14, reconnaît en effet le principe de *human oversight* qui doit être assuré sur les systèmes d'IA utilisés pour établir un

¹⁵ Consultable au lien suivant : https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0020.02/DOC_1&format=PDF

¹⁶ Commission européenne, « Proposition de règlement du parlement européen et du conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'union », Bruxelles, 21 avril 2021 [consultable au lien suivant : <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>].

¹⁷ (1) Chapitre « Numérique et santé » de l'avis n° 129 du CCNE *Contribution du Comité consultatif national d'éthique à la révision de la loi de bioéthique 2018-2019*, 18 septembre 2018, pp. 94-106.

(2) Rapport CERNA/CCNE, *Numérique et santé : quels enjeux éthiques pour quelles régulations ?*, 19 novembre 2018, 102 p.

(3) Avis n° 130 du CCNE du 29 mai 2019, *Données massives et santé : État des lieux, prospective et nouvelles questions éthiques*, 94 p.

¹⁸ CNPEN, Bulletin de veille n° 3 du 21 juillet 2020, *Enjeux d'éthique liés aux outils numériques en télémédecine et télésoin dans le contexte de la Covid-19*, 21 p.

¹⁹ Organisation mondiale de la santé (OMS), *Éthique et gouvernance de l'intelligence artificielle en santé*, 28 juin 2021, p. 13.

diagnostic médical, principe que la traduction française par la Commission désigne par l'expression « contrôle humain ».

Version anglaise de l'Article 14 du projet d'IA Act

Article 14
Human oversight

1. High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use.

2. Human oversight shall aim at preventing or minimising the risks to health, safety or fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, in particular when such risks persist notwithstanding the application of other requirements set out in this Chapter.

3. Human oversight shall be ensured through either one or all of the following measures:

(a) identified and built, when technically feasible, into the high-risk AI system by the provider before it is placed on the market or put into service;

(b) identified by the provider before placing the high-risk AI system on the market or putting it into service and that are appropriate to be implemented by the user.

4. The measures referred to in paragraph 3 shall enable the individuals to whom human oversight is assigned to do the following, as appropriate to the circumstances:

(a) fully understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation, so that signs of anomalies, dysfunctions and unexpected performance can be detected and addressed as soon as possible;

(b) remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias'), in particular for high-risk AI systems used to provide information or recommendations for decisions to be taken by natural persons;

(c) be able to correctly interpret the high-risk AI system's output, taking into account in particular the characteristics of the system and the interpretation tools and methods available;

(d) be able to decide, in any particular situation, not to use the high-risk AI system or otherwise disregard, override or reverse the output of the high-risk AI system;

(e) be able to intervene on the operation of the high-risk AI system or interrupt the system through a "stop" button or a similar procedure.

5. For high-risk AI systems referred to in point 1(a) of Annex III, the measures referred to in paragraph 3 shall be such as to ensure that, in addition, no action or decision is taken by the user on the basis of the identification resulting from the system unless this has been verified and confirmed by at least two natural persons.

Version française de L'article 14 du projet d'IA Act

Article 14

Contrôle humain

1. La conception et le développement des systèmes d'IA à haut risque permettent, notamment au moyen d'interfaces homme-machine appropriées, un contrôle effectif par des personnes physiques pendant la période d'utilisation du système d'IA.
2. Le contrôle humain vise à prévenir ou à réduire au minimum les risques pour la santé, la sécurité ou les droits fondamentaux qui peuvent apparaître lorsqu'un système d'IA à haut risque est utilisé conformément à sa destination ou dans des conditions de mauvaise utilisation raisonnablement prévisible, en particulier lorsque de tels risques persistent nonobstant l'application d'autres exigences énoncées dans le présent chapitre.
3. Le contrôle humain est assuré au moyen d'une ou de la totalité des mesures suivantes :
 - (a) lorsque cela est techniquement possible, des mesures identifiées et intégrées par le fournisseur dans le système d'IA à haut risque avant la mise sur le marché ou la mise en service de ce dernier ;
 - (b) des mesures identifiées par le fournisseur avant la mise sur le marché ou la mise en service du système d'IA à haut risque et qui se prêtent à une mise en œuvre par l'utilisateur.
4. Les mesures prévues au paragraphe 3 donnent aux personnes chargées d'effectuer un contrôle humain, en fonction des circonstances, la possibilité :
 - (a) d'appréhender totalement les capacités et les limites du système d'IA à haut risque et d'être en mesure de surveiller correctement son fonctionnement, afin de pouvoir détecter et traiter dès que possible les signes d'anomalies, de dysfonctionnements et de performances inattendues ;
 - (b) d'avoir conscience d'une éventuelle tendance à se fier automatiquement ou excessivement aux résultats produits par un système d'IA à haut risque (« biais d'automatisation »), en particulier pour les systèmes d'IA à haut risque utilisés pour fournir des informations ou des recommandations concernant les décisions à prendre par des personnes physiques ;
 - (c) d'être en mesure d'interpréter correctement les résultats du système d'IA à haut risque, compte tenu notamment des caractéristiques du système et des outils et méthodes d'interprétation disponibles ;
 - (d) d'être en mesure de décider, dans une situation particulière, de ne pas utiliser le système d'IA à haut risque ou de négliger, passer outre ou inverser le résultat fourni par ce système ;
 - (e) d'être capable d'intervenir sur le fonctionnement du système d'IA à haut risque ou d'interrompre ce fonctionnement au moyen d'un bouton d'arrêt ou d'une procédure similaire.

5. Pour les systèmes d'IA à haut risque visés à l'annexe III, point 1 a), les mesures prévues au paragraphe 3 sont de nature à garantir qu'en outre, aucune mesure ou décision n'est prise par l'utilisateur sur la base de l'identification résultant du système sans vérification et confirmation par au moins deux personnes physiques.

De surcroît, notons que le CNPEN a relevé que les méthodologies dites de « collège de garantie humaine » associant professionnels de santé et représentants des patients permettraient d'assurer ce contrôle humain au stade de l'application de l'IA²⁰.

Le présent avis s'appuie plus globalement sur des textes fondateurs dans le domaine de l'éthique biomédicale comme le code de Nuremberg, la déclaration d'Helsinki, le rapport Belmont ou la Déclaration universelle sur la bioéthique et les droits de l'homme. Cette longue tradition biomédicale de la réflexion éthique a inspiré et parfois même influencé la manière dont s'est construite, beaucoup plus récemment, la réflexion éthique centrée sur le numérique. Ces deux champs de l'éthique ont le souci commun, racine de la démarche éthique, de préserver la dignité humaine. Cet avis sur les enjeux d'éthique liés à l'apparition et la diffusion de systèmes d'intelligence artificielle appliqués au diagnostic médical invite à approfondir les questionnements autour de l'articulation de ces deux domaines de la réflexion éthique. Ce texte veille à ce que les cadres éthiques utilisés en matière de bioéthique et d'éthique du numérique soient applicables à l'intelligence artificielle.

Les SIADM peuvent être utilisés à plusieurs étapes du parcours de soin d'un patient :

- lors de l'orientation initiale des patients, **en amont de la consultation médicale**, dans une optique de priorisation, notamment en permettant d'identifier les situations requérant une prise en charge urgente ; il peut s'agir de dispositifs proposés aux patients eux-mêmes, ou utilisés dans les services d'urgence par les professionnels de santé ²¹ ;
- lors des **étapes diagnostiques**, et en particulier de celles ayant recours à un examen d'imagerie ;
- lors du **suivi médical**, si des dispositifs de surveillance permettant d'ajuster un traitement ou de détecter une modification de l'état du patient sont utilisés.

Notons que l'intelligence artificielle est maintenant utilisée aussi dans le cadre de campagnes de dépistage en population générale, par exemple pour le cancer du sein. En deux ans, les publications sur le sujet se sont multipliées avec une dynamique exponentielle et un élargissement de l'offre à la disposition des radiologues²². Il en est de même pour le développement de ces techniques dans le cadre de la biologie médicale²³. Néanmoins, nous réservons la réflexion éthique portant sur ces usages spécifiques de l'IA

²⁰ Avis émis en juillet 2020 par le CNPEN en réponse au Livre Blanc de la Commission européenne sur l'IA (Commission européenne, (2020), « Livre blanc. Intelligence artificielle. Une approche européenne axée sur l'excellence et la confiance », 31 p.).

²¹ D'autres enjeux du parcours de soins pourraient mobiliser des préconisations du CCNE, mais cet avis se concentre sur les enjeux des SIADM dans le cadre diagnostique.

²² Thomassin-Naggara I., Ceugnart, L., Tardivon L., Verzaux L., Balleyguier C. et al., « Intelligence artificielle : Place dans le dépistage du cancer du sein en France », *Bulletin du Cancer* 109, n° 7 (1 juillet 2022) : 780-85. <https://doi.org/10.1016/j.bulcan.2022.04.008>.

²³ <https://lapbm.org/intelligence-artificielle-et-biologie-medicale/>.

à des travaux ultérieurs en choisissant de se concentrer, dans cet avis, sur la relation clinique entre le patient et son médecin.

D'autre part, il faut distinguer les SIADM utilisés dans le cadre d'un parcours de soin, et qui ont été soumis à une certification médicale, des multiples autres applications. Ces dernières, qui se situent à la frontière entre le bien-être, l'utilisation récréative, ou la surveillance de paramètres de santé, bénéficient d'un marquage CE (« Conformité Européenne ») mais ne suivent pas le parcours de certification des dispositifs médicaux.

La plupart des approches présentées au cours du texte sont encore au stade de l'expérimentation. Mais si leur transfert dans la pratique clinique n'est pas encore une réalité, il pourrait le devenir dans un avenir proche, aussi il convient de s'interroger sur les questions éthiques que l'utilisation des SIADM présentés peut susciter²⁴. Nous discutons notamment des principaux freins à leur implémentation, qu'ils soient d'ordre technique ou réglementaire, en développant particulièrement les tensions et enjeux éthiques rencontrés.

²⁴ Keane P.A., Topol E., (2018), "With an eye to AI and autonomous diagnosis", *npj Digital Medicine*,1, 40. ; He J., et al., (2019), "The practical implementation of artificial intelligence technologies in medicine", *Nature Medicine*, 25:3.0-36.

II. EXEMPLES DE SYSTEMES D'INTELLIGENCE ARTIFICIELLE D'AIDE AU DIAGNOSTIC MEDICAL : BENEFICES ET ENJEUX ETHIQUES

La reconnaissance d'images est particulièrement adaptée aux systèmes d'intelligence artificielle fondés sur l'apprentissage machine. Ces systèmes représentent l'évolution logique, dans le domaine de l'imagerie, des logiciels d'aide à la détection assistée par ordinateur utilisés par les médecins et particulièrement par les radiologues depuis une dizaine d'années. Le rôle de ces outils d'alerte – dont les règles de détection sont prédéterminées par les humains d'après leurs connaissances et leur expérience – est d'améliorer les performances diagnostiques et d'attirer l'attention du radiologue sur de possibles anomalies. Leur utilité clinique ne fait cependant pas consensus, comme nous allons le voir.

Une amélioration importante est apportée par l'apprentissage machine sur une quantité considérable de données radiologiques numérisées. L'entraînement des algorithmes peut se faire de deux façons :

- dans le cas de l'apprentissage supervisé, l'algorithme de **reconnaissance d'images utilise des données annotées par une personne**, c'est-à-dire qu'à la donnée brute est associée une information *a priori* sur le contenu de l'image ;
- au contraire certaines approches non supervisées ne nécessitent pas d'étiquetage préalable des caractéristiques d'images, et les critères pertinents pour parvenir à un résultat **sont déterminés par l'algorithme lui-même sans intervention humaine après sa conception**.

Néanmoins, cela accentue le **risque d'opacité du processus d'obtention du résultat**, ce que nous discutons plus loin.

2.1 Les systèmes apprenants d'analyse d'images

L'utilisation de logiciels d'aide à la détection assistée par ordinateur dans le domaine de l'imagerie a pris de l'ampleur il y a une dizaine d'années à la faveur de l'avènement d'approches d'apprentissage profond (*deep learning*)²⁵. Ces dernières se prêtent particulièrement à la classification et à la segmentation d'images issues d'exams radiologiques tels que les scanners (tomodensitométries ou TDM), CT scans (*computerized tomography*), IRM (imagerie par résonance magnétique), radiographies conventionnelles, échographies, examens de rétine, de peau, et à l'analyse d'images histologiques.

Les performances de ces systèmes apprenants d'aide à la détection et au diagnostic sont supérieures à celles des outils précédemment disponibles. Actuellement, seuls sont autorisés par la réglementation française les systèmes dont les algorithmes sont sans

²⁵ Un système d'apprentissage profond utilise des couches cachées de réseaux de neurones artificiels. L'apprentissage par *deep learning* peut être supervisé, ou non.

possibilité d'apprentissage en continu, mais l'émergence de logiciels qui « apprennent » automatiquement offre une perspective d'évolution importante de ce secteur, tout en posant des questionnements éthiques spécifiques²⁶.

2.1.1 L'oncologie

Les cancers du poumon, du sein et de la prostate représentent un problème majeur de santé publique. Le critère déterminant pour réduire la mortalité due à ces cancers est la détection précoce de petites lésions, avant toute traduction symptomatique, dans les groupes à risque de développer un cancer du poumon ou de la prostate, ou dans le dépistage du cancer du sein, en population générale.

Or le dépistage massif en population générale – ou dans un groupe à risque numériquement important – requiert un grand nombre de praticiens pour interpréter les données ; de plus, cette approche nécessite un temps supplémentaire lors d'un examen ou nécessite une consultation particulière de prévention, démarche dont on sait qu'elle est particulièrement conditionnée aux ressources des personnes (en termes d'information, de moyens financiers, etc.) qui n'ont de fait pas la même facilité d'accès à des consultations ciblées de prévention.

Par ailleurs, des micro lésions peuvent échapper à l'œil humain, ce qui complexifie l'interprétation des images et requiert une expérience ciblée importante de la part des radiologues, dont la formation, l'expérience et la disponibilité sont très hétérogènes. De ce fait, certaines lésions ne sont pas diagnostiquées, malgré une double lecture des images de mammographie²⁷. L'efficacité des systèmes experts classiques d'aide à la décision n'a quant à elle pas été démontrée.

Au cours des cinq dernières années, des SIADM reposant sur des techniques d'apprentissage profond supervisé ont été développés et testés dans différentes cohortes et types d'imagerie²⁸. Outre la reproductibilité et le gain de temps, ces systèmes se caractérisent par une précision de détection proche de 95%^{29,30}. Cela leur confère une sensibilité élevée, notamment en matière de détection d'images anormales, voire de marqueurs imperceptibles par l'œil et par les systèmes traditionnels, permettant d'orienter le praticien vers la nature bénigne ou maligne, et d'affiner le diagnostic et la prise en

²⁶ Bitterman D.S. et al., (2020), "Approaching autonomy in medical artificial intelligence", *The Lancet digital health*, 2: e447-9.

²⁷ Hickman S.E. et al., (2021), "Adoption of artificial intelligence in breast imaging: evaluation, ethical constraints and limitations", *British Journal of Cancer*, 125(1):15-22 ; Lotter W., et al., (2021), "Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach", *Nature Medicine*, 27, 244-249 ; Kim H.E., et al., (2020), "Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study", *The Lancet digital health*, 2, 138-48 ; Yu K.H., et al., (2018), "Artificial intelligence in healthcare", *Nature Biomedical Engineering*, 2 : 719-31 ; Thomassin-Nagara I., et al., (2019), "Artificial intelligence and breast screening: French Radiology Community position paper", *Diagnostic and Interventional Imaging*, 100, 553-66.

²⁸ Voir par exemple la solution développée par la start-up française Therapixel : <https://www.therapixel.com/product-technology/> .

²⁹ E. Svoboda, (2020), "Deep learning delivers early detection", *Nature*, 587, S20-2.

