

# L'intelligence artificielle dans le quotidien médical

Domaines d'application en médecine:  
utilité, défis et exigences de la FMH



**Éditrice**

FMH – Fédération des médecins suisses, Berne

**Finalisation rédactionnelle**

hallokern.ch

**Concept, mise en page et illustrations**

atelierrichner.ch

**Publication**

Septembre 2022

[www.fmh.ch](http://www.fmh.ch) | [ehealth@fmh.ch](mailto:ehealth@fmh.ch)

**Code couleur des illustrations**

Les représentations graphiques de cette brochure sont réalisées dans les couleurs ocre et bleu.

- L'ocre fait référence à l'homme.
- Le bleu fait référence à la machine.

# Table des matières

Abréviations	4
5	Préface
6	Quels avantages les médecins attendent-ils de l'intelligence artificielle?
8	Qu'est-ce que l'intelligence artificielle?
9	Méthodes d'intelligence artificielle
16	Dans quels domaines de la médecine peut-on utiliser l'intelligence artificielle?
20	Quels sont les défis que l'intelligence artificielle présente en médecine?
	1. Équité 20
	2. Autonomie et contrôle 21
	3. Transparence 24
	4. Fiabilité 25
	5. Sécurité 26
	6. Protection des données 27
	Conclusion intermédiaire: Quelle partie prenante humaine doit veiller à ce qu'un système d'IA se montre et demeure digne de confiance? 29
30	Comment l'intelligence artificielle est-elle réglementée en médecine?
32	Quelles sont les exigences de la FMH en matière d'intelligence artificielle dans le domaine médical?
35	Quelle sera l'évolution du profil professionnel du médecin?
40	Références bibliographiques

## Abréviations

<b>ADR</b>	<i>Adverse Drug Reaction</i> (réaction indésirable à un médicament)
<b>AMA</b>	<i>American medical association</i> (Association médicale américaine)
<b>ASSM</b>	Académie suisse des sciences médicales
<b>ATF</b>	Arrêts principaux du Tribunal fédéral suisse
<b>CanMEDS</b>	<i>Canadian medical education directives for specialists</i> (directives canadiennes en matière d'éducation médicale pour les spécialistes)
<b>CDSS</b>	<i>Clinical Decision Support Systems</i> (systèmes d'aide à la décision clinique)
<b>CNN</b>	<i>Convolutional Neural Network</i> (réseau neuronal convolutif)
<b>CONSORT-AI</b>	<i>Consolidated Standards of Reporting Trials-Artificial Intelligence</i> (Normes consolidées de déclaration des essais cliniques – intelligence artificielle)
<b>CPME</b>	Comité permanent des médecins européens
<b>CPRD</b>	<i>Clinical Practice Research Datalink</i>
<b>DDI</b>	<i>Drug-drug interaction</i> (interaction médicament-drogue)
<b>DNN</b>	<i>Deep Neural Network</i> (réseau neuronal profond)
<b>EbM</b>	<i>Evidence-based medicine</i> (médecine fondée sur les preuves)
<b>ECG</b>	Électrocardiogramme
<b>EHR</b>	<i>Electronic health record</i> (dossier électronique de santé)
<b>e-person</b>	Personne électronique
<b>FDA</b>	<i>U.S. Food and Drug Administration</i>
<b>GAN</b>	<i>Generative Adversarial Network</i> (réseau antagoniste génératif)
<b>GPU</b>	<i>Graphical Processing Unit</i> (unité de traitement graphique)
<b>HLEG</b>	<i>High-Level Expert Group on AI</i> (Groupe d'experts de haut niveau sur l'intelligence artificielle)
<b>IA</b>	Intelligence artificielle
<b>ICMRA</b>	<i>International Coalition of Medicines Regulatory Authorities</i> (Coalition internationale des autorités de réglementation des médicaments)
<b>IP</b>	Propriété intellectuelle ( <i>Intellectual Property, IP</i> )
<b>IRGC</b>	<i>International Risk Governance Center</i> (Centre international de gouvernance des risques)
<b>IRM</b>	Imagerie par résonance magnétique
<b>MI-CLAIM</b>	<i>Minimum information about clinical artificial intelligence modelling</i> (Informations minimales sur la modélisation clinique de l'intelligence artificielle)
<b>ML</b>	<i>Machine Learning</i> (apprentissage automatique)
<b>NDD</b>	<i>Neural network-based method for drug-drug interaction prediction</i> (méthode basée sur les réseaux neuronaux pour la prédiction de l'interaction médicament-médicament)
<b>NLP</b>	<i>Natural Language Programming</i> (programmation du langage naturel)
<b>NN</b>	<i>Neural Network</i> (réseau neuronal)
<b>ODim</b>	Ordonnance sur les dispositifs médicaux 812.213
<b>PPG</b>	Photopléthysmographie
<b>RD</b>	Rétinopathie diabétique
<b>RDIV</b>	Règlement (UE) 2017/746 du Parlement européen et du Conseil du 5 avril 2017 relatif aux dispositifs médicaux de diagnostic in vitro ( <i>Regulation on in-vitro diagnostic medical devices, IVDR</i> )
<b>RDM</b>	Réglementation des dispositifs médicaux ( <i>Medical Device Regulation, MDR</i> )
<b>RF</b>	<i>Random Forest</i> (forêt aléatoire, forêt d'arbres décisionnels)
<b>RNN</b>	<i>Recurrent Neural Network</i> (réseau neuronal récurrent)
<b>SATW</b>	Académie suisse des sciences techniques
<b>SPIRIT-AI</b>	<i>Standard Protocol Items: Recommendations for Interventional Trials-Artificial Intelligence</i> (Éléments standard du protocole : recommandations pour les essais interventionnels – intelligence artificielle)
<b>SVM</b>	<i>Support Vector Machine</i> (machine à vecteurs de support)
<b>TRIPOD</b>	<i>Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis</i> (Rapport transparent d'un modèle de prédiction multivariable pour le pronostic ou le diagnostic individuel)