³⁰ Ardila, D. et al., (2019), "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography", *Nature Medicine*, 25, 954-61 ; Blanc D., et al., (2020), "Artificial intelligence solution to classify pulmonary nodules on CT", *Diagnostic and Interventional Imaging*, 101 : 803-810.

charge du patient (même si, comme nous le verrons plus bas, cette sensibilité peut se payer par une perte de spécificité).

Certains SIADM sont approuvés par les autorités compétentes en Europe et aux États-Unis, mais la comparaison des résultats obtenus par ces SIADM avec une analyse faite uniquement par des radiologues n'est pas systématique, et la variabilité du niveau de validation entrave une généralisation de l'implémentation clinique comme nous le discutons plus loin. La Société française de radiologie a récemment publié ses recommandations à ce sujet³¹. À long terme, on peut espérer parvenir à une individualisation de l'analyse du risque en combinant pour chaque patient, données cliniques, épidémiologiques et radiologiques, dans un algorithme « multitâche » de façon à adapter les modalités de dépistage.

Beaucoup d'autres applications basées sur le traitement d'images connaissent une importante diffusion dans le champ de l'oncologie (par exemple, l'interprétation des images de coloscopie), comme dans d'autres spécialités médicales.

Du fait des performances et du gain de temps que permettent les SIADM, leur développement apparaît ainsi particulièrement adapté au dépistage précoce de certains cancers, en particulier aux campagnes de dépistage à large échelle dans une visée de santé publique.

2.1.2 Les maladies rétiniennes

Le dépistage à large échelle de trois maladies rétiniennes responsables de cécité précoce représente certainement l'une des applications parmi les plus avancées et les plus prometteuses des SIADM. Il s'agit de la rétinopathie diabétique (complication grave touchant 50 % des diabétiques de type 2), la dégénérescence maculaire liée à l'âge (DMLA, touchant 20 % des personnes âgées de plus de 80 ans) et du glaucome³². Ces pathologies disposent d'un traitement efficace si elles sont détectées à un stade précoce.

Les SIADM développés pour l'interprétation d'imagerie pour ces maladies sont appliqués dans le cadre de dépistages de routine³³. Ce dépistage par rétinographie numérique peut déjà être réalisé par des professionnels de santé non médecins ou non spécialistes d'ophtalmologie et peut bénéficier d'une interprétation en différé et à distance des images par l'ophtalmologiste en cas de difficultés d'accès aux soins.

La performance et la fiabilité de certains algorithmes de détection de la rétinopathie diabétique à partir d'images de rétine ont été démontrées³⁴. De plus, une étude prospective³⁵ en conditions réelles a testé et démontré l'efficacité d'un SIADM qui réalise l'examen et l'interprète localement en simultané, sans intervention d'un spécialiste

³¹ Thomassin-Nagara I., et al., (2019), "Artificial intelligence and breast screening: French Radiology Community position paper", *Diagnostic and Interventional Imaging*, 100, 553-66.

³² Il s'agit d'une dégénérescence progressive du nerf optique qui, en l'absence de traitement, aboutit au rétrécissement du champ visuel et à la cécité. Un dépistage systématique est recommandé à partir de 40 ans.

³³ Le *deep learning* est appliqué aux images de fond d'œil, de tomographie en cohérence optique (*optical coherence tomography*) de la rétine en 3D, ou encore d'analyse du champ visuel.

³⁴ Quellec G., et al., (2017), "Deep image mining for diabetic retinopathy screening", *Medical Image Analysis*, 39:178-193.

³⁵ Abramoff M.D., et al., (2018), "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices", *npj Digital Medicine*, 1:1-8.

humain, ce qui facilite une décision médicale rapide. La *Food and Drug Administration* (FDA) a validé et autorisé la mise sur le marché d'un tel SIADM (IDx-DR) en 2017. Il dispose également du marquage « CE » (conformité aux exigences fixées par la réglementation de la communauté européenne) à l'image d'un SIADM français, OphtAI qui a obtenu le marquage CE en 2019. Toutefois, la certification d'un tel SIADM est pour l'instant une exception.

2.1.3 La dermatologie

L'importance de la reconnaissance précoce de lésions de la peau et plus particulièrement de trois types de cancers, dont le mélanome, tumeur maligne redoutable, justifie les campagnes actuelles de dépistage de ces maladies par les dermatologues. Depuis plusieurs années, il a été démontré qu'un SIADM peut faire la distinction entre lésion maligne et bénigne de la peau dans des conditions expérimentales³⁶. Toutefois, les applications proposées en ambulatoire, notamment celles qui interprètent les images de lésions cutanées collectées par un *smartphone*, n'ont pas le degré d'avancement ni la rigueur d'évaluation des applications en ophtalmologie et sont encore au stade expérimental. L'agence fédérale américaine des produits alimentaires et médicamenteux (FDA) n'a encore autorisé aucun système de détection, et des mises en garde ont été publiées³⁷.

2.1.4 La cardiologie

L'électrocardiogramme (ECG, tracé obtenu par l'enregistrement et la transcription des courants électriques du cœur) et l'échocardiographie (échographie des structures du cœur) sont des pratiques courantes pour le diagnostic, le suivi et la détection de facteurs de risque de maladies cardiovasculaires. Si l'ECG n'est pas *stricto sensu* une image, l'interprétation des caractéristiques du tracé, très stéréotypées, se prête particulièrement à l'application d'algorithmes, et, en particulier à la détection de troubles du rythme cardiaque³⁸ (comme la fibrillation auriculaire) qui exposent les patients à des complications graves³⁹. Des enregistrements ECG prolongés répétés avec des dispositifs portables permettent leur détection. Couplés à des algorithmes d'apprentissage profond qui automatisent l'interprétation, ils peuvent aider au dépistage de ces arythmies et prévenir la survenue de complications invalidantes. Plusieurs de ces dispositifs ont été homologués⁴⁰.

³⁶ Esteva A., et al., (2017), "Dermatologist-level classification of skin cancer with deep neural networks", *Nature*, 542: 115-18.

³⁷ Matin R.N., Dinnes J., (2021), "AI-based smartphone apps for risk assessment of skin cancer need more evaluation and better regulation", *British Journal of Cancer*, 124, 1749-1750 ; Walter F.M., et al., (2020), "Further evaluation is required for smartphone-aided diagnosis of skin cancer", *The Lancet Digital Health*, 2 ; e104.

³⁸ Citons la fibrillation auriculaire (FA), souvent asymptomatique et fugace, qui favorise par exemple la survenue d'accidents vasculaires cérébraux thrombotiques (30 % des cas d'AVC) ou une décompensation cardiaque, et qui peut être traitée si elle est dépistée.

³⁹ Attia Z.I. et al., (2019), "An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction", *The Lancet*, 394 : 861-7 ; Hannun A.Y., et al., (2019), "Cardiologist level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network", *Nature Medicine*, 25 : 65-69.

⁴⁰ La start-up française Cardiologs a ainsi développé une plateforme web d'analyse ECG, utilisant des techniques d'intelligence artificielle, récemment homologuée par la FDA, ainsi que par un marquage CE, consultable au lien suivant : <https://cardiologs.com/>.

2.1.5 L'analyse d'images en histopathologie

En pratique, l'analyse histopathologique⁴¹ est essentielle pour consolider la certitude diagnostique, particulièrement en oncologie. Le diagnostic histologique se fonde sur des techniques de marquage tissulaire associées à des techniques moléculaires ou génomiques offrant une classification très fine, mais aussi très complexe, des cancers. L'interprétation du diagnostic histopathologique peut, comme en tous domaines, être subjective lorsqu'elle est réalisée par l'humain⁴². Or, elle est essentielle puisqu'elle fonde non seulement la décision thérapeutique mais aussi la prédiction de la réponse à un traitement. Plusieurs études ont montré que les SIADM appliqués à des images microscopiques numérisées à partir de lames d'histopathologie permettaient l'optimisation de la classification histologique, et pouvaient distinguer une lésion bénigne d'une lésion maligne. En combinant l'analyse des données cliniques, microscopiques et moléculaires, les technologies d'intelligence artificielle permettent également de prédire le comportement d'une tumeur^{43,44}. En outre, comme en radiologie et tous les champs d'application analysés ci-dessus, les logiciels optimisent la gestion du temps des professionnels en automatisant la lecture et en ne sélectionnant pour un examen par l'humain que des échantillons dont l'interprétation est sujette à discussion.

2.1.6 Systèmes d'intelligence artificielle apprenants, incertitudes et performances

Le fonctionnement des systèmes d'intelligence artificielle dont il est question dans cet avis est basé sur un traitement statistique de données, qui sont en général en très grand nombre. Parmi ces SIADM, nombreux sont ceux qui utilisent des réseaux de neurones artificiels organisés en couches, qui implémentent des fonctions mathématiques, comme des classificateurs.

Les données d'entrée peuvent être de différentes natures : des images, du texte, des sons, etc. Dans l'apprentissage dit « supervisé », le plus communément utilisé, chaque donnée d'entraînement est contourée, puis étiquetée. Une fracture sur une radiographie sera ainsi annotée « fracture », une tumeur maligne sur un scanner sera annotée « tumeur », etc.

Dans une première phase, dite d'apprentissage ou d'entraînement, les nombreux paramètres du réseau (le nombre de poids synaptiques peut être de plusieurs milliers, millions ou milliards selon la taille du réseau de neurones utilisé) sont calculés grâce à un algorithme d'optimisation de façon à obtenir une classification satisfaisante des données d'apprentissage, d'où l'appellation d'apprentissage « supervisé ». Considérons comme exemple une classification des parties d'images étiquetées « fracture » ou étiquetées

⁴¹ Elle est établie à partir d'échantillons de tissus ou de cellules prélevés lors d'une intervention chirurgicale, d'une endoscopie, ou lors de tout autre prélèvement, fixés sur lame de verre. Le diagnostic se fonde sur la distinction des différents constituants de l'architecture tissulaire et cellulaire au microscope optique après coloration et marquages. Ces lames sont aujourd'hui numérisées.

⁴² Laurent C., et al., (2017), "Impact of expert pathologic review of lymphoma diagnosis: study of patients from the French lymphopath Network", *Journal of Clinical Oncology*, 35, 2008-2017 ; Elmore, J. G. et al., (2015), "Diagnostic concordance among pathologists interpreting breast biopsy specimens", *Journal of the American Medical Association*, 313, 1122-32.

⁴³ Sryk C., et al., (2020), "Accurate diagnosis of lymphoma on whole-slide histopathology images using deep learning", *npj Digital Medicine*, 3, 63 ; Coudray N., et al., (2018), "Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning", *Nature Medicine*, 24, 1559-67.

⁴⁴ Colling R., et al., (2019), "Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice", *The Journal of Pathology*, 249: 143-50 ; Van der Laak J., et al., (2021), "Deep learning in histopathology: the path to the clinic", *Nature Medicine*, 27, 775-84.

« tumeur maligne ». L'objectif de cette phase est que le système ajuste ses paramètres pour identifier les éléments communs, en termes de paramètres caractérisant les composantes des images (intensité, alignements de pixels, etc.), fréquemment retrouvés dans telle ou telle caractéristique des images. On obtient ainsi un « modèle » statistique des données représenté par les valeurs optimisées des poids synaptiques.

Certains modèles nécessitent de fixer un certain nombre de paramètres externes à l'apprentissage, appelés hyperparamètres (typiquement le nombre de neurones ou la taille des lots ou « *batch size* », etc.) ; ces hyperparamètres sont déterminés sur un ensemble séparé de données (dit de validation).

Le modèle est alors prêt à l'emploi pour classer de nouvelles données selon les catégories apprises. On parle de prédiction : il s'agit de prédire à quelle catégorie appartient la nouvelle donnée. Les performances de cette prédiction dépendent de paramètres probabilistes caractérisant le système. Ces prédictions sont donc toujours assorties d'incertitudes. Pour les évaluer, on utilise dans le cas d'une classification avec deux catégories les notions suivantes :

- Vrai positif (vp) : le nombre de cas où la prédiction trouve une caractéristique effectivement présente dans la donnée ;
- Faux négatif (fn) : le nombre de cas où la prédiction ne trouve pas une caractéristique alors qu'elle est présente dans la donnée ;
- Faux positif (fp) : le nombre de cas où la prédiction trouve une caractéristique alors qu'elle n'est pas présente dans la donnée ;
- Vrai négatif (vn) : le nombre de cas où la prédiction ne trouve pas une caractéristique qui n'est effectivement pas présente dans la donnée.

Les performances du système sont alors évaluées par plusieurs métriques dont les plus communes sont les suivantes :

- Justesse (*accuracy*) $A = (vp+vn)/(vp+vn+fp+fn)$ soit le taux de prédictions correctes par rapport à l'ensemble des prédictions ;
- Précision (*precision*) $P = vp/(vp+fp)$ soit le taux de prédictions correctes des caractéristiques d'intérêt parmi l'ensemble des caractéristiques trouvées. La précision est une mesure de la pertinence ;
- Rappel (*recall*) $R = vp/(vp+fn)$ (ou sensibilité) soit le taux de prédictions correctes des caractéristiques d'intérêt parmi l'ensemble des prédictions de ces caractéristiques. Le rappel est une mesure de l'exhaustivité ;
- F_1 (appelée mesure F_1) $= 2 P.R/(P+R)$ soit la moyenne harmonique de la précision et du rappel qui permet de les combiner en leur accordant une importance égale.

Ces systèmes peuvent être utiles dans la mesure où ils permettent d'atteindre un bon niveau de reproductibilité de la décision⁴⁵.

L'aide des SIADM risque aussi d'être déterminante pour les centres moyennement spécialisés. En effet, pour les centres très spécialisés et compétents, il s'agit d'une aide certes intéressante, mais complémentaire. En revanche, insistons sur la nécessité de

⁴⁵ Morris, Alan H., Brian Stagg, Michael Lanspa, James Orme, Terry P. Clemmer, Lindell K. Weaver, Frank Thomas, et al. « Enabling a Learning Healthcare System with Automated Computer Protocols That Produce Replicable and Personalized Clinician Actions ». *Journal of the American Medical Informatics Association: JAMIA* 28, n° 6 (12 juin 2021): 1330-44. <https://doi.org/10.1093/jamia/ocaa294>.

démocratiser l'utilisation de ces SIADM qui pourront parfois contribuer à pallier le relatif défaut de compétences spécialisées.

Pourtant, ils soulèvent dans le même temps une tension éthique du fait que **les performances des SIADM comportent toujours une probabilité d'erreur**. Il est donc capital que l'équipe médicale qui l'utilise soit consciente de ces possibilités d'erreurs et exerce un contrôle sur les résultats proposés par le SIADM. Cette tension peut être explicitée au moyen des notions d'**incomplétude** et d'**incidentalome**.

Certes, les humains aussi commettent des erreurs, ce qui implique des questions éthiques. Cependant, cela incite à penser que ce n'est pas le risque d'erreur en lui-même qui pose problème, puisque l'humain en fait. Mais il existe chez l'humain des procédures intellectuelles et mentales permettant de relativiser ses décisions. Le problème posé par les dispositifs de décision utilisant les SIADM est qu'ils peuvent être binaires lorsqu'ils ne sont pas utilisés avec un regard critique. Cela soulève un problème éthique qui tient au fait qu'il n'existe pas à ce jour de manière éthiquement satisfaisante de faire en sorte que le diagnostic émis suite à l'utilisation de ces technologies puisse être apprécié de manière juste, en minimisant le risque, pour fournir une décision médicale dont l'impact sur le patient est souvent considérable. Par ailleurs, la question de la décharge de responsabilité faite à la machine soulève la question morale de l'origine et du sens de l'action humaine.

Premièrement, tout SIADM peut commettre des **faux négatifs** c'est-à-dire que le rappel (ou sensibilité) peut être trop faible. On parle alors d'**incomplétude**. Par exemple, une lésion ou une anomalie échappe au SIADM, qui indique que le cas est « négatif » au sens où il n'est pas problématique, bien que le cas soit en fait problématique. Deuxièmement, l'algorithme peut faire des **faux positifs**, c'est-à-dire identifier des lésions qui en fait n'en sont pas.

En outre, l'imagerie du corps humain, quelle qu'elle soit et quelle que soit la technique utilisée, expose à la découverte de **certains éléments fortuits, normaux ou anormaux, qu'elles contiennent et qui n'étaient pas recherchés**. Ces **découvertes imprévues**, sans relation aucune avec le contexte de réalisation de l'imagerie, sont classiquement nommées **incidentalomes** au sein de la communauté médicale. C'est le cas par exemple lors de la découverte d'un kyste osseux banal au cours de l'exploration d'une masse hépatique au scanner. Ces découvertes se répandent de plus en plus largement, ce qui implique de considérer leur survenue dans le cas d'analyses d'images par un SIADM.

Les taux d'incomplétude et de faux positifs vont en sens inverse. Plus on cherche à ce que le SIADM capture toutes les images de lésions, plus il en pourra en capturer de fausses ; inversement, plus on voudra qu'il ne capture que les vraies lésions, plus il pourra en oublier. Le SIADM est donc réglé, le plus souvent par les développeurs, rarement par les médecins, pour trouver un juste milieu entre les deux.

La découverte des incidentalomes s'est intensifiée aussi bien en médecine clinique que dans le domaine de la recherche biomédicale, corrélativement au développement d'examen d'imagerie⁴⁶ et à l'amélioration de la résolution des images. Une découverte fortuite est souvent source de situations complexes pour le patient et l'équipe médicale. Cette question, particulièrement prégnante pour la radiologie et la génétique, a été

⁴⁶ O'Sullivan J.W. et al., (2018), "Prevalence and outcomes of incidental imaging findings: umbrella review", *British Medical Journal*, 361: k2387.

précédemment traitée par le CCNE dans son avis n° 133 sur les « Enjeux éthiques des modifications ciblées du génome : entre espoir et vigilance ».

Il est probable que le recours aux SIADM modifie la fréquence des incidentalomes, soit parce qu'il en minimisera la détection (par exemple, le SIADM n'est pas entraîné à détecter une lésion osseuse lors d'une exploration hépatique au scanner), soit au contraire parce qu'il en détectera de façon inappropriée (exemple de la détection d'un cartilage de croissance normal en pédiatrie pris à tort pour une fracture).

Point de vigilance :

1) Pour éviter les risques d'incomplétude et d'incidentalome, il est capital que la signification d'une mesure qui évalue le compromis entre précision et sensibilité soit parfaitement clarifiée par les industriels.

2.1.7 L'interprétation de données génomiques

Les technologies d'intelligence artificielle ont la capacité singulière d'exploiter des quantités considérables de données générées par les techniques de séquençage d'ADN, dites de nouvelle génération (*Next Generation Sequencing*, NGS). Cette convergence entre numérique et génomique⁴⁷ a permis une avancée considérable des connaissances, non seulement pour la prédiction de la fonction des gènes, mais aussi pour la description d'associations entre le génotype et le phénotype dans de multiples maladies, de biomarqueurs pour certains patients, et la prise en charge thérapeutique individualisée des patients.

2.1.8 Implications éthiques : incomplétude et incidentalome

Si, techniquement, plusieurs systèmes apprenants existants sont en mesure de détecter une lésion et parfois d'en suggérer la nature maligne ou bénigne sans intervention directe humaine, le transfert complet du diagnostic à l'algorithme en pratique clinique soulève de multiples questionnements éthiques et juridiques et n'est pas encore autorisée en particulier en France à l'exception du secteur de l'ophtalmologie. De même, les algorithmes qui déterminent de façon non supervisée (sans annotation humaine des données) les critères leur permettant d'arriver à un résultat (*unsupervised learning*) ne sont pas actuellement autorisés.

L'équipe médicale doit être en capacité d'exercer ce que nous avons appelé plus haut sa pensée « globale » qui lui donne un point d'appui pour évaluer les propositions de diagnostic faites par un SIADM. Il est essentiel que le développement de ces systèmes d'aide au diagnostic n'en vienne pas à se substituer aux compétences professionnelles des praticiens et ne fasse pas courir le risque d'une perte d'éléments diagnostiques contextualisés pour les patients, et d'une perte d'expertise au cours du temps pour les médecins.

⁴⁷ La génomique est traitée au point 4.2. de ce document.

Point de vigilance :

2) Un SIADM, s'il peut rassurer par son fonctionnement rigoureux et automatique, ne plonge pas moins le patient comme l'équipe soignante dans un certain degré d'incertitude. Le maintien d'un contrôle humain lors de l'utilisation d'un SIADM apparaît comme indispensable mais ne lèvera pas forcément les incertitudes.

Recommandations :

- 1) Il est essentiel que **les méthodes diagnostiques déjà établies**, ne faisant pas intervenir a priori de SIADM, **continuent d'être enseignées et de faire l'objet de recherches** visant à les faire progresser.
- 2) La **place grandissante des SIADM** dans le champ des compétences médicales **nécessite des études approfondies** sur l'interaction entre les humains et les technologies d'intelligence artificielle pour évaluer l'impact des SIADM dans l'exercice de la médecine.
- 3) Dans un souci de **transparence et de traçage**, il faut indiquer l'utilisation d'un SIADM dans le compte rendu médical d'une consultation.
- 4) Ces éléments convergent en faveur d'un **contrôle humain** à toutes les étapes du soin, de l'indication des examens aux résultats des analyses et à l'interprétation contextuelle de ces résultats.

2.2 Les SIADM dans l'amélioration du parcours de soins

En dehors des SIADM réservés aux professionnels de santé, des applications ou des dispositifs sont proposés aux personnes elles-mêmes dans le cadre de leur parcours de soins. Outre qu'ils facilitent et sécurisent le maintien à domicile des patients, par leurs vertus pédagogiques ces nouveaux outils pourraient également contribuer à améliorer l'observance de certains traitements et à associer les patients à la prise en charge de leur maladie au quotidien. Ces bénéfices médicaux sont toutefois étroitement liés aux bénéfices économiques qui en résultent et cette intrication doit constituer un point de vigilance éthique important, sur lequel nous reviendrons.

2.2.1 L'orientation pré-diagnostique

De manière concomitante avec la démocratisation d'Internet, les patients se sont appropriés l'information médicale en ligne et ont adopté, en amont de leurs consultations médicales, des comportements de recherche active concernant leur état de santé.

Cette mutation des relations entre patients et médecins a suscité le développement, par des entreprises, d'une offre d'orientation pré-diagnostique (*diagnostic support system*) mise à disposition par le biais d'applications qui, à partir de questionnaires ou d'agents conversationnels (ou *chatbots*) collectant leurs symptômes, formulent une orientation diagnostique et une stratégie à adopter, en se fondant sur une base de connaissances associant données de la littérature et modèles de maladies. L'objectif premier est d'orienter correctement les patients vers le professionnel adapté, et, en particulier, de repérer les situations nécessitant une intervention d'urgence. Citons par exemple au

Royaume-Uni, Ada⁴⁸ ou Babylon⁴⁹, des plateformes qui se présentent comme universelles et préparent leur développement à l'international, dont la France.

Ces dispositifs représentent un champ économique très important qui pose de nombreuses questions éthiques (voir partie 2.2.4).

2.2.2 Aide à la gestion de la prise en charge de l'urgence vitale

À l'échelle des services d'urgence, le défi est de repérer les patients nécessitant une prise en charge immédiate, le délai d'instauration d'un geste thérapeutique vital pouvant être très court lors d'un accident vasculaire cérébral ou d'un choc septique par exemple. Cette priorisation est un exercice complexe et peut être ralentie par la surcharge des services d'urgences⁵⁰ ou l'absence d'informations sur les comorbidités antérieures du patient.

Plusieurs algorithmes de priorisation ont été testés, entraînés sur des données intégrant paramètres démographiques, constantes vitales, symptômes, histoire médicale et comorbidités, voire plus récemment l'analyse textuelle des dossiers médicaux partagés⁵¹ et/ou observations de professionnels de santé, ou encore des données d'imagerie cérébrale⁵². Ils ont été comparés aux arbres de décision traditionnels avec comme critère d'efficacité une priorisation efficace des patients arrivant aux urgences, leur bonne orientation vers les soins intensifs ou une hospitalisation classique, ou encore le délai de prise en charge thérapeutique. Ce dernier critère est essentiel aux urgences, notamment en cas d'accident vasculaire cérébral ou cardiovasculaire où l'on ne dispose que de trois ou quatre heures⁵³ avant l'instauration d'un traitement. De même, ces systèmes pourraient contribuer à mieux détecter les patients susceptibles d'une détérioration rapide du pronostic vital dans les unités de soins intensifs.

Au moins deux expériences françaises d'application de SIADM au sein de services de régulation du SAMU sont en cours. À Besançon, le projet franco-suisse SIA REMU ambitionne d'optimiser la régulation des appels d'urgence médicale en aidant les médecins dans leurs prises de décisions à partir de l'évaluation de la gravité des appels, s'appuyant sur un logiciel d'aide à la décision médicale développé grâce à l'IA. Une autre étude menée conjointement par le CHU de Bordeaux et l'Inserm prévoit le déploiement d'un système de surveillance d'événements de santé publique (accidents de voitures,

⁴⁸ Consultable au lien suivant : <https://ada.com/>

⁴⁹ Consultable au lien suivant : <https://www.babylonhealth.com/en-us>

⁵⁰ D'après plusieurs études, « 40 % des diagnostics d'admission lors de la première présentation aux urgences ne concordent pas avec le diagnostic final du patient », voir cette synthèse : Faqar-Uz-Zaman S.F. et al., (2021), "Study protocol for a prospective, double-blinded, observational study investigating the diagnostic accuracy of an app-based diagnostic health care application in an emergency room setting: the eRadaR trial", *British Medical Journal Open*, 11:e041396.

⁵¹ Klang E., et al., (2021), "Predicting adult neuroscience intensive care unit admission from emergency department triage using a retrospective, tabular-free text machine learning approach", *Scientific Reports*, 11, 1381 ; Liang H., et al., (2019), "Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence", *Nature Medicine*, 25 : 433-438.

⁵² Lynch C.L., et al., (2018), "New machine-learning technologies for computer-aided diagnosis", *Nature Medicine*, 24, 1304-12.

⁵³ Raita Y., et al., (2019), "Emergency department triage prediction of clinical outcomes using machine learning models", *Critical Care*, 23:64; Fernandes M., et al., (2020), "Predicting Intensive Care Unit admission among patients presenting to the emergency department using machine learning and natural language processing", *PLoS One*, 15 : e0229331.

symptômes respiratoires, etc.) visant à alléger le travail des régulateurs. Ces systèmes d'IA en question ont « appris » en se fondant sur l'analyse des comptes rendus des appels.

Les recherches citées ci-dessus en sont toutefois encore à un niveau expérimental. Elles sont principalement fondées sur des données rétrospectives (collectées précédemment dans un autre objectif) et monocentriques (provenant d'une seule structure ou d'un seul groupe de patients). Leur validation dans des études prospectives (pour lesquelles de nouvelles données sont collectées) en situation réelle, multicentriques (données provenant de plusieurs centres ou groupes de patients), est indispensable pour juger de leur efficacité et de leur avantage pour la pertinence de la décision médicale. Leur performance et leur capacité à accélérer la décision tout en améliorant les flux doivent être comparées à celles des méthodes en usage. Ces stratégies pourront ensuite être intégrées dans la pratique clinique au bénéfice du patient.

2.2.3 Aide au suivi des patients à domicile

Indépendamment d'un contexte d'urgence, des dispositifs de suivi permettent aux médecins de surveiller en continu les indicateurs biologiques de leurs patients (tension artérielle, glycémie, etc.), mesurés dans un contexte de vie réelle qui peut aller du logement à la pharmacie et même au supermarché où des cabines d'autodiagnostic commencent à être déployées. Cette démarche essentielle, compte tenu de la prévalence des maladies chroniques et de la désertification médicale, pourrait être particulièrement intéressante en cardiologie ou en diabétologie. Elles sont, de manière générale, pertinentes dans les domaines pour lesquels les facteurs environnementaux ou comportementaux sont décisifs pour l'évolution de la maladie (diabète, maladies cardiovasculaires, cancer, ou maladie de Parkinson par exemple via l'interprétation du mouvement ou de la tonalité de la voix) ou à domicile après une intervention chirurgicale.

Au-delà du simple suivi, les SIADM pourraient permettre d'automatiser et d'adapter le traitement en fonction de l'activité du patient. Des systèmes de détection de symptômes ou signaux inhabituels (baisse d'activité, fréquence accrue du rythme cardiaque, baisse de mobilité) peuvent aussi alerter les soignants et contribuer ainsi à une intervention plus rapide. C'est le cas par exemple de l'algorithme auto-apprenant produit par Diabeloop pour la gestion automatisée du diabète de type 1, qui a obtenu en 2020 l'aval de la Haute Autorité de Santé (HAS) pour le remboursement par l'Assurance maladie⁵⁴. On peut également citer la gestion du suivi des patients à domicile après une intervention chirurgicale, notamment grâce à une télésurveillance par SMS⁵⁵ et chatbots⁵⁶.

Les SIADM ont contribué au déploiement d'un suivi automatisé à domicile des patients hospitalisés en ambulatoire, en amont et en aval de leur séjour à l'hôpital, qu'il s'agisse de surveillance post-opératoire ou de traitements anticancéreux par exemple. Dans ce dernier cas, l'évaluation de la qualité de vie du patient et la prise en compte de son

⁵⁴ « Le capteur mesure toutes les 5 minutes le glucose interstitiel qui est transmis au terminal mobile. Ce terminal mobile vise à piloter automatiquement la pompe à insuline, ou proposer des recommandations à partir des données du capteur, des données historiques du patient, de la pompe à insuline et des données rentrées par le patient », voir : HAS, Avis de la Commission Nationale d'Evaluation des Dispositifs Médicaux et des Technologies de Santé (CNEDI-MTS) du 28 janvier 2020, 45 p.

⁵⁵ L'outil MemoQuest de Calmedica, adopté par plusieurs hôpitaux de l'APHP, est un robot conversationnel (ou « chatbot ») capable d'échanger par SMS avec les patients avant et après leur hospitalisation en ambulatoire.

⁵⁶ Voir : CNPEN, Avis n° 3 du 15 septembre 2021, *Agents conversationnels : enjeux d'éthique*, 38 p.

comportement et de sa tolérance au traitement dans son environnement quotidien sont des facteurs importants.

Des applications dans le domaine de la prédiction peuvent aussi être envisagées. Qu'il s'agisse de la détection précoce ou de l'évolution d'une maladie, de la réponse d'un patient à un traitement personnalisé ou encore du stockage et de l'intégration de multiples mesures réalisées en continu, le SIADM permet de constituer un système d'alerte. Il est alors utilisable en service de soins intensifs pour prédire avec davantage de précision des événements défavorables, ou la détérioration de l'état du patient. Il facilite ainsi une décision appropriée et en temps opportun.

Comme nous le mentionnions en introduction de cette partie, les considérations économiques dans un contexte de ressources hospitalières rares ont un rôle crucial dans ce transfert « ambulatoire » de soins précédemment réalisés à l'hôpital. L'influence de cette pression économique dans la transformation du système de santé par le numérique représente un questionnement éthique majeur.

2.2.4 Implications éthiques : bénéfices économiques versus bénéfices de santé

Les SIADM font l'objet d'une attention particulière dans le champ médical, notamment en raison du gain de temps qu'ils promettent dans un contexte de raréfaction des ressources hospitalières. Ce temps dégagé pour les médecins pourrait être utilisé pour investir des situations médicales complexes et/ou singulières dont les sensibilités échappent à l'approche proposée par l'IA. Le temps libéré par les SIADM permettrait ainsi de redécouvrir et de développer une source d'information majeure qui naît de la relation que le médecin – et tout professionnel de santé – noue avec la personne malade. Ce temps libéré permettrait en outre d'améliorer le temps d'information et d'explication auprès du patient pour servir la notion de consentement éclairé. Enfin, il pourrait être mis à profit pour développer les échanges interdisciplinaires et interprofessionnels qui fondent les décisions justes dans les situations à forte densité éthique.