# Préface

L'intelligence artificielle (IA) semble omniprésente. C'est du moins l'impression que nous évoque la fréquence à laquelle on retrouve ce terme. Les médias l'exposent sous des formes aussi multiples que la conduite automatisée, la notation des citoyen-ne-s, ou encore la gestion des plaintes et des décisions juridiques assistée par intelligence artificielle.

Nous ne faisons pas que rencontrer le terme; nous employons l'IA au quotidien. Un texte en langue étrangère peut se traduire en quelques secondes dans la langue souhaitée. Nous pouvons ainsi accéder aux reportages du monde entier, sans plus nous heurter aux barrières linguistiques des régions et des pays. Une machine peut également convertir instantanément le langage parlé en un texte écrit, de sorte à pouvoir exploiter son contenu. En voiture, la voix de notre téléphone «intelligent» nous guide jusqu'à notre destination, pour nous épargner la recherche de notre itinéraire sur la carte. Même les embouteillages ne devraient plus nous ralentir, puisque la voie qui permet de les contourner nous est tout simplement proposée.

En revanche, ne vous est-il pas non plus arrivé de vous demander pourquoi cette gentille voix vous a fait parvenir jusqu'à une étroite route forestière, tellement cahoteuse qu'elle mériterait le qualificatif de sentier forestier, en vous inspirant le sentiment désagréable qu'il s'agit ici d'un raccourci bien inhabituel? Suivre une route secondaire reste en soi une surprise bien inoffensive – tant qu'elle ne se termine pas dans un précipice ou un lac. Sur le chemin de la médecine, l'absence d'esprit critique pourrait toutefois conduire à des issues fatales. C'est pourquoi il s'agit d'éviter les détours inutiles et prévenir ainsi tout risque de dommages.

Partons donc ensemble à la recherche des réponses aux questions suivantes:

- Qu'est-ce qui se cache derrière le terme «intelligence artificielle»?
- Quelles sont ses possibilités et ses limites?
- À quels défis la mise en œuvre d'IA dans le domaine médical est-elle confrontée?
- Comment pouvons-nous en minimiser les risques et optimiser les avantages?
- En tant que médecins, quelles exigences en déduisons-nous en ce qui concerne l'utilisation de l'IA?

Si l'on s'intéresse de plus près à l'intelligence artificielle, on se rend compte qu'elle ne constitue pas une huitième merveille du monde pour la médecine. Elle n'est pas non plus à craindre comme un golem qui, une fois hors de contrôle, deviendrait un danger pour son créateur. Cette brochure a pour objectif de démystifier l'IA, de sorte que ses lectrices et lecteurs puissent se forger une représentation réaliste des possibilités et des limites de son utilisation dans le domaine médical. C'est ainsi qu'elle pourra contribuer, en premier lieu et en fin de compte, à améliorer la santé de nos patient-e-s.