Des questionnements demeurent concernant les performances réelles des SIADM. Celles-ci ne sont pas définitivement établies, et, au contraire, de nombreuses études ont montré leurs limites techniques. Nous reviendrons en partie 3 sur ces problèmes d'évaluation. D'autre part, les SIADM peuvent soulever des questions éthiques sur la situation et le rôle du patient. Enfin, elles ne manqueront pas de questionner les assurances complémentaires – deux points que nous abordons ci-dessous.

En étant considéré comme « acteur de sa santé » par les outils numériques dans le cadre de télé-médecine et de télésoins utilisant de l'IA, le patient peut se sentir investi d'une responsabilité. Cela peut engendrer de l'anxiété et faute de compétences médicales et/ou numériques, il peut se sentir démuni, sans interlocuteur de l'équipe de soins. Les personnes vulnérables peuvent particulièrement pâtir de cette situation⁵⁷.

C'est pourquoi les SIADM posent la double question éthique des inégalités face au numérique, mais aussi face à un accès de plus en plus difficile au système de soins incluant la question des déserts médicaux. Les obstacles à l'accès aux soins ne peuvent

⁵⁷ Voir : CNPEN, Bulletin de veille n° 3 du 21 juillet 2020, *Enjeux d'éthique liés aux outils numériques en télémédecine et télésoin dans le contexte de la Covid-19*, 21 p.

être supprimés par les seuls outils numériques dont l'appropriation par les patients est inégale et dépend de multiples déterminants économiques, géographiques, sociaux, générationnels, etc. Or on sait que ces difficultés ont tendance à se cumuler sur les mêmes populations. Les personnes qui rencontrent des difficultés d'accès à l'hôpital seront aussi celles qui peineront à utiliser seules les dispositifs numériques en santé ou les cabines d'autodiagnostic, pour ne prendre que cet exemple. En outre, le recueil à distance du consentement de ces patients soulève également de nouveaux points de tension. Enfin, le risque de mésusage et de dérive vers des mesures coercitives (concernant la non-observance d'un traitement ou d'une mesure prescrite, les conduites à risques, etc.) doit être pris en compte.

Une autre tension spécifique soulevée par ces dispositifs est celle de la balance entre le bénéfique en termes de prise en charge médicale promise aux patients et le bénéfique économique visé (en termes de ressources en personnel médical). L'orientation pré-diagnostique, qui propose une intervention optimisée dans un contexte de crise hospitalière et de raréfaction des praticiens, pourrait devenir un obstacle à l'admission des patients à l'hôpital si les critères établis étaient trop exigeants, ou dans les cas où les caractéristiques spécifiques à certains patients ne seraient pas prises en compte dans les algorithmes établis.

Les possibles erreurs au stade d'orientation pré-diagnostique peuvent générer un risque potentiellement grave pour le patient. Ce risque est d'autant plus important que les structures hospitalières connaissent une progression exponentielle d'utilisateurs dont les pathologies pourraient être prises en charge par la médecine de ville, laquelle n'est pas organisée pour répondre à toute la demande. Certains patients n'ayant pas accès à un médecin généraliste sont aussi plus susceptibles d'éprouver des difficultés d'accès à l'hôpital et de se retrouver par conséquent sans alternative de soins possible. Les technologies d'intelligence artificielle appliquées à la santé auraient alors pour effet de transférer le poids de la désorganisation du système de soins vers ces patients les plus vulnérables.

La possibilité accrue par les SIADM de suivi en soins ambulatoire peut inciter à recourir à cette forme de soin bien meilleur marché que l'hospitalisation en occultant les spécificités de certaines catégories de patients ou de cas spécifiques nécessitant une hospitalisation. Ainsi, généralement, les systèmes de gestion des flux, essentiellement promus par les hôpitaux pour des raisons d'économie, risquent de s'ajouter aux multiples barrières d'accès aux soins pour les usagers les plus vulnérables. Il est vrai cependant qu'un patient ne doit pas être gardé dans un lit de soins critiques s'il n'en a pas besoin. On se retrouve donc dans une situation où la technologie est considérée comme un moyen de pallier l'organisation déficiente d'un système de soins qui évoluerait au gré des innovations numériques.

Il est par ailleurs capital que la qualité des soins prévale sur les intérêts économiques associés au développement des SIADM. Cela concerne tant le secteur de l'innovation technologique et industrielle qui soulève des tensions éthiques traitées plus haut, que celui des assurances complémentaires, à propos desquelles d'autres problèmes surgissent.

En effet, la diffusion de l'IA de diagnostic médical amène à souligner deux types de questionnements éthiques s'agissant du rôle que peuvent jouer ou non les assurances santé complémentaires. Notons en premier lieu les risques de discrimination du fait que les acteurs de la couverture santé complémentaire ont accès à des données sur l'état de

santé des usagers. Ces données peuvent théoriquement permettre aux acteurs de l'assurance santé complémentaire de proposer des offres individualisées et évolutives. Le cadrage juridique français actuel proscrit, en principe, une telle individualisation⁵⁸.

Pour autant, une réflexion éthique sur les SIADM implique d'intégrer également ces enjeux dans leurs perspectives à moyen terme. Or, cette évolution possible des assurances santé complémentaires nous confronte à de nouvelles problématiques, et notamment à un risque de discrimination si la possibilité de cette modification n'est pas pensée en amont. Le cadre actuel vise, en particulier, à éviter les risques de sélection adverse des profils en fonction de l'état de santé. Par ailleurs, une telle évolution pourrait induire un effet d'amplification des biais existants et serait également possiblement attentatoire à la protection des données personnelles des usagers.

Notons en second lieu que le rôle des acteurs de l'assurance santé complémentaire ne doit pas être pensé uniquement sous le prisme de risques éthiques. L'absence d'intervention de leur part se révélerait également éminemment problématique. Les assurances santé complémentaires ont, en effet, également un rôle majeur à jouer pour permettre l'accessibilité pour tous à l'innovation en santé. Or, si elles ne prennent pas part au financement de l'innovation, nous pourrions assister à une rupture d'égalité entre les usagers concernant l'accès à l'innovation en IA de diagnostic médical ou au développement d'une forme de système de soins alternatif matérialisé par une multiplication de SIADM non homologués ou de confort.

Or, un meilleur accès à l'innovation pour tous est indispensable. En l'absence de prise en compte du principe d'égalité d'accès des usagers, nous pourrions être confrontés à une situation où certaines personnes pourraient bénéficier d'une médecine algorithmique de pointe et où d'autres se retrouveraient mises de côté pour des raisons économiques. L'assurance santé dans sa globalité a un rôle essentiel à jouer sur ces questions même si naturellement une prise en charge, au moins partielle, de l'accès à l'innovation en IA de diagnostic médical par l'assurance santé obligatoire apparaît une condition *sine qua non* s'agissant de la préservation du principe de solidarité régissant notre système de santé.

Points de vigilance :

- 3) Les SIADM doivent être envisagés en tant qu'**outils complémentaires** des réponses à apporter aux insuffisances du système de soins, en particulier de la désertification médicale, mais ne **doivent pas être considérés comme des solutions de substitution** des équipes médicales.
- 4) Le processus décisionnel conduisant à l'utilisation de SIADM doit impérativement prendre en compte **le souci du soin avant les considérations économiques**. Il doit être considéré comme un moyen de lutte contre les inégalités sociales et territoriales de santé.
- 5) Maintenir une vigilance contre les risques de sélection adverse qui pourraient être opérés par les assurances complémentaires en particulier au moyen d'outils de diagnostic actuariels fondés sur des techniques d'IA.

⁵⁸ CCNE, avis n°124 du 21 janvier 2016, *Réflexion éthique sur l'évolution des tests génétiques liée au séquençage de l'ADN humain à très haut débit*, 84 p.

Recommandation :

5) Engager une **concertation nationale** sur la mise en accessibilité économique des SIADM pour les usagers du système de soins. À cet égard, l'identification d'une **section consacrée au financement de l'IA dans les lois de financement de la Sécurité sociale** pourrait être expertisée. Dans ce débat, le rôle des acteurs de la protection complémentaire pour identifier et stimuler la mise en accessibilité de l'innovation en IA devrait être attentivement examiné.

2.3 Dispositifs portables⁵⁹ utilisant des SIADM hors du parcours de soins

On constate aujourd'hui la multiplication de capteurs, applications mobiles, réseaux sociaux, géolocalisation, ou simples questionnaires, qui recueillent régulièrement des informations en temps réel ainsi que la mesure de multiples paramètres corporels⁶⁰ qu'ils intègrent pour offrir un résultat qu'on peut qualifier de paramètres de santé ou de bien-être. Ces usages peuvent entrer dans le cadre du suivi ou du diagnostic de maladies (voir ci-dessus) ou être d'ordre « récréatif » et se situer hors de toute supervision médicale. Parmi la palette de dispositifs disponibles, certains permettent d'estimer le degré d'anémie en analysant des photos du lit de l'ongle, d'autres proposent des systèmes d'échographie sur *smartphone* potentiellement manipulables par des non-médecins et permettent de générer et d'interpréter des échographies ou des électrocardiogrammes à distance, offrant au patient une première orientation diagnostique. Enfin, d'autres applications sont utilisables en dermatologie (voir 2.1.3. de ce document). Ces dispositifs portables (*digital wearables* tels que montres connectées, bracelets enregistreurs, applications de smartphone, etc.) sont parfois qualifiés de « selfies médicaux ».

Les données mesurées par ces dispositifs peuvent être recueillies par des professionnels non médicaux. Ceux-ci peuvent donc aider au dépistage de maladies chroniques (ou qui ne le sont pas, comme les maladies neuro-évolutives) souvent négligées ou diagnostiquées avec retard (ex. : hypertension artérielle, diabète) mais également au suivi de ces affections lorsque l'accès aux soins est difficile et/ou coûteux. À ce titre, ces outils peuvent influencer positivement sur notre système de santé, en évitant des déplacements longs, coûteux, fatigants et en diminuant les délais de rendez-vous. À l'inverse, le développement grand public de ces dispositifs s'accompagne d'un risque d'augmentation des consultations médicales dues à l'anxiété que peuvent générer chez certaines personnes les données issues des dispositifs portables utilisant des SIADM.

Pour être véritablement utiles aux professionnels de santé, outre une validation régulière et rigoureuse, il faudrait pouvoir croiser ces données mobiles à celles du dossier médical partagé du patient, ce qui soulève des questions logistiques, juridiques et éthiques considérables concernant les données de santé, qui sortent du cadre de cet avis.

⁵⁹ En anglais : *wearables*.

⁶⁰ Rythme cardiaque, respiratoire, pression artérielle, température, phases du sommeil, voix, mouvement, etc. Par exemple, une montre connectée peut permettre la détection d'anomalies du rythme cardiaque (fibrillation auriculaire). C'est le cas de l'Apple watch : Champion E.W., Jarcho J. A., (2019), "Watched by Apple", *The New England Journal of Medicine*, 381; 1964-5. ; Perez M.V., et al., (2019), "Large-scale assessment of a smartwatch to identify atrial fibrillation", *The New England Journal of Medicine*, 381 : 1909-17. Voir également : Rodin J.M., et al., (2021), "The hopes and hazards of using personal health technologies in the diagnosis and prognosis of infections", *The Lancet Digital Health*, 3, 7, e455-e461.

Des données plus subjectives peuvent être aussi collectées à l'aide de questionnaires mis en œuvre dans des applications pour *smartphone* et objets connectés ou par la collecte d'indicateurs de comportements supposés détecter des signes de dépression, par exemple. L'analyse de la démarche d'une personne ou de sa voix peut par exemple être utilisée pour diagnostiquer la maladie de Parkinson ou un syndrome dépressif. Beaucoup d'études se sont également intéressées au domaine du comportement (*personal sensing*), où les biomarqueurs détectant le sommeil, le contexte social, l'humeur ou le stress sont le plus communément utilisés. Ces données pourraient aider à repérer plus facilement et rapidement les personnes nécessitant une prise en charge, mais le traitement ou la diffusion de ces informations peuvent également les exposer à des risques en termes de protection des données à caractère personnel et posent donc d'importantes questions éthiques. D'autre part, leur qualité, qui n'aura souvent pas été contrôlée en tant que dispositif de soin, est très inégale.

Comme nous l'évoquons ci-dessus, certains de ces dispositifs sont déjà utilisés pour le suivi à distance de patients atteints de diverses maladies chroniques. L'intérêt clinique ou personnel (dans le cas d'applications de bien-être) de ces outils doit d'autant plus faire l'objet d'une analyse critique et s'inscrire dans le cadre d'une politique de santé. Leur développement, dans un marché en pleine expansion, ne doit pas constituer une fin en soi et il est important de s'assurer que leur développement vise à en faire un dispositif médical et ne répond pas à de seules fins récréatives ou commerciales.

2.3.1 Implications éthiques : risques d'avis non professionnels

On observe qu'un marché d'objets et de capteurs qui se trouve à la limite du système sanitaire institué, voire à l'extérieur, se diffuse et pourrait peut-être bénéficier à la santé de la population. Néanmoins certaines études relèvent de sérieuses limites⁶¹ et des risques d'usage inappropriés. Insistons sur trois des questions d'éthique qui accompagnent le développement de ce marché.

Premièrement, la **protection des données** personnelles produites et collectées dans ce cadre doit être assurée. Or la plupart du temps, aucun outil pédagogique ne permet de comprendre aisément ce que font ces dispositifs des données dont ils disposent : comment ils les stockent, avec qui ils les partagent, etc. Il faut donc veiller à ce que les données soient traitées, au même titre que celles qui sont produites dans le cadre du système de soins, comme des données de santé sensibles dont le traitement est plus contraignant en termes de protection que les données simples que nous produisons dans d'autres cadres récréatifs (ex. : réseaux sociaux). Attendu que les données peuvent devenir médicales par destination, il importe de mettre en œuvre les moyens de les identifier comme telles lorsque c'est le cas. Les utilisateurs doivent donc être informés des procédures de sécurité qui encadrent la production et le stockage de ces données et ils doivent pouvoir en contrôler l'usage dans le cadre de l'expression d'un consentement, ou au moins d'une possibilité d'opposition. L'une des questions qui se posent est de savoir dans quelle mesure ces données pourraient à terme être associées à l'espace numérique en santé « Mon espace santé » en cours de déploiement⁶². S'il s'avérait qu'un rapprochement entre les deux constituait un objectif, il conviendrait de faire converger les

⁶¹ Par exemple : Y. Demeure, « Selon une étude, l'auto-diagnostic sur le Web marche dans seulement un tiers des cas », *SciencePost*, mai 2020.

⁶² Consultable au lien suivant : <https://esante.gouv.fr/strategie-nationale/mon-espace-sante>.

règles de production, de stockage et d'utilisation dans le sens le plus protecteur pour les données personnelles des usagers et de veiller à en informer les usagers.

Deuxièmement et plus généralement, au-delà du seul problème de la protection des données, la diffusion à terme de **ces SIADM portables pourrait contribuer à éloigner les patients les plus avertis (en termes d'usage numérique) du système de soins conventionnel strictement encadré**. Leurs performances en tant qu'outils de santé ne sont que rarement homologuées et, d'autre part, ils sont susceptibles de creuser un fossé entre l'offre de santé soumise à régulation étatique et d'autres pratiques de soins non régulées. Il y a un risque alors de désintermédiation médicale ou de bouleversement de la relation patient-médecin. À l'évidence, les SIADM portables ne remplaceront néanmoins pas la relation qu'un patient peut avoir avec un médecin. La voie d'avenir à soutenir est, au contraire, celle d'un appui à l'intervention des professionnels de santé pour libérer, par le concours de la technologie, davantage de temps humain auprès du patient.

Troisièmement, **ces SIADM peuvent générer des inégalités de traitement entre les personnes** les plus favorisées – en capital économique, mais aussi en capital relationnel et culturel (incluant la culture numérique) – et les autres. Les premières pourraient en effet s'offrir les services fournis par les entreprises privées pour répondre à leurs besoins spécifiques, alors que les autres s'en verraient exclues – les entreprises privées ne développant des services que si ceux-ci sont rentables.