Je vous souhaite une lecture intéressante et agréable!



Dr méd. Alexander Zimmer,  
membre du comité central de la FMH,  
responsable du département Numérisation/  
eHealth

# Quels avantages les médecins attendent-ils de l'intelligence artificielle ?

Il ne fait aucun doute que l'intelligence artificielle provoque des changements profonds dans tous les domaines de la médecine, changements qui sont déjà en cours. Les médecins de Suisse sont cependant résolument d'avis que la relation personnelle entre le médecin et les patient-e-s doit rester centrale, même dans un monde marqué par l'intelligence artificielle.

On estime que les connaissances médicales doublaient encore tous les 50 ans, en 1950. En 2010, la durée nécessaire pour cela s'est réduite à 3,5 ans, et en 2020, le doublement a eu lieu tous les 73 jours<sup>1</sup>. La numérisation dans le domaine de la santé génère de plus en plus de données disponibles sous forme numérique. Celles-ci peuvent en principe être intégrées dans les processus décisionnels, afin d'améliorer la qualité du diagnostic et du traitement. Les dernières générations d'ordinateurs disposent d'une énorme puissance de calcul et peuvent tester des hypothèses liées à la maladie à l'aide de grands ensembles de données multimodales. Par rapport aux humains, les ordinateurs sont capables de traiter les informations de manière plus systématique, de prendre des décisions plus cohérentes et de réagir plus rapidement aux changements<sup>2</sup>.

La littérature spécialisée décrit l'impact de l'IA en médecine au moyen d'une classification en trois niveaux:

1. l'impact sur les médecins: sous la forme de systèmes de reconnaissance et de décision assistés par ordinateur, par exemple pour interpréter rapidement et précisément des images;
2. l'impact sur les systèmes de santé: par l'amélioration des processus de travail et la minimisation du potentiel d'erreurs médicales, ainsi que par l'adaptation des soins, par exemple vers un suivi à domicile basé sur l'IA et combiné à de la télémédecine;
3. l'impact sur les patient-e-s: par la possibilité de traiter leurs propres données afin de promouvoir leur santé, et en modifiant la relation entre médecins et patient-e-s<sup>3,4</sup>.

**Du point de vue des médecins, l'intelligence artificielle devrait servir en premier lieu à optimiser les techniques et les processus au long du parcours du patient.**

L'opinion du corps médical et de la population sur l'utilité des applications numériques en matière de santé a fait l'objet du *Digital Trends Survey 2021*. Les résultats de cette enquête menée par la FMH évoquent, quant à eux, de manière impressionnante que l'acceptation de l'IA se heurte à certaines limites. Certes, les médecins voient l'utilité de telles applications dans la simplification des processus administratifs ou des procédures de travail. En revanche, ils s'opposent aux applications qui s'apprêtent à progressivement remplacer la visite médicale. Le corps médical est convaincu que le «facteur humain» revêt une grande importance pour la réussite du traitement et que l'IA n'est pas en mesure de le remplacer. C'est ainsi que 94% des médecins interrogés s'opposent à ce que le diagnostic soit exclusivement établi au moyen d'un logiciel intelligent. Quant au plan de traitement, 85% des médecins refusent d'en confier entièrement l'élaboration à un logiciel. Tant de la perspective du corps médical que de celle de la population, l'enquête a souligné la principale attente à l'égard de la numérisation: aménager du temps au médecin pour lui permettre de se consacrer davantage à la prise en charge individualisée<sup>5</sup>.

Pour les différentes étapes qui jalonnent le parcours du patient (*patient journey*), le corps médical estime que l'IA pourrait se révéler utile pour l'amélioration et le développement dans les domaines suivants:

- 
- **Prévention/pronostic** L'IA aide à identifier de nouveaux mécanismes biologiques d'une maladie, des facteurs de risque ainsi que des facteurs protecteurs.
- 
- **Diagnostic** Les systèmes d'aide à la décision clinique (*Clinical decision support systems*) permettent notamment une interprétation rapide et précise des résultats. Les CDSS peuvent minimiser la variabilité de la décision et optimiser sa cohérence et son efficacité.
- 
- **Traitement** L'utilisation de CDSS permet d'améliorer la surveillance des médicaments et contribue à augmenter la sécurité médicamenteuse en identifiant les effets secondaires et les interactions. En recherchant dans les dossiers électroniques de santé (ou EHR, pour l'anglais *electronic health records*), les systèmes d'IA pourraient simplifier l'identification et le recrutement des patient-e-s appropriés pour les essais cliniques.
- 
- **Surveillance des patient-e-s** L'IA peut simplifier le suivi de l'évolution du traitement et de l'observance des médicaments. En effet, les consultations médicales simples pourraient être partiellement remplacées par des consultations de télémédecine réalisées à distance.
- 
- **Tâches administratives** L'IA peut réduire la charge de travail administratif répétitif, telles que la saisie des données des patient-e-s (extraites à partir des dossiers de santé) et la vérification automatisée des résultats de laboratoire et d'imagerie médicale. Ceci soulage le personnel, et permet ainsi de libérer de la capacité et des ressources pour mieux les répartir en faveur d'une médecine individualisée et à visage humain.
-

# Qu'est-ce que l'intelligence artificielle ?

L'intelligence artificielle n'est pas une nouveauté; dans le monde scientifique, l'idée d'une machine pensante était déjà à l'ordre du jour il y a bien plus de 80 ans. C'est avec l'augmentation rapide des performances des ordinateurs que l'IA a fini par acquérir son importance actuelle.

À ce jour, il n'existe pas de définition de l'intelligence artificielle qui soit généralement établie<sup>6</sup>. Le mathématicien anglais Alan Turing a probablement été le premier à s'intéresser aux machines intelligentes. Lors d'un symposium à Manchester, en 1947 déjà, il a formulé la question centrale concernant l'IA: « Les machines peuvent-elles penser ? » L'un des pères fondateurs de l'IA, John McCarthy, en a inventé le terme, qu'il a défini en tant que « science et technique de la fabrication de machines intelligentes, en particulier de programmes informatiques intelligents »<sup>7</sup>. On définit aussi couramment l'IA comme « l'étude des calculs qui permettent de percevoir, de penser et d'agir »<sup>8</sup>.

N'étant donc pas apparue récemment, pourquoi l'IA est-elle devenue un thème omniprésent dans de nombreux domaines de la vie humaine au cours des dernières années ? C'est la nette amélioration de la puissance de calcul et des technologies de stockage, qui permettent de traiter de grandes quantités de données, qui a rendu possible son essor fulgurant. En un mot : *big data*. Pour le secteur médical, ce terme englobe notamment l'imagerie médicale à haute résolution, le séquençage du génome, les biocapteurs, ou encore les dossiers électroniques de santé (EHR).

**La fonction de l'intelligence artificielle concerne essentiellement les options en vue d'une action et s'articule autour de leur perception, leur analyse et leur développement.**

L'IA, ou la machine intelligente, peut être décrite en tant que système informatisé (système d'IA) capable d'élaborer des prédictions et des recommandations, ou de prendre des décisions. Elle peut ainsi influencer ainsi des environnements réels, dans un objectif

défini par l'humain. En se basant sur des paramètres de départ fournis par la machine et/ou l'humain, l'IA est en mesure de:

1. percevoir des environnements réels;
2. analyser de telles perceptions de manière automatisée, dans la perspective d'un certain objectif ou résultat (p. ex. avec l'apprentissage automatique), ainsi qu'en déduire des modèles;
3. tirer des conclusions à partir du modèle (inférence de modèle), afin de rassembler des connaissances et des possibilités d'action<sup>6</sup>.

En outre, les systèmes d'IA sont conçus de manière à fonctionner avec différents degrés d'autonomie<sup>6,9</sup>.

La conduite automatisée illustre clairement les trois caractéristiques de l'IA qui viennent d'être mentionnées.

1. Les véhicules autonomes disposent d'un grand nombre de capteurs et de caméras, qui fonctionnent notamment avec des ultrasons ou des ondes radio. Cela leur permet de percevoir leur environnement.
2. Au moyen de deep learning (ou apprentissage profond), un sous-domaine de l'apprentissage automatique (c'est-à-dire une méthodologie fondamentale de l'IA), ces images et ces impulsions provenant de capteurs sont analysées et transformées en un modèle visant une conduite automatisée.
3. Ce modèle se sert des données collectées pendant la conduite afin que le système d'IA puisse prendre la décision de rouler, de freiner ou de s'arrêter, dans le but d'une « conduite sans accident ». En cas d'urgence, la personne au volant du véhicule autonome garde la possibilité d'intervenir pour contourner le système d'IA et reprendre le contrôle<sup>10</sup>.