Point de vigilance :

6) Prendre en compte le **risque de dérives d'avis non professionnels, de traitement non sécurisé des données, et d'exercice illégal de la médecine** accentuant la vulnérabilité des personnes. Les SIADM doivent relever de pratiques médicales strictes.

III. UN ENCADREMENT DES SIADM QUI RESTE A CONSTRUIRE

Dans les disciplines médicales mentionnées ci-dessus, la littérature traitant des SIADM dans l'approche diagnostique reste avant tout une littérature de recherche, dont la qualité est très hétérogène. La majorité des applications étudiées n'a pas de validation officielle⁶³. Le fossé est encore grand entre les promesses annoncées⁶⁴ – souvent à l'initiative des industriels, compte tenu de la croissance très importante de ce marché – lors de la conception d'un algorithme scientifiquement robuste, et la réalité de son transfert dans la pratique clinique⁶⁵. On peut identifier deux étapes interdépendantes par lesquelles un SIADM doit passer pour être adopté dans la pratique. Il s'agit premièrement de l'évaluation de l'efficacité clinique du SIADM qui ouvre la voie à son éventuel remboursement par la sécurité sociale et, deuxièmement, de la certification de conformité qui autorise sa mise sur le marché.

3.1 Pour une évaluation des SIADM qui protège la relation de soin

L'utilisation des SIADM se trouve au croisement de deux domaines, les technologies d'intelligence artificielle et la médecine, qui ont des procédures d'évaluation scientifique et techniques différentes. Il y a une difficulté spécifique, encore en discussion au niveau international, à définir une évaluation des SIADM qui soit acceptable scientifiquement et socialement. En particulier, le développement clinique ne peut être fait sans que le SIADM n'ait été évalué au regard de son efficacité réelle en matière de soins. Concernant cette évaluation, une double question capitale intervient. D'abord celle de savoir comment produire une évaluation sincère et utile de ses performances, et ensuite de contrôler l'influence prédominante du marché dans ces jugements, avec le risque que celui-ci ne dicte les transformations du système de soins.

3.1.1 Le rôle de la Haute Autorité de Santé (HAS) dans l'évaluation des SIADM

En France, l'organisme chargé d'évaluer les SIADM sous le rapport de leurs bénéfices cliniques est la Haute Autorité de Santé (HAS). Son rôle consiste à identifier l'apport d'un produit sous l'angle de l'utilité pour le patient⁶⁶. Elle procède à l'évaluation des algorithmes d'IA lorsqu'ils sont inclus dans un dispositif médical. Toutefois, elle ne procède qu'à l'évaluation des dispositifs médicaux qui font l'objet de remboursement par l'Assurance maladie auprès des patients. **De fait, la presque totalité des algorithmes utilisés par les**

⁶³ D'après Rezazade Mehrizi, M.H., et al., (2021), "Applications of artificial intelligence (AI) in diagnostic radiology: a technography study", *European Radiology*, 31, 1805–1811 : « Plus de la moitié des applications (60 %) n'ont pas été validées par un organisme officiel, et parmi celles qui le sont, 60 % le sont par la FDA, 62 % ont un marquage CE, et 32 % les deux ». Ainsi, en octobre 2020, 29 de ces systèmes ont été validés par la FDA et 35 sont en attente (contre 2 en 2012).

⁶⁴ Wilkinson J., et al., (2020), "Time to reality check the promises of machine learning-powered precision medicine", *The Lancet Digital Health*, e677-e680.

⁶⁵ Keane P.A., Topol E., (2018), "With an eye to AI and autonomous diagnosis", *npj Digital Medicine*, 1, 40. ; He J., et al., (2019), "The practical implementation of artificial intelligence technologies in medicine", *Nature Medicine*, 25:3.0-36.

⁶⁶ Audition de représentants de la HAS par le groupe de travail du CCNE et du CNPEN ; voir aussi : https://www.has-sante.fr/jcms/c_928541/fr/comprendre-l-evaluation-des-dispositifs-medicaux.

professionnels de santé se voient donc exclus de l'évaluation par la HAS puisqu'il n'ont pour la plupart pas vocation à être vendus aux patients⁶⁷.

Ce défaut d'évaluation – aujourd'hui facultative et à l'initiative des entreprises – est susceptible d'entraver légitimement la confiance de la population envers les SIADM. Pour répondre à ces manques, la HAS a développé une grille socle permettant une évaluation qui contient une partie éthique complémentaire à celle prévue par le règlement européen ; elle vise à évaluer ce qui est une véritable innovation utile au patient et qui sera peut-être financée par l'Assurance maladie.

Sans un élargissement du périmètre d'évaluation de la HAS à tous les établissements de santé ou professionnels de santé qui veulent s'équiper de logiciels d'IA, les industriels peuvent avancer des promesses qui s'avèrent difficilement mesurables. Deux problèmes majeurs se posent : d'abord, pour les dispositifs destinés aux professionnels, aucun contrôle des bénéfices cliniques n'est imposé et ensuite, selon la HAS, la plupart des dispositifs étant présentés comme contenant de l'IA n'ont été soumis à aucune évaluation.

3.1.2 Implications éthiques : pour une meilleure explicabilité et une tempérance des promesses

L'une des solutions imaginées pour évaluer les SIADM serait de pouvoir garantir un certain niveau d'explicabilité du fonctionnement des algorithmes et de leurs résultats. L'explicabilité est la possibilité de fournir une raison compréhensible par l'utilisateur humain du résultat produit par le système qui soit également représentative de son fonctionnement. **Ainsi, si un SIADM d'analyse d'images identifie une lésion, il faut que le médecin soit capable d'expliquer cette lésion et si elle doit être ignorée ou prise en compte.** L'explicabilité est la capacité de donner un sens clinique aux résultats informatiques. Elle dépend à la fois de ce qui est expliqué et du degré de technicité informatique du praticien. En effet, pour indiquer si une lésion repérée est une erreur ou non, il faut d'une part connaître la spécialité et, de l'autre, comprendre ce qui, dans le fonctionnement du SIADM, l'a conduit au résultat produit. Ce niveau du médecin en informatique peut être nul ou presque, il peut aussi être informé, ce à quoi s'attachent certains chercheurs en informatique en produisant des « guides » d'aide à l'explicabilité⁶⁸ ainsi que de nombreuses recherches⁶⁹.

Une intermédiation pourrait également être nécessaire. Elle consisterait à créer sinon de nouveaux métiers, au moins de nouvelles compétences pour le médecin afin de couvrir la compréhension des dispositifs de SIADM.

Notons que cette notion d'explicabilité ne peut pas être étendue du diagnostic à court terme au traitement à plus long terme. En effet on peut imaginer, en particulier dans le cas des maladies rares, qu'un traitement soit proposé à partir des informations recueillies, qui ne sera pas très compréhensible et dont on ne verra jamais les résultats immédiatement. Celui-ci ne sera donc pas « explicable » au sens proposé ci-dessus mais il reste bien

⁶⁷ Voir le site de la Haute autorité de santé : https://www.has-sante.fr/jcms/p_3212876/fr/un-nouvel-outil-pour-l-evaluation-des-dispositifs-medicaux-embarquant-de-l-intelligence-artificielle.

⁶⁸ Bennetot, Adrien, Ivan Donadello, Ayoub El Qadi, Mauro Dragoni, Thomas Frossard, Benedikt Wagner, Anna Saranti, et al. « A Practical Guide on Explainable Ai Techniques Applied on Biomedical Use Case Applications ». *SSRN Electronic Journal*, 2022.

⁶⁹ A. Barredo Arrieta et al., (2020), "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI", *Information Fusion*, 58, 82–115.

entendu nécessaire. Il faut donc clairement distinguer les deux situations, à savoir l'explicabilité à court terme, que l'on peut contrôler, et l'explicabilité à moyen et long terme, qui est aujourd'hui impossible à contrôler.

En parallèle, il convient de se demander si l'état de santé des patients et, plus largement le système de soins, bénéficient directement de l'usage de ces technologies d'intelligence artificielle. Peu d'études font une comparaison rigoureuse de leur utilité ou de leur pertinence par rapport à l'arsenal existant, et peu analysent la performance de l'algorithme sur des cohortes externes indépendantes, en multicentrique. Des critères importants non directement médicaux tels que des critères économiques, de gestion du temps ou des moyens humains comme matériels doivent être intégrés à la grille d'évaluation de ces systèmes.

L'augmentation de la charge de travail des radiologues, l'explosion de la demande d'examens d'imagerie, la sophistication des techniques d'imagerie jouent un rôle important dans la multiplication de ces systèmes d'aide à l'interprétation d'images radiologiques. C'est particulièrement vrai en ce qui concerne les campagnes de dépistage à large échelle, qui ne peuvent être gérées que *via* des outils numériques. Un gain de temps a souvent été mentionné au bénéfice des professionnels de santé, plus disponibles pour l'étude de cas complexes ou l'échange avec leurs patients.

En ce sens, les promesses parfois exorbitantes faites par les industriels pour commercialiser leurs produits soulèvent de profondes tensions éthiques. Beaucoup d'économistes évoquent aujourd'hui une *économie de la promesse* en matière de nouvelles technologies pour désigner les financements parfois spectaculaires accordés à des SIADM qui ne sont encore que des promesses de solution.

Cette économie doit être accompagnée par **une éthique de la promesse**. D'abord, la promesse ne doit être exprimée que lorsqu'un ensemble de conditions pragmatiques ont été réunies pour qu'elle soit réaliste. En particulier, l'accessibilité aux bases de données ayant servi au développement du SIADM doit être prouvée. D'autre part, une promesse non tenue doit être sanctionnée d'une façon ou d'une autre. Or les industriels sont, en l'état actuel, rarement tenus de prouver le réalisme de leur projet. Pour pallier ces lacunes, il faudrait exiger a minima que les industriels puissent rendre compte de l'usage qui a été fait des financements et expliquer pourquoi le but poursuivi n'a pas été atteint.

Recommandations :

6) Créer les **conditions de la confiance** en incitant les développeurs à fournir un certain niveau d'**explicabilité** du SIADM qu'ils mettent sur le marché.

7) Promouvoir le **développement de nouvelles compétences professionnelles** permettant d'explicitier aux soignants les propriétés des SIADM.

8) Créer les conditions de la confiance en promouvant le développement de **justifications étayant les promesses** et, d'autre part, la **publication de résultats de recherche négatifs**, c'est-à-dire ceux qui ne confirment pas les promesses annoncées.

3.2 La certification, étape cruciale pour des SIADM de confiance

Le secteur médical fait l'objet de processus de régulation stricts, particulièrement concernant les produits de santé à destination des patients et principalement les médicaments. *A contrario*, dans le cas des dispositifs innovants reposant sur des techniques d'IA, les autorités de santé sont confrontées à un flou juridique associé à leur nouveauté et à des procédures peu adaptées.

En ce qui concerne la certification des SIADM, en amont de la mise sur le marché, il existe deux régulations à l'échelle européenne, le RGPD et le marquage CE. La conformité au RGPD ne rentre pas dans le cadre de cet avis. Pour le marquage CE, la certification des SIADM requiert qu'ils soient catégorisés comme dispositifs médicaux ou dispositifs médicaux de diagnostic *in vitro*. Il vise à évaluer la balance bénéfice-risque associée à un SIADM. Au niveau national, l'autorité compétente en la matière en France est l'Agence nationale de sécurité du médicament et des produits de santé (ANSM). Il n'est pas question de remettre ici en cause la caractérisation des SIADM comme dispositifs médicaux, mais d'interroger les processus de marquage qui sont mis en œuvre pour ces systèmes très singuliers comparativement à, par exemple, un abaisse-langue ou des béquilles.

3.2.1 Une évolution de la réglementation qui peine à suivre le rythme de diffusion des SIADM

L'évolution de la réglementation concernant les SIADM s'est accélérée à partir de 2007 avec la publication d'une directive⁷⁰ du Conseil de l'Union européenne relative aux dispositifs médicaux. Celle-ci introduit une définition du logiciel en tant que dispositif médical à part entière et intègre l'idée qu'un logiciel peut, au regard de sa destination, avoir une destination compatible avec un dispositif de diagnostic médical ou un dispositif médical de diagnostic *in vitro*.

En 2010, des travaux européens et internationaux viennent compléter ces textes en questionnant la définition du logiciel de dispositif médical, et la manière d'évaluer la balance bénéfice-risque en la matière. Le positionnement des technologies d'IA reste balbutiant : on parle alors de logiciels externes qui fournissent des informations légèrement plus subtiles qu'un simple affichage de résultats, le terme utilisé est alors « aide au diagnostic ».

En 2017, un nouveau règlement commence à évoquer l'IA et un certain nombre de points sont revus pour adapter plus spécifiquement la réglementation au logiciel. Parmi les étapes importantes de ces nouveaux textes, on note l'introduction d'une classification de risques (selon 4 classes) concernant l'utilisation du dispositif médical, dont l'objet est de diriger le fabricant vers des procédures d'évaluation de la conformité.

Les mêmes exigences s'appliquent aux dispositifs médicaux et aux dispositifs de diagnostic *in vitro*, néanmoins, par rapport à des dispositifs qui montrent un faible risque, un niveau plus élevé de mode de preuve est demandé au fabricant. Les directives précédentes incluaient ces dispositifs dans des classes de risques très bas, y compris des

⁷⁰ Directive 2007/47/CE du Parlement européen et du Conseil du 5 septembre 2007 modifiant la directive 90/385/CEE du Conseil concernant le rapprochement des législations des états membres relatives aux dispositifs médicaux implantables actifs, la directive 93/42/CEE du Conseil relative aux dispositifs médicaux et la directive 98/8/CE concernant la mise sur le marché des produits biocides.

logiciels qui faisaient par exemple des traitements d'images de mammographie auparavant classés au même niveau qu'un simple abaisse-langue.

On observe donc une évolution notable depuis les conditions de classification élaborées dans les années 1990 concernant le diagnostic médical physique (échographe, bistouri ou abaisse-langue par exemple). Avec le nouveau règlement, la classification est faite en fonction du risque lié à l'utilisation de ce logiciel en prenant en compte, par exemple, dans les aspects de conception, la protection contre des failles de cybersécurité.

Dans ce contexte en forte évolution, le **terme d'IA n'apparaît pas actuellement dans la réglementation française : on parle de logiciel d'aide à la décision diagnostique ou thérapeutique.**

3.2.2 Une procédure de conformité matérialisée par le marquage « CE »

Le système d'autorisation de mise sur le marché (AMM) appliqué aux médicaments n'est pas valable pour les SIADM. Ces dispositifs impliquent une réglementation « nouvelle approche » qui fait appel à une **procédure de conformité reposant sur un marquage CE**, procédure suivie par le fabricant avec des considérations plus ou moins importantes en fonction de la classe de risque du dispositif médical.

En France, et dans plusieurs pays d'Europe, les demandes font l'objet d'autorisations coordonnées au cours desquelles des processus d'évaluation font appel à des comités d'éthique se prononçant sur les aspects éthiques des SIADM. L'autorité compétente - en France il s'agit de l'Agence nationale de sécurité du médicament et des produits de santé (ANSM) - vient ensuite évaluer le dispositif au regard de ses bénéfices et des questions auxquelles ils doivent répondre. Le contrôle de conformité requiert l'intervention d'un organisme tiers dit notifié⁷¹. La pertinence de la démonstration de l'évaluation clinique résulte ainsi à la fois d'investigations cliniques propres et réalisées par l'industriel et de travaux bibliographiques qui sont évalués par l'organisme notifié qui statue sur la conformité ou non. Au cours de ce processus, le fabricant doit par exemple justifier de la pertinence d'une banque d'images sur laquelle il pourrait valider son logiciel en justifiant que son échantillon est conforme.

Ainsi, il revient au fabricant de prouver que son dispositif est capable, par exemple sur des radiographies, de détecter tel type de fracture. On suppose alors qu'en étant détenteur du savoir-faire de son dispositif, le fabricant doit être en mesure de procéder à l'évaluation de sa conformité.

Parmi les lacunes du contrôle de conformité : les biais engendrés par les bases de données

La question de possibles biais engendrés par les bases de données, par exemple, qui est censée être traitée lors du contrôle de conformité ne l'est que très imparfaitement. Des **dilemmes éthiques** peuvent survenir dans ce cas précis. Développons cette étape : tout algorithme d'apprentissage machine est d'abord conçu en s'appuyant sur une première

⁷¹ Selon l'article 35 du Règlement (UE) 2017/745 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux, modifiant la directive 2001/83/CE, le règlement (CE) n° 178/2002 et le règlement (CE) n° 1223/2009 et abrogeant les directives du Conseil 90/385/CEE et 93/42/CEE, consultable au lien suivant : <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32017R0745&from=FR>.

série de données, appelée base d'apprentissage. Dans le cas d'un algorithme de détection de cancer du sein par exemple, des échographies vont être soumises à un algorithme après avoir été étiquetées par un humain.

À partir de ces images, l'algorithme va dégager des régularités spécifiques aux images désignées par les humains comme étant celles où il y a présence d'un cancer, par opposition à celles n'en présentant pas. Un second jeu de données, dit « base de test » sera alors soumis à l'algorithme afin de voir s'il classe correctement les images. Lorsque les résultats obtenus sont satisfaisants, l'algorithme pourra alors éventuellement être testé sur une base de validation et ensuite être amené à traiter des données réelles. Les paramètres de traitement sont pondérés en fonction des particularités des bases de données à traiter, ce qui nécessite d'autres données spécifiques.

D'autre part, la base d'apprentissage initiale ou la nature des analyses algorithmiques successives doivent aussi être évaluées au regard de l'égalité de traitement des cas possibles (c'est-à-dire qu'elles n'introduisent pas de biais de représentation de certaines sous-populations) et être améliorées le cas échéant.

Ces différents jeux de données peuvent en effet comporter des biais de différentes sortes qu'il faut pouvoir identifier et pondérer et auxquels il faut tenter de remédier. Ces biais peuvent porter sur différentes caractéristiques. Par exemple les symptômes des crises cardiaques ne sont en moyenne pas tout à fait les mêmes pour les hommes et les femmes. Si l'une des deux catégories est moins représentée dans la base d'apprentissage, le SIADM risquera de commettre davantage de faux négatif pour cette catégorie⁷². Les différences entre catégories sociales peuvent aussi engendrer ce type de problème. Par exemple, les personnes défavorisées se trouvent en situation de sous-recours aux soins, ce qui peut être interprété par le SIADM comme un état de bonne santé⁷³. Enfin, on peut aussi observer des biais liés à l'origine des patients⁷⁴. Pour éviter et contrôler ce genre de biais, il faut accéder non seulement aux bases de données d'entraînement, mais aussi à celles de test et de validation. Si ces difficultés ne sont pas propres à l'apprentissage machine appliqué au diagnostic médical, la sensibilité des données en question, s'agissant de données de santé, rend les possibilités d'audit d'autant plus complexes. En outre, la comparaison des résultats obtenus par deux systèmes d'IA est d'autant plus pertinente qu'ils sont évalués à partir des mêmes bases de données d'entraînement ou de test – ce point milite pour la définition d'un certain nombre de bases de données fixes, partageables et analysées en double aveugle.

Enfin, ces systèmes produisent également des données. Dans le cas de l'imagerie médicale, ces données étiquetées par le SIADM sont parfois renvoyées dans le système d'archivage central (PACS pour Picture Archiving and Communication System) quand bien même les algorithmes qui les produisent n'auraient pas été certifiés. Lorsque c'est le cas,

⁷² Cirillo, Davide, Silvina Catuara-Solarz, Czuee Morey, Emre Guney, Laia Subirats, Simona Mellino, Annalisa Gigante, et al. « Sex and Gender Differences and Biases in Artificial Intelligence for Biomedicine and Healthcare ». *Npj Digital Medicine* 3, n° 1 (1er juin 2020): 1-11. <https://doi.org/10.1038/s41746-020-0288-5>.

⁷³ Gianfrancesco, Milena A., Suzanne Tamang, Jinoos Yazdany, et Gabriela Schmajuk. « Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data ». *JAMA internal medicine* 178, n° 11 (1 novembre 2018): 1544-47. <https://doi.org/10.1001/jamainternmed.2018.3763>.

⁷⁴ Sjoding, Michael W., Robert P. Dickson, Theodore J. Iwashyna, Steven E. Gay, et Thomas S. Valley. « Racial Bias in Pulse Oximetry Measurement ». *New England Journal of Medicine* 383, n° 25 (17 décembre 2020): 2477-78. <https://doi.org/10.1056/NEJMc2029240>.

se pose alors la question de l'accessibilité de ces images marquées pour tous les acteurs de la chaîne médicale via le système d'archivage central, sans contrôle préalable de la pertinence du marquage. À ce jour, la mise à disposition dans un système d'archivage d'images marquées par un système d'IA – par exemple, des images étiquetées comme présentant une fracture ou une lésion mammaire – ne fait pas l'objet d'une validation préalable par un radiologue, ce qui ne répond pas au principe de contrôle humain (voir partie 5).

En conclusion, concernant la réglementation « IA » qui est en cours de discussion à l'échelle communautaire, il apparaît que des exigences complémentaires sur les dispositifs médicaux vont venir enrichir le règlement actuel. Cela implique que, pour les fabricants d'un logiciel ayant des fonctionnalités qui répondent à la définition de l'IA donnée dans ce règlement, le dispositif devra être mis en conformité et répondre, auprès d'un organisme notifié, aux critères du marquage CE pour certaines classes de risque. *In fine*, l'aboutissement du contrôle de conformité impliquant un marquage CE pour les SIADM s'apparente à une autorisation de mise sur le marché.

3.2.3 Implications éthiques : un contrôle de conformité en cours d'élaboration

La procédure de contrôle de conformité soulève plusieurs questions éthiques :

- Le contrôle de conformité est en cours de construction et nécessite des compétences professionnelles impliquant un travail de formation et le développement de recherches académiques et scientifiques afin de parfaire sa définition.
- Les fabricants sont appelés à concevoir eux-mêmes les démonstrations de conformité de leurs dispositifs, ce qui implique que le contrôle que la démonstration n'est pas biaisée est faible, d'autant que ce contrôle est réalisé en premier lieu par un organisme notifié privé.
- La conformité d'un SIADM ne signifie en rien qu'il y a un bénéfice clinique à son usage. Peut-être fait-il très bien ce qu'il prétend pouvoir faire, mais rien n'assure que ce qu'il prétend pouvoir faire est utile ou bénéfique pour le patient ou pour le système de soins.

Recommandations :

9) Encourager la **recherche pluridisciplinaire et la formation professionnelle** pour mettre au point une définition contrôle de conformité des SIADM dénuée de biais.

10) Promouvoir l'utilisation d'essais cliniques pour **évaluer le rapport bénéfice-risque** des SIADM de la même façon que cela se fait pour les médicaments.

3.3 Régime de responsabilité

La Commission européenne a publié, le 28 septembre 2022, une proposition de directive portant sur l'établissement d'un régime harmonisé d'indemnisation au niveau communautaire des dommages susceptibles d'être causés par des systèmes d'IA. Cette directive renforcerait très sensiblement la portée de l'obligation de contrôle humain pour les systèmes d'IA à haut risque, ce qui s'applique au champ particulier de la santé. L'absence de système établi de contrôle humain permettrait au patient de s'appuyer sur

une présomption réfragable du lien de causalité entre la faute du fabricant et son dommage, lui facilitant ainsi l'obtention d'une réparation.

Dans ces conditions, les concepteurs et utilisateurs de solutions d'IA seraient fortement incités au déploiement de dispositifs de supervision comme les « collèges de garantie humaine » recommandés par l'avis n° 129 du CCNE et par l'OMS, et qui associent concepteurs d'algorithmes, représentants des professionnels et représentants des bénéficiaires finaux du système d'IA.

Le dispositif du projet de règlement européen sur l'IA prévoit déjà un mécanisme de sanction de l'absence de contrôle humain dans le cadre d'un régime de sanctions financières – jusqu'à 4 % du chiffre annuel mondial du groupe – décidés par les autorités de régulation sur le modèle du RGPD. La directive viendrait ainsi compléter l'opposabilité du contrôle humain en ouvrant aux patients en santé et plus largement aux bénéficiaires finaux de l'IA une capacité à agir en responsabilité civile pour, en cas de dommage, s'appuyer sur l'absence ou l'insuffisance du contrôle humain pour obtenir réparation.

3.4 Pour un contrôle de conformité et une évaluation des SIADM plaçant la qualité du soin avant toute autre considération

Comme nous l'avons vu plus haut, en amont de leur introduction sur le marché de la santé, et donc de leur utilisation en situation réelle, les SIADM sont soumis à un contrôle de conformité (marquage « CE ») et d'évaluation (portant sur l'efficacité clinique) au cours duquel les questions éthiques sont appelées à occuper une place particulière.

En matière de **contrôle de conformité** (voir 3.2), les systèmes européen (Union européenne) et américain (*Food and drug administration*, FDA) diffèrent, mais aucune comparaison précise des deux systèmes n'est disponible. Il serait utile que des recherches soient développées sur ce sujet.

En l'état actuel de la situation en Europe, il est souhaitable qu'un certain degré d'explicabilité soit atteint lors de la présentation du SIADM à l'institution de contrôle (Agence nationale de sécurité du médicament (ANSM) pour le cas français) afin que cette dernière puisse s'assurer que les utilisateurs finaux du SIADM aient accès à la meilleure compréhension possible de son fonctionnement.

De plus, le contrôle doit spécifier le degré de performance diagnostique atteint par le SIADM. Ce degré de performance est souvent issu d'une comparaison entre un diagnostic fait par un praticien ou de manière collégiale et les résultats calculés par le SIADM. Ce degré de performance peut ensuite s'exprimer de très diverses manières : sensibilité, spécificité, stabilité, robustesse, etc. Cependant, la définition d'une référence opposable et très largement admise (ou *gold standard*) du contrôle est délicate pour les SIADM, en particulier en raison de la complexité et de la diversité des données d'apprentissage. Pour autant, il est essentiel de tendre à stabiliser les étapes et de rechercher, autant que possible, des critères stables et opérants du contrôle.

Concernant l'**évaluation**, elle est conduite par la Haute autorité de santé (HAS) dans le cas français. Or celle-ci ne s'intéresse qu'aux SIADM susceptibles d'être remboursés par la sécurité sociale. La grande quantité de SIADM utilisés par les soignants – la plupart de ceux que nous avons présentés dans cet avis – lui échappent donc puisqu'ils ne sont pas

vendus aux patients. La question se pose de savoir s'il faudrait élargir les compétences de la HAS, ce qui impliquerait bien sûr de lui en fournir les moyens.

D'autre part, l'évaluation doit prendre en compte les propriétés des bases de données qui doivent aussi être adaptées, quantitativement et qualitativement, à la prévalence de la maladie étudiée. En effet, ces bases de données peuvent masquer ou même engendrer des biais de différentes natures : biais de représentativité de la population, biais d'activité (par ex. : orienté vers les soins les mieux valorisés par la structure de soins) ou encore biais organisationnels (ex. : on collecte les données faciles à recueillir mais non nécessairement utiles au soin) (voir l'encart au 3.2.2). Il est donc nécessaire que ces bases soient elles-mêmes évaluées.

En outre, lorsqu'ils sont soumis à évaluation, les SIADM sont accompagnés de promesses faites par leurs promoteurs (voir 3.2.2.). Certes, ces promesses répondent aux contraintes financières d'investissement dans les projets, mais elles peuvent être exagérées au regard des bénéfices réellement apportés aux patients. En outre, ce décalage pose un problème spécifique d'information qui devrait être pointé par les autorités d'évaluation.

Enfin, les SIADM s'insèrent dans les relations interprofessionnelles du système de soins qu'ils transforment. Il est donc essentiel que l'institution évaluatrice dispose d'informations de nature socio-économique sur ces transformations de façon à anticiper les changements organisationnels qui résultent de leur utilisation. Cette remarque prend une ampleur plus grande encore si l'on considère que les SIADM sont des produits qui peuvent être délocalisés et être utilisés dans des pays sous-connectés pour lesquels ils ne seront pas adaptés. Il faut en effet prendre en considération l'un des angles morts cruciaux de l'IA, à savoir que les bases de données des systèmes d'IA proviennent principalement des pays occidentaux largement numérisés, excluant de fait des données de santé concernant les personnes habitant dans des régions sous-dotées en infrastructures de recueil de données de santé comme l'Afrique et l'Amérique latine, ce qui peut engendrer des biais. Il est alors fort probable que des implicites sociaux enchâssés dans un SIADM ne correspondent pas aux caractéristiques de la société où il est utilisé. Par exemple, un quart des hôpitaux africains⁷⁵ ne dispose pas d'inventaire de leur matériel informatique, ce qui restreint les capacités d'agrégation des données (puisqu'il est alors difficile de savoir où elles peuvent être produites et stockées) et limite drastiquement le recours à des SIADM.

Très généralement, l'évaluation telle qu'elle devra être pratiquée, aura pour fonction de vérifier que la rationalité sanitaire non seulement n'est pas occultée par la rationalité algorithmique, mais aussi qu'elle n'est pas remplacée par des rationalités économiques d'une part, et organisationnelles de l'autre. L'évaluation devra avoir en permanence le principe de bienfaisance.

L'un des critères d'évaluation des SIADM sur lequel la France s'est démarquée par sa réactivité⁷⁶, notamment avec la loi relative à la bioéthique du 2 août 2021, réside dans ce

⁷⁵ Sagbo J-Y., (2007), « Etude sur les contraintes de gestion des équipements biomédicaux en Afrique », 64 p.

⁷⁶ Dans la loi relative à la bioéthique du 2 août 2021, les dispositions concernées figurent à l'article 17. Les mots « Garantie Humaine » ont été proposés puis ont disparu de la loi elle-même au cours des navettes parlementaires mais sont, en revanche, très directement repris dans l'exposé des motifs du projet de loi [1] et dans son étude d'impact [2]. Ce positionnement correspond à la distinction que l'on peut effectuer entre

que la Commission européenne a traduit par l'expression de « **contrôle humain** » ou *Human Oversight*⁷⁷ au titre de l'article 14 de son projet de réglementation commune sur l'IA. Après de très nombreuses hésitations et discussions au sein du CCNE comme du CNPEN, qui prouvent que le concept n'est pas encore fermement établi, nous avons retenu l'expression de « contrôle humain » car c'est la traduction du concept de *human oversight* qui a été retenu par la Commission européenne dans le texte cité ci-dessus. Le paragraphe 1 de cet article énonce ainsi que les systèmes d'intelligence artificielle doivent être conçus et développés de façon à pouvoir être contrôlés par des humains. Le paragraphe suivant précise que le contrôle humain permettra de prévenir ou de minimiser les risques pour la santé, la sécurité ou les droits fondamentaux pouvant émerger d'un système d'IA susceptible de présenter un niveau de risque élevé. Le paragraphe 3 donne des indications sur la mise en application du contrôle humain des SIADM en préconisant, lorsque cela est techniquement possible, des mesures identifiées et si possible intégrées au système d'IA par le fournisseur avant la mise sur le marché ou la mise en service de ce dernier et qui se prêtent à une mise en œuvre par l'utilisateur.