Après avoir examiné dans les grandes lignes la nature de l'IA, nous pouvons explorer plus en détail la méthodologie qui la sous-tend. Quelles sont les méthodes d'IA existantes, et en quoi se distinguent-elles les unes des autres ?

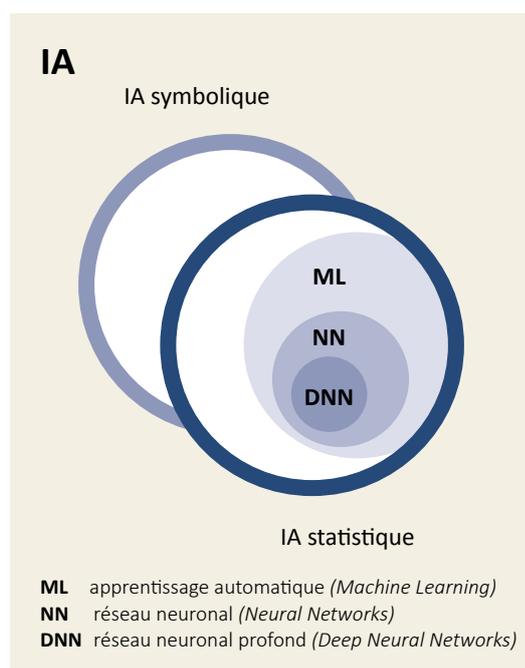
# Méthodes d'intelligence artificielle

Dans l'histoire récente de l'intelligence artificielle, la recherche a suivi deux orientations fondamentalement différentes: l'approche symbolique et l'approche statistique. Aujourd'hui, l'accent est mis sur l'IA statistique, qui consiste à résoudre des problèmes complexes à l'aide de modèles mathématiques et de méthodes statistiques.

Par le passé, la recherche en matière d'intelligence artificielle a distingué l'IA symbolique de l'IA statistique<sup>6</sup> (Fig. 1). Du milieu des années 1950 à la fin des années 1980, c'est l'approche symbolique qui a dominé la recherche en IA. Elle avait pour objectif de déterminer si l'intelligence humaine pouvait être réduite à la représentation de symboles, et, le cas échéant, dans quelle mesure. La recherche en IA reposait sur l'hypothèse que tout système intelligent – tant l'humain que la machine – doit fonctionner en traitant des structures de données constituées de symboles<sup>11</sup>. Pour amorcer l'étude de l'IA symbolique, il a fallu développer de manière précise les structures qui sous-tendent la prise de décision. Il s'agit là de représenter la complexité du monde réel et d'amener les machines à aboutir à des décisions semblables à celles que prennent les humains<sup>6</sup>. L'IA symbolique comprend des techniques telles que la programmation logique (*logic programming*), la logique de description (*description logic*), l'analyse syntaxique symbolique (*symbolic parsing*) et l'induction de règles (*rule induction*)<sup>12</sup>.

À partir des années 1990, l'approche statistique a gagné de plus en plus d'importance<sup>11</sup>. Alors que, dans l'approche symbolique, le système d'IA reproduit des règles humaines, il tente, dans l'approche statistique, de reconnaître des schémas parmi les données de sorte à en déduire un modèle<sup>6</sup>. L'IA statistique comprend des techniques telles que les réseaux bayésiens, les modèles de Markov cachés, les processus de décision markoviens, l'analyse statistique et les réseaux neuronaux<sup>12</sup>. Il existe également des systèmes d'IA qui se basent sur une combinaison de l'approche symbolique et de l'approche statistique. C'est le cas, par exemple, de la programmation du langage naturel qui s'appuie généralement sur de grandes quantités de données (aspect statistique) tout en tenant compte des règles de la grammaire (aspect symbolique)<sup>6</sup>.