Le principe de contrôle humain se situe dans le prolongement des recommandations émises par le CCNE dans son avis n° 129 à propos de ce qu'il qualifiait de « supervision humaine »⁷⁸. Dans son avis n° 130, le CCNE avait également précisé le contenu de ce principe s'agissant plus spécifiquement des traitements de données en précisant « qu'une « garantie humaine » est essentielle pour répondre de la rigueur méthodologique des différentes étapes du processus de recueil et de traitement des données que sont : (i) la qualité et l'adéquation des données sélectionnées pour entraîner les algorithmes ; (ii) l'adéquation du choix des traitements algorithmiques à la question posée ; (iii) la vérification sur un jeu de données indépendantes de la robustesse et de l'exactitude du résultat donné par l'algorithme. »

Le CNPEN, de son côté, dans sa contribution en réponse au Livre blanc de la Commission européenne sur l'intelligence artificielle⁷⁹, avait repris ces recommandations et précisé, s'agissant de la mise en œuvre du principe, les attendus des « collègues de garantie humaine » en soulignant l'intérêt d'associer innovateurs en IA, représentants des professionnels de santé et représentants des patients à l'exercice de ce contrôle.

un principe éthique général et ses concrétisations opérationnelles en droit positif qui se retrouvent, quant à elles, directement dans le corps des dispositions législatives.

[1] L'article « vise à sécuriser la bonne information du patient lorsqu'un traitement algorithmique de données massives (« intelligence artificielle ») est utilisé à l'occasion d'un acte de soins. Il décline également la garantie d'une intervention humaine. »

[2] <https://www.legifrance.gouv.fr/dossierlegislatif/JORFDOLE000038811571/>.

⁷⁷ Voir : Commission européenne, « Proposition de règlement du parlement européen et du conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'union », Bruxelles, 21 avril 2021 [consultable au lien suivant : <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>].

⁷⁸ CCNE, avis n° 129 du 18 septembre 2018, *Contribution du Comité consultatif national d'éthique à la révision de la loi de bioéthique 2018-2019*, 165 p.

⁷⁹ Commission européenne, (2020), *Livre blanc. Intelligence artificielle. Une approche européenne axée sur l'excellence et la confiance*, 31 p.

Recommandations :

11) Veiller à limiter les biais potentiels liés à la programmation même des algorithmes et aux différentes bases de données utilisées pour la mise au point des SIADM, dont les données d'apprentissage sont spécifiques à leur date de conception.

12) Promouvoir la définition, par les industriels, de critères qui permettront d'évaluer un SIADM dès sa conception (« **Ethics by design** »). Ceux-ci-peuvent être génériques ou spécifiques à un SIADM.

13) Établir un certain nombre de **critères sociaux d'évaluation** des SIADM (critères médicaux pour le patient, organisationnels pour les établissements et professionnels de santé, sanitaires pour la population, économiques pour l'ensemble des acteurs, etc.) permettant d'en dresser un bilan éthique.

VI. ENJEUX ETHIQUES POSES PAR LES SIADM DANS LA TRANSFORMATION DE LA DEMARCHE DIAGNOSTIQUE

Dans cette partie consacrée aux mutations induites par les SIADM dans la démarche diagnostique, nous proposons de décrire les enjeux qui nous sont apparus comme les plus saillants au regard des données actuelles à la fois en termes de potentialités, d'état d'avancement des procédés techniques mais aussi en termes d'attentes et de besoins sociaux.

4.1 Enjeux de la radiomique

Outre leur apport en matière d'aide au diagnostic dans la détection de lésions, les algorithmes appliqués à l'analyse d'images peuvent aussi extraire, à partir des données d'imagerie médicale (échographie, scanner ou IRM), un grand nombre de paramètres quantitatifs tissulaires, inaccessibles à l'œil humain dont la signification en terme de pronostic et de réponse au traitement notamment dans les tumeurs fait actuellement l'objet de recherches prometteuses.

L'image d'une tumeur est un « phénotype », c'est-à-dire une « apparence » correspondant à de nombreux paramètres tissulaires hétérogènes. Cette analyse informatique paramétrique dite « radiomique » (terme construit sur le modèle de la génomique pour le génome et de la protéomique pour le protéome) peut à terme permettre de trouver des biomarqueurs capables de prédire le pronostic d'une tumeur, (risque de métastases notamment), et donc d'adapter le traitement en fonction puis de suivre finement la réponse thérapeutique. Parallèlement, la radiomique vise à étudier les corrélations paramétriques avec le profil génétique d'une tumeur.

Ainsi, la radiomique vise *in fine* à développer une réponse thérapeutique individuellement et précisément adaptée à chaque tumeur grâce à la connaissance de sa « carte d'identité » paramétrique (radiomique) et génétique (génomique).

La majorité des études dans ce champ sont encore au stade de l'expérimentation, la plupart des biomarqueurs potentiels doivent être validés mais le potentiel d'application de la radiomique est important, avec pour prérequis la nécessité d'organiser et de traiter des grandes bases de données d'imagerie en équipes pluridisciplinaires (médicales, bio informatiques, bio statistiques etc.). Corrélativement, l'ouverture de tels champs suppose des défis pour l'interprétation de ces données massives : d'une part, la suraccumulation peut saturer le professionnel de santé et, d'autre part, une possible surinterprétation si le champ d'utilisation de ces nouvelles données n'est pas parfaitement maîtrisé. À ce titre, nous attirons l'attention sur l'importance de la contextualisation de ces nouvelles informations par le médecin référent.

4.2 Enjeux dans l'interprétation des données génomiques

Depuis peu, les chercheurs envisagent la mesure à large échelle de multiples autres constituants de l'organisme humain grâce à des techniques dites « omiques », mais aussi

environnementales (alimentation, pollution, climat, etc.) autour du concept d'exposome⁸⁰, reconnu comme enjeu majeur de santé publique. Alors qu'autrefois, face à un phénotype donné (symptomatologie de maladie), le séquençage d'une région ciblée du génome venait confirmer l'existence de mutations, la démarche opposée est aujourd'hui appliquée : on sait que la survenue d'une maladie repose à la fois sur des facteurs génétiques et environnementaux, au sens d'exposome. Si l'on dispose de mesures fiables de ces éléments, il devient possible non seulement de faire un diagnostic précoce, mais aussi de proposer un « score » corrélé à un risque de développer cette maladie⁸¹.

4.2.1 Du diagnostic du symptôme au diagnostic du risque

Dans le domaine de la génomique, une minorité de maladies dites « monogéniques » parce qu'elles ne concernent qu'un seul gène, sont causalement liées à des mutations génétiques rares. Mais la plupart des maladies chroniques sont des maladies complexes dites polygéniques, c'est-à-dire impliquant plusieurs variants⁸² sur plusieurs gènes identifiés. Des méthodes statistiques qui agrègent le risque cumulé par la présence de ces variants donnent un « score polygénique » pour une maladie donnée et pour un individu donné. Pour certaines maladies, un score polygénique élevé est corrélé à un risque aussi significatif de développer la maladie que la présence d'une mutation rare⁸³. Ces variations ne s'exprimeront en « maladie symptomatique » que sous la pression d'autres facteurs, notamment environnementaux. Ces données reposent toutes sur l'analyse par des technologies d'IA de différents types de données massives (génomiques, cliniques, environnementales) et la détermination de corrélations.

L'élaboration d'un diagnostic pré-clinique d'une pathologie, et l'anticipation de l'apparition de maladies en mesurant leur « risque » de survenue pour éviter leur développement font partie des défis éthiques posés par l'essor des modèles prédictifs en santé en matière d'information des personnes potentiellement concernées. Ce « score polygénique » peut représenter un marqueur clinique majeur et il pourrait être envisagé que l'analyse du génome fasse partie, dans quelques années, du parcours de soins de tout individu, voire de stratégies de dépistages préventifs. Cet accent mis sur l'évaluation précoce du risque est ce qui fonde la constitution de grandes banques génomiques, souvent soutenues par les États et les institutions publiques⁸⁴. Cette démarche initialement pratiquée en oncologie, et aujourd'hui appliquée à d'autres maladies ou déficiences, est en pleine expansion. Il faut toutefois garder à l'esprit que l'analyse algorithmique du génome, si elle permet un diagnostic ou une évaluation du risque plus précis, utilise une telle quantité d'informations qu'il est difficile pour le généticien de comprendre comment la prédiction a

⁸⁰ Vermeulen R., et al., (2020), "The exposome and health: Where chemistry meets biology", *Science*, 367: 392-396.

⁸¹ Voir : Steinfeldt J. et al., (2022), "Neural network-based integration of polygenic and clinical information: development and validation of a prediction model for 10-year risk of major adverse cardiac events in the UK Biobank cohort", *The Lancet Digital Health*, 4, 2, e84-e94.

⁸² On appelle variant une modification de la séquence d'un ou plusieurs acides nucléiques de l'ADN, par rapport à une séquence dite de référence.

⁸³ Khera A.V., et al., (2018), "Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations", *Nature Genetics*, 50 : 1219-24.

⁸⁴ Citons la UK Biobank (500 000 personnes), mais aussi aux États-Unis le *100K Wellness Project* et le *All of Us Research Program* (un million d'individus), en Asie la biobanque Kadoorie financée par le Wellcome Trust-funded China (515 000 chinois), et d'autres encore. Au Royaume-Uni, le NHS (*National Health Service*) et le ministère de la Santé encouragent aussi les individus à payer pour faire l'analyse de leur génome (Genomics England) avec des garanties de fiabilité. Ces données sont mises à disposition des chercheurs depuis 2017.

été faite. Cet élément renvoie au questionnement de l'explicabilité des solutions proposées par l'intelligence artificielle abordé en section 3.1.2. Il soulève également des questions éthiques délicates, en particulier au sujet de la gestion des données incidentes (voir la section 2.1.8. à propos de l'incidentalome) thème abordé par le CCNE dans un avis antérieur (voir avis n° 130 du CCNE).

Les algorithmes d'IA ont également permis, notamment en oncologie, de mettre en évidence des corrélations entre la survenue de maladies et d'autres potentiels marqueurs (microARN, métabolites circulants, protéines, fragments d'ADN génomique circulants) mesurables chez les individus par une simple prise de sang. De ces corrélations sont déduits des biomarqueurs, dont l'intégration avec le score polygénique et une « surveillance » par capteurs, pourrait augmenter la précision d'un diagnostic précoce de certaines maladies, mais aussi prédire leur évolution. Mais au-delà de la prouesse technique⁸⁵, demeure la question de ce que l'individu fait de cette information, ce qui soulève de sérieuses questions éthiques sur lesquelles nous reviendrons (voir partie 4.3).

4.2.2 Implications éthiques : explicabilité, explication et importance du conseil en génétique

La diffusion des technologies d'IA, la banalisation de l'accès à des données génétiques et l'accélération de l'identification de facteurs de risques induit un besoin croissant d'accompagnement médical et d'explication des diagnostics qui s'y rapportent auprès des patients et de leurs familles. Cette accélération de l'identification de facteurs de risques induit une modification potentielle de la perception de ce qu'est un « état sain ». Cette détection précoce du risque est une question majeure d'un point de vue éthique, tant à l'échelle de la société que du point de vue individuel, mais aussi notamment du point de vue de l'autonomie du patient et de la gestion de l'anxiété.

En outre, comme développé plus tôt dans cet avis (section 3.1.2), le besoin d'une explicabilité des SIADM pour les professionnels est significatif en génomique au regard de la diffusion de plus en plus rapide des traitements algorithmiques dans la discipline. Des tâches comme l'établissement de la significativité clinique des variants génomiques leur sont maintenant dévolues et impliquent un grand nombre d'hypothèses qui doivent être explicitées à l'utilisateur. Pour les patients, le développement de la profession de « conseiller en génétique » répondra peut-être en partie à ce point en particulier si ses compétences sont étendues à l'appréhension de la complexité des SIADM.

L'interprétation des données génomiques, en particulier résultant du séquençage haut débit du génome entier ou de l'exome, pose de façon particulièrement saillante la question des découvertes fortuites, ou incidentalomes. Si ce sujet n'est pas spécifique à l'intelligence artificielle et aux données génomiques, l'accélération des traitements de données génomiques grâce à l'IA provoque un changement d'échelle de ces enjeux particulièrement sensibles. Il se peut ainsi qu'un SIADM opérant une lecture systématique du génome dans le cadre d'une recherche clinique identifie des facteurs de risques réels sur des brins du génome *a priori* sans rapport avec les symptômes pour lesquels le patient consulte.

⁸⁵ Par exemple de telles analyses ont pu montrer que des anomalies étaient détectables plus de 10 ans avant les manifestations cliniques (cancer, Alzheimer), ou encore que la date de décès pouvait être anticipée.

Point de vigilance :

7) Une attention particulière doit être portée **aux risques de surveillance et de sélection adverse en matière assurantielle et de santé publique** (cet élément fait l'objet d'un développement en partie 5).

4.3 Enjeux concernant l'émergence de nouveaux espaces de démocratie en santé

La transformation numérique du secteur de la santé en général, et la diffusion de l'intelligence artificielle dans le diagnostic médical en particulier, vont nécessairement modifier la relation entre les professionnels de santé et leurs patients. L'article 17 de la loi de bioéthique introduit à cet effet un devoir d'information du patient sur le recours à un traitement algorithmique dans son protocole de prise en charge⁸⁶. Par ailleurs, sur un plan plus collectif, la diffusion de systèmes d'IA en santé peut aussi être l'occasion d'ouvrir des espaces de participation aux patients et à leurs représentants dans la réflexion sur l'éthique et la régulation de l'usage de l'intelligence artificielle dans le diagnostic médical. En complément, de nouveaux droits pourraient être énoncés dans le cadre d'une législation relative à la démocratie en santé, c'est-à-dire à la participation du patient dans la décision médicale le concernant, dans la filiation de la loi n° 2002-303 du 4 mars 2002 relative aux droits des malades et à la qualité du système de santé.

4.3.1 L'émergence d'un nouveau statut : l'auxiliaire numérique

Un nouveau statut pourrait être promu par le système de soins, celui « d'auxiliaire numérique » ou « aidant numérique » qui aiderait les patients à mieux comprendre les enjeux des SIADM auxquels ils pourraient être confrontés. Il s'agirait d'un auxiliaire dont la fonction serait d'éviter que la difficulté d'accès au numérique ne vienne se surajouter à une difficulté d'accès aux soins. Il conviendrait aussi d'établir une modalité de certification des compétences associées à ce rôle.

La personne l'exerçant pourrait être choisie dans l'entourage du patient. Mais si la solution d'une personne de l'entourage n'est pas envisageable pour le patient, il faudrait pouvoir lui en proposer une : des membres d'associations de patients pourraient alors intervenir. Ce rôle nouveau pourrait aussi être attribué à la « personne de confiance », instituée par la loi du 4 mars 2002, afin de permettre notamment à un patient de bénéficier de son appui dans l'expression d'un consentement plus complexe du fait de la multiplicité des circonstances. Cette fonction pourrait enfin être attribuée à un représentant d'associations de patients compétentes dans la communication d'informations relatives au traitement algorithmique des données. À cet égard, l'effectivité de cette nouvelle attribution pour les représentants de patients implique l'engagement d'un large programme de formation aux enjeux de l'intelligence artificielle en santé au sein des associations. L'organisme de

⁸⁶ L'article mentionne ainsi : « Le professionnel de santé qui décide d'utiliser, pour un acte de prévention, de diagnostic ou de soin, un dispositif médical comportant un traitement de données algorithmique dont l'apprentissage a été réalisé à partir de données massives s'assure que la personne concernée en a été informée et qu'elle est, le cas échéant, avertie de l'interprétation qui en résulte ».

représentation des patients et usagers du système de santé, France Assos Santé, s'est d'ores et déjà engagé dans cette direction⁸⁷.

Le contexte de crise sanitaire a montré une diffusion de plus en plus large des outils numériques tout en révélant la persistance de situations pour lesquelles le recours à ces outils peut s'avérer problématique (manque d'acculturation au numérique, défaillance d'infrastructures réseaux, conditionnalité économique et sociale pour l'accès au numérique). Ainsi, en se développant de plus en plus largement, l'IA de diagnostic médical rend encore plus impérieuse la résolution de la fracture numérique et de l'illectronisme, dont l'Insee considérait en 2019 qu'il concernait 17 % de la population française⁸⁸.

4.3.2 Implications éthiques : promouvoir les initiatives de démocratie en santé

La faible acculturation de la population et des professionnels de santé au champ du numérique en santé et particulièrement à l'intelligence artificielle implique que la diffusion de SIADM comporte un risque intrinsèque d'exclusion de patients, de leurs auxiliaires et des professionnels de santé. Le danger du développement et de l'accessibilité des dispositifs médicaux intégrant de l'IA est celui d'un approfondissement de la distance entre les professionnels et les techniciens en IA et en santé d'une part, et les patients d'autre part, ou encore entre les médecins connaissant le numérique et ceux qui lui restent étrangers.

Pour autant, nous constatons, à ce stade, un mouvement inverse d'émergence de nouveaux espaces de démocratie en santé et de consolidation de la participation des patients à un niveau individuel mais aussi sur un plan collectif. Cette dynamique n'est bien entendu pas acquise par principe et elle doit faire l'objet de soutiens, en particulier pour consolider l'effort de formation des représentants des patients aux enjeux de l'intelligence artificielle en santé.

Recommandation :

14) Promouvoir le **statut d'auxiliaire numérique**. Cette personne fournirait une raison compréhensible par le patient, ou son représentant légal, ou sa personne de confiance, du résultat produit par le SIADM qui soit également représentative de son fonctionnement. Il pourrait être rapproché du statut « d'infirmier de pratique avancé » en développement actuellement.

⁸⁷ France Assos santé, (2019), *Livre blanc. Contribution des outils numériques à la transformation des organisations de santé*, 189 p.

⁸⁸ Vie-publique.fr, « Fracture numérique : l'illectronisme touche 17% de la population selon l'INSEE », 13 novembre 2019 [consultable au lien suivant : <https://www.vie-publique.fr/en-bref/271657-fracture-numerique-lillelectronisme-touche-17-de-la-population>].

4.4 Les SIADM ne doivent pas dégrader les qualités du cadre diagnostique

Le SIADM doit rester une **aide à la décision humaine**. Le SIADM peut et doit être ignoré délibérément par un praticien s'il le juge nécessaire⁸⁹. Cette question de la préservation de la prudence dans le recours à l'IA constitue l'une des justifications du contrôle humain mais elle induit également un certain nombre de corollaires dans le positionnement du professionnel vis-à-vis de l'outil d'IA.

Lors d'un parcours de soins, il faut s'assurer que le patient a bien été informé que l'équipe médicale peut avoir recours à un SIADM, pour quelle raison, et quels en sont les bénéfices et les risques. C'est d'ailleurs le sens de l'article 17 de la loi de bioéthique qui introduit un devoir d'information du patient par le professionnel sur le recours à un traitement algorithmique, cette information étant mise en œuvre à l'occasion du recueil du consentement sur le protocole de soins dans son ensemble.

Les relations entre les équipes médicales et le patient ne doivent pas pâtir de l'usage d'un SIADM par les soignants ou par le patient⁹⁰. Si ce dernier a utilisé un SIADM hors parcours de soins, il ne peut se contenter d'une information non professionnelle pour prendre en charge sa pathologie. Il faut d'abord qu'il puisse comprendre les limites des informations qui sont produites hors cadre médical et ensuite, s'il suspecte un problème à partir de ces informations, que le SIADM le dirige vers un professionnel de santé. Cette question se pose particulièrement pour les systèmes de confort.

L'utilisation d'un SIADM au cours du processus diagnostique engendre un certain nombre de conséquences, qui elles aussi soulèvent des questions éthiques. Les données et informations produites par le SIADM, particulièrement si elles ne sont pas attendues ou recherchées dans le cadre du processus diagnostique initial, peuvent inciter l'équipe médicale à prendre des décisions qui ne s'imposent pas nécessairement (cas de l'incidentalome). L'équipe médicale doit se garder d'ordonner des interventions sans avoir vérifié la réalité du risque identifié et demandé son consentement éclairé au patient.

Pour mieux appréhender ces transformations structurelles associées à la diffusion des technologies d'IA appliquées au diagnostic médical, une nouvelle approche de la formation des professionnels de santé est impérative. Les impacts de la diffusion de l'intelligence artificielle pour les métiers de la santé sont manifestes voire disruptifs⁹¹. Ces technologies innovantes ne se substituent pas nécessairement à l'intervention humaine mais peuvent augmenter les capacités du médecin (analyse, diagnostic, prescription). Pourtant, elles questionnent la **responsabilité de la décision**. En effet, si le SIADM a une très bonne précision, un risque de confiance exagérée de la part du praticien peut survenir. À l'inverse, si le praticien a davantage confiance dans son propre jugement, c'est un risque de sous-utilisation de la technologie qui apparaît. De ce point de vue, la formation au débat collégial autour de l'IA, à travers, le cas échéant, du vecteur classique des réunions de concertation pluridisciplinaires (RCP) où différentes disciplines médicales se rencontrent, constitue un

⁸⁹ Cela figure dans le projet de réglementation européenne pour l'IA : (d) d'être en mesure de décider, dans une situation particulière, de ne pas utiliser le système d'IA à haut risque ou de négliger, passer outre ou inverser le résultat fourni par ce système.

⁹⁰ B. Mittelstadt, *op. cit.*

⁹¹ Voir notamment l'étude de l'Institut Montaigne sur les impacts RH de l'IA et de la robotisation en santé : <https://www.institutmontaigne.org/publications/ia-et-emploi-en-sante-quoi-de-neuf-docteur>

enjeu essentiel pour l'appropriation et le contrôle de ces outils. Il conviendra alors d'intégrer dans tout ou partie des RCP d'autres corps de métiers (ingénieurs, informaticiens) et des usagers du système de soins⁹².

Les conditions d'exercice des spécialités médicales et des fonctions support aux soins pourraient durablement changer, même si les impacts de l'utilisation des SIADM diffèrent selon les spécialités. L'impact peut être important pour les disciplines médicales déjà très requérantes en numérique comme la radiologie, l'histopathologie, et l'ophtalmologie. En conséquence, ces spécialités sont naturellement les premières à être transformées par l'IA mais sans pour autant qu'on puisse appréhender, à ce stade, toutes les implications concrètes de cette plus large diffusion. Le développement des SIADM intervient dans un contexte de raréfaction de ces spécialistes. Une conséquence probable est que l'activité de ces professionnels se concentrera sur l'analyse de situations diagnostiques complexes, ou le développement d'actes à plus haute valeur ajoutée, les SIADM étant utilisés pour effectuer les tâches diagnostiques simples ou de routine, sous un contrôle humain, dont les modalités précises devront être précisées. À court et moyen terme, si les technologies d'IA ne menacent pas directement d'entraîner la disparition des médecins, la profession subira des évolutions et leurs effets seront plus manifestes encore sur les fonctions support aux soins, comme cela a été observé dans d'autres secteurs où plusieurs systèmes d'IA en matière de gestion des tâches administratives ou logistiques existent déjà.

Recommandations :

15) Veiller à ce que les concepteurs de SIADM indiquent clairement les performances et les limites de leurs SIADM et les expliquent en termes compréhensibles pour leurs utilisateurs.

16) Adapter les cursus d'enseignement (formation initiale et continue) pour les professions médicales et paramédicales pour qu'ils forment aux technologies d'IA en intégrant la prise en compte de leurs enjeux éthiques.

⁹²HAS, « Réunion de concertation disciplinaire », Novembre 2017, 3 p.

CONCLUSION

Les SIADM sont pourvoyeurs de nombreux progrès. Ils permettent d'identifier des lésions qui échappent à l'œil humain. Ils permettent de traiter très rapidement des cas assez aisés à diagnostiquer mais répétitifs, et libèrent ainsi du temps médical pour échanger avec le patient. Certains SIADM donnent aussi la possibilité de repenser l'organisation des soins, particulièrement à l'hôpital, et permettent ainsi d'améliorer les phases d'accueil des patients.

Mais un certain nombre d'enjeux éthiques associés aux technologies d'IA appliquées au diagnostic médical ont été présentés dans les quatre parties de ce document, en regard des différentes situations d'usages possibles de ces technologies. Deux principaux axes de tensions éthiques sont apparus.

D'une part, on distingue les tensions qui concernent le **respect de la personne**, qui doit être respectée avec ses particularités et dans son contexte singulier et qui, en particulier, ne peut être transformée en simple « consommatrice de soins » agissant dans le cadre d'un marché. Le patient est d'une autre nature que le consommateur.

D'autre part, certaines tensions sont relatives à la **nécessité d'un traitement équitable de toutes les personnes** concernées de façon à ne pas créer de discriminations. On reconnaît un enjeu fondamental du diagnostic en intelligence artificielle qui fait écho à une tension constante en santé publique entre l'individuel et le collectif : comment traiter chaque personne envisagée dans une situation spécifique tout en traitant toutes les personnes équitablement ? C'est à cet exercice difficile que l'introduction des SIADM dans le diagnostic médical doit permettre de répondre, associant ainsi la pensée du singulier et la pensée algorithmique, telles que décrites en première partie de cet avis.

Pour finir, le CCNE et le CNPEN attirent l'attention sur le fait que de très nombreux développements de l'IA dans le domaine de la prévention sanitaire sont en train d'émerger. Il serait de toute première importance de compléter cet avis portant sur le diagnostic par une réflexion qui couvrirait cette question tout à fait complémentaire.

ANNEXES

Annexe 1 : Membres du groupe de travail

Catherine Adamsbaum – Invitée au titre du CCNE
Gilles Adda (CCNE)
Erik Boucher de Crèvecœur – Invité au titre du CNPEN
Raja Chatila (CNPEN)
Laure Coulombel (CCNE)
Laurence Devillers (CNPEN)
Emmanuel Didier (rapporteur CCNE)
Karine Dognin-Sauze (CNPEN)
Valeria Faure-Muntian (CNPEN)
Christine Froidevaux (CNPEN)
Jean-Gabriel Ganascia (CNPEN)
David Gruson (rapporteur CNPEN)
Emmanuel Hirsch (CNPEN)
Claude Kirchner (CNPEN)

Camille Darche (rédactrice CNPEN)
Lucie Guimier (rédactrice CCNE)
Marguerite Schweizer (stagiaire CCNE)

Annexe 2 : Liste des personnes auditionnées

Isabelle Adenot, présidente de la Commission Nationale d'Évaluation des Dispositifs Médicaux et des Technologies de Santé (CNEDiMETS) ;

Hélène Bruyere, cheffe d'équipe, évaluateur contrôle de conformité (ANSM) ;

Guillaume Chassagnon, Maître de conférences des universités-praticien hospitalier, Unité d'imagerie thoracique, APHP, Centre-Université de Paris - Hôpital Cochin ;

Corinne Collignon, Cheffe de service de la Mission numérique en santé à la Haute Autorité de Santé (HAS) ;

Stéphanie Combes, directrice du *Health Data Hub* ;

Frédéric Dittenit, délégué à la protection des données (ANSM) ;

Antoine Flahault, épidémiologiste et professeur de santé publique à l'université de Genève ;

Virginie Gaiffe, cheffe de l'équipe produits diagnostic, des systèmes radiogènes et des systèmes d'information (ANSM) ;

Dominique Leguludec, présidente du collège de la HAS ;

Thierry Sirdey, chef de l'unité évaluation et contrôle du marché à la direction de l'évaluation des dispositifs médicaux (ANSM).

Annexe 3 : Recommandations et points de vigilance de l'avis

Points de vigilance :

- 1) Pour éviter les risques d'incomplétude et d'incidentalome, il est capital que la signification d'une mesure qui évalue le compromis entre précision et sensibilité soit parfaitement clarifiée par les industriels.
- 2) Un SIADM, s'il peut rassurer par son fonctionnement rigoureux et automatique, ne plonge pas moins le patient comme l'équipe soignante dans un certain degré d'incertitude. Le maintien d'un contrôle humain lors de l'utilisation d'un SIADM apparaît comme indispensable mais ne lèvera pas forcément les incertitudes.
- 3) Les SIADM doivent être envisagés en tant qu'outils complémentaires des réponses à apporter aux insuffisances du système de soins, en particulier de la désertification médicale, mais ne doivent pas être considérés comme des solutions de substitution des équipes médicales.
- 4) Le processus décisionnel conduisant à l'utilisation de SIADM doit impérativement prendre en compte le souci du soin avant les considérations économiques. Il doit être considéré comme un moyen de lutte contre les inégalités sociales et territoriales de santé.
- 5) Maintenir une vigilance contre les risques de sélection adverse qui pourraient être opérés par les assurances complémentaires en particulier au moyen d'outils de diagnostic actuariels fondés sur des techniques d'IA.
- 6) Prendre en compte le risque de dérives d'avis non professionnels, de traitement non sécurisé des données, et d'exercice illégal de la médecine accentuant la vulnérabilité des personnes. Les SIADM doivent relever de pratiques médicales strictes.
- 7) Une attention particulière doit être portée aux risques de surveillance et de sélection adverse en matière assurantielle et de santé publique.

Recommandations :

- 1) Il est essentiel que les méthodes diagnostiques déjà établies, ne faisant pas intervenir a priori de SIADM, continuent d'être enseignées et de faire l'objet de recherches visant à les faire progresser.
- 2) La place grandissante des SIADM dans le champ des compétences médicales nécessite des études approfondies sur l'interaction entre les humains et les technologies d'intelligence artificielle pour évaluer l'impact des SIADM dans l'exercice de la médecine.
- 3) Dans un souci de transparence et de traçage, il faut indiquer l'utilisation d'un SIADM dans le compte rendu médical d'une consultation.
- 4) Ces éléments convergent en faveur d'un contrôle humain à toutes les étapes du soin, de l'indication des examens aux résultats des analyses et à l'interprétation contextuelle de ces résultats.
- 5) Engager une concertation nationale sur la mise en accessibilité économique des SIADM pour les usagers du système de soins. À cet égard, l'identification d'une section consacrée

au financement de l'IA dans les lois de financement de la Sécurité sociale pourrait être expertisée. Dans ce débat, le rôle des acteurs de la protection complémentaire pour identifier et stimuler la mise en accessibilité de l'innovation en IA devrait être attentivement examiné.

6) Créer les conditions de la confiance en incitant les développeurs à fournir un certain niveau d'explicabilité du SIADM qu'ils mettent sur le marché.

7) Promouvoir le développement de nouvelles compétences professionnelles permettant d'expliquer aux soignants les propriétés des SIADM.

8) Créer les conditions de la confiance en promouvant le développement de justifications étayant les promesses et, d'autre part, la publication de résultats de recherche négatifs, c'est-à-dire ceux qui ne confirment pas les promesses annoncées.

9) Encourager la recherche pluridisciplinaire et la formation professionnelle pour mettre au point une définition contrôlée de conformité des SIADM dénuée de biais.

10) Promouvoir l'utilisation d'essais cliniques pour évaluer le rapport bénéfice-risque des SIADM de la même façon que cela se fait pour les médicaments.

11) Veiller à limiter les biais potentiels liés à la programmation même des algorithmes et aux différentes bases de données utilisées pour la mise au point des SIADM, dont les données d'apprentissage sont spécifiques à leur date de conception.

12) Promouvoir la définition, par les industriels, de critères qui permettront d'évaluer un SIADM dès sa conception (« Ethics by design »). Ceux-ci peuvent être génériques ou spécifiques à un SIADM.

13) Établir un certain nombre de critères sociaux d'évaluation des SIADM (critères médicaux pour le patient, organisationnels pour les établissements et professionnels de santé, sanitaires pour la population, économiques pour l'ensemble des acteurs, etc.) permettant d'en dresser un bilan éthique.

14) Promouvoir le statut d'auxiliaire numérique. Cette personne fournirait une raison compréhensible par le patient ou son représentant légal ou sa personne de confiance du résultat produit par le SIADM qui soit également représentative de son fonctionnement. Il pourrait être rapproché du statut « d'infirmier de pratique avancé » en développement actuellement.

15) Veiller à ce que les concepteurs de SIADM indiquent clairement les performances et les limites de leurs SIADM et les expliquent en termes compréhensibles pour leurs utilisateurs.

16) Adapter les cursus d'enseignement (formation initiale et continue) pour les professions médicales et paramédicales pour qu'ils forment aux technologies d'IA en intégrant la prise en compte de leurs enjeux éthiques.