Fig. 1 Aperçu des approches méthodologiques de l'intelligence artificielle



**Dans l'approche statistique de l'IA, on trouve les technologies d'apprentissage automatique. L'ordinateur y apprend à élaborer des prédictions basées sur des données, sans programmation explicite.**

L'apprentissage automatique, ou *machine learning* (ML), constitue un exemple de l'IA statistique. Les méthodes regroupées sous ce terme permettent aux ordinateurs d'apprendre à partir de données, sans programmation explicite, afin d'élaborer ou d'améliorer des prédictions. L'apprentissage automatique incarne un changement de paradigme. À partir de la programmation « normale », qui doit fournir explicitement toutes les instructions à l'ordinateur, il inaugure la « programmation indirecte »,

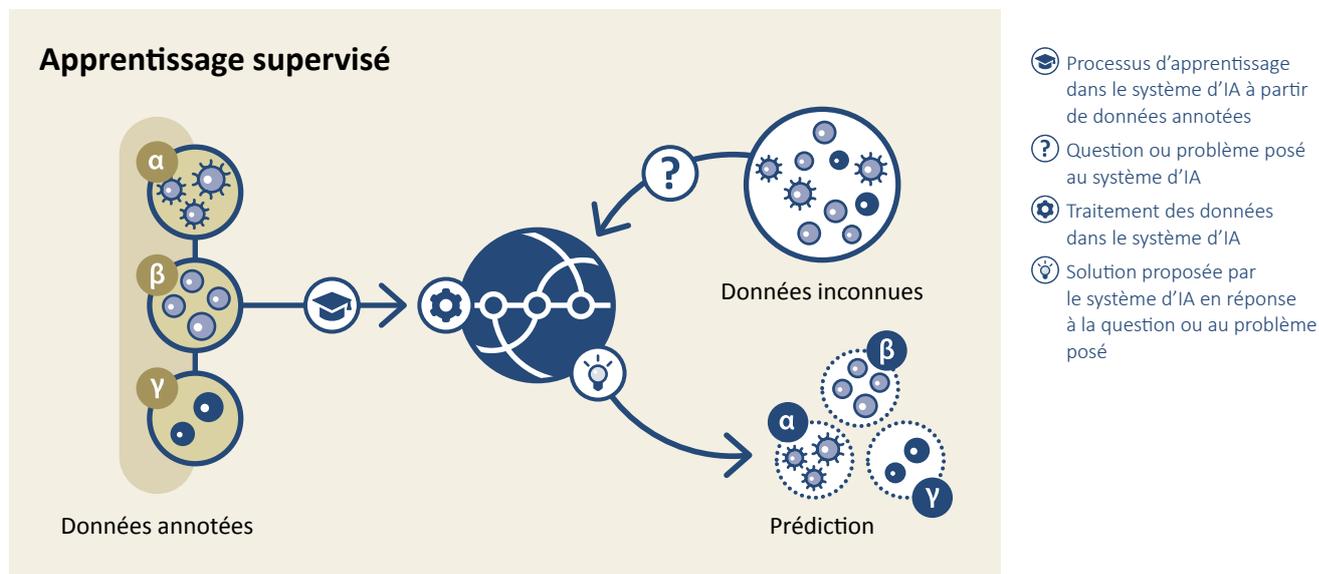
qui élabore des prédictions fondées sur des données<sup>13</sup>. Les méthodes d'apprentissage automatique permettent aux ordinateurs d'acquérir leurs propres connaissances. Pour ce faire, ils reconnaissent des schémas qui émergent des données brutes, forment des prédictions basées sur de tels modèles, ou dégagent des conclusions à partir de données complexes<sup>14</sup>. Parmi les techniques utilisées dans l'apprentissage automatique, on trouve notamment la régression linéaire et la régression logistique, les arbres de décision, l'analyse en composantes principales (*principle component analysis*), les réseaux neuronaux (*neural networks*, NN), ainsi que les réseaux neuronaux profonds (*deep neural networks*, DNN)<sup>6</sup>.

La manière d'apprendre de ces systèmes se classe actuellement en trois approches différentes (Fig. 2a-c) :

- apprentissage supervisé
- apprentissage non supervisé
- apprentissage par renforcement

Les systèmes d'IA peuvent s'appuyer sur différentes méthodes. Par exemple, un système d'IA peut intégrer des algorithmes d'apprentissage automatique. Ces éléments lui permettent de résoudre des problèmes complexes en les divisant en problèmes «plus simples», chacun étant traité par un algorithme d'apprentissage automatique<sup>15</sup>.

Fig. 2a Approche d'apprentissage en machine learning: apprentissage supervisé  
Supervised learning



### Caractéristiques

L'apprentissage supervisé regroupe les algorithmes d'IA<sup>a</sup> qui requièrent pour leur entraînement un ensemble de données annotées. L'annotation s'effectue typiquement par l'ajout de métadonnées. Il s'agit d'informations supplémentaires (p. ex. des codes de diagnostic) qui décrivent le contenu d'un ensemble de données, comme des images médicales, des textes ou d'autres données. Un tel ensemble de données doit donc contenir ce que l'on appelle des « paires d'entrée/sortie », par exemple les images numériques d'une lésion cutanée pathologique appariées au code diagnostique exact qui leur correspond.

<sup>a)</sup> Un algorithme est une méthode qui sert à résoudre un problème mathématique par un nombre fini d'étapes. En d'autres termes, l'algorithme constitue une séquence de règles qu'un ordinateur ou une machine doit suivre afin d'atteindre un objectif donné<sup>23</sup>.

Comme l'annotation des données doit être réalisée au préalable par intervention humaine, on qualifie ce type d'apprentissage automatique de « supervisé »<sup>16</sup>. L'ensemble de données annoté sert à entraîner l'algorithme, qui en extrait des modèles sur la base des corrélations reconnues entre l'entrée et la sortie. Ainsi, l'algorithme est désormais en mesure de réaliser des prédictions à partir d'ensembles de données inconnus et non annotés<sup>14</sup>. Dans l'apprentissage supervisé, on s'attend à des résultats d'un certain type connu à l'avance<sup>17</sup>.

### Données utilisées

**Ensemble de données annoté avec paires d'entrée/sortie** Les données sous-jacentes peuvent être structurées (p. ex. valeurs de laboratoire) ou non structurées (p. ex. images).

### Techniques utilisées

- Régression linéaire et régression logistique
  - Classification naïve bayésienne (naive Bayes)
  - Arbres de décision<sup>17,18</sup>
  - Réseaux neuronaux (NN)
  - Machine à vecteurs de support (SVM)
  - Forêt aléatoire (RF)<sup>19</sup>
- (Liste non exhaustive)

### Avantages

**Classification** Les données sont classées dans une catégorie spécifique.

**Régression** Les données sont associées à une valeur numérique.

### Domaines d'application

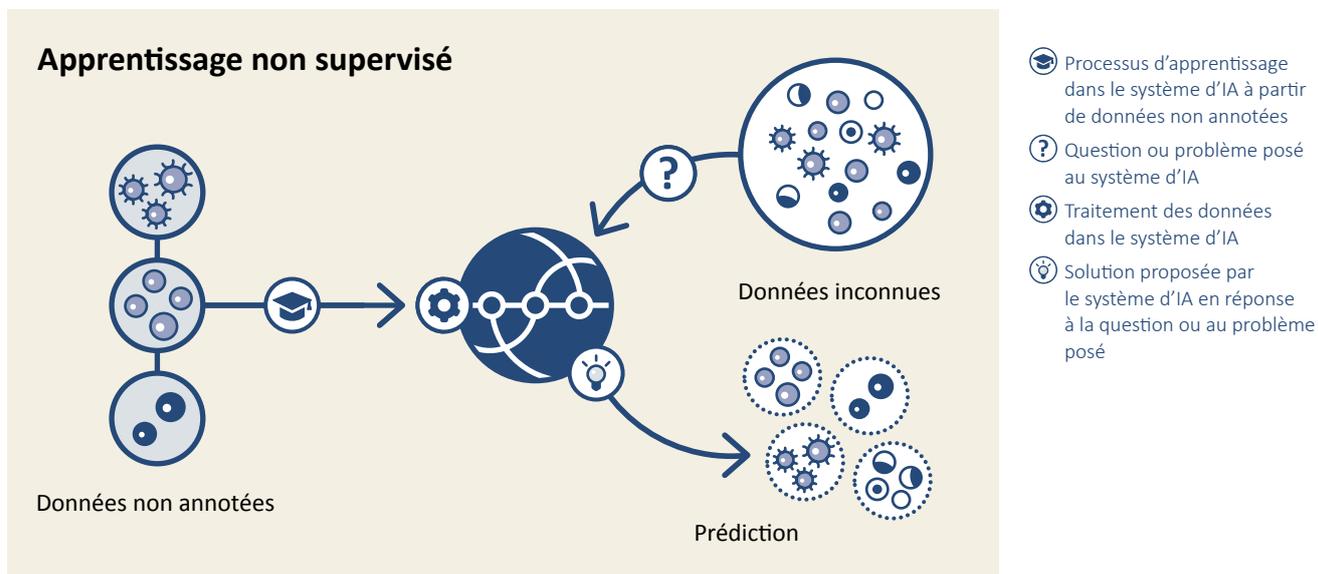
**Classification** Diagnostic médical, reconnaissance faciale<sup>16</sup>

**Régression** Posologie de pharmacothérapie, prédiction de l'heure du décès<sup>20</sup>, prédiction de la température<sup>16</sup>

### Limites

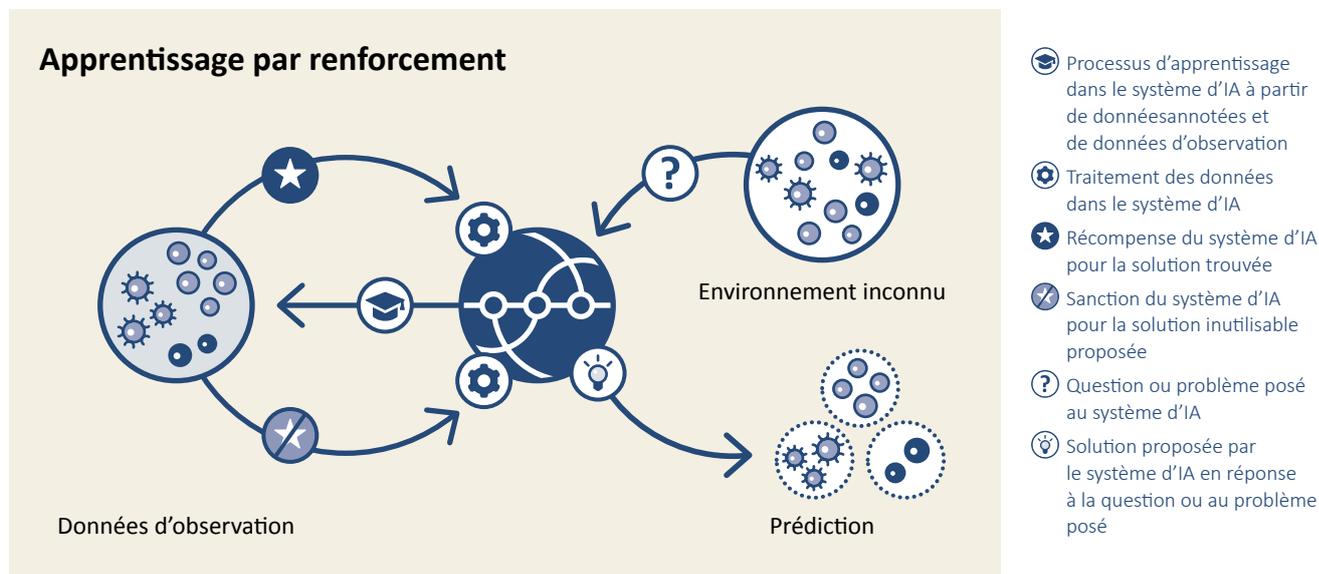
Pour l'entraînement et la validation, l'apprentissage supervisé requiert de grands ensembles de données annotées – dans l'idéal, au moyen d'une cohorte indépendante<sup>19</sup>.

Fig. 2b **Approche d'apprentissage en machine learning: apprentissage non supervisé**  
*Unsupervised learning*



<b>Caractéristiques</b>	L'apprentissage non supervisé regroupe les algorithmes d'IA qui tentent de reconnaître des modèles ou de tirer des conclusions à partir de données non annotées. Dans l'apprentissage non supervisé, il n'y a pas de « bonnes » réponses, contrairement à l'apprentissage supervisé <sup>16</sup> . L'apprentissage non supervisé vise à obtenir de nouvelles connaissances à partir de grandes quantités de données. L'algorithme d'apprentissage non supervisé détermine lui-même ce qui est différent ou intéressant dans l'ensemble de données <sup>17</sup> .
<b>Données utilisées</b>	<i>Ensemble de données non annotées</i> Les données peuvent être structurées (p. ex. valeurs de laboratoire) ou non structurées (p. ex. images).
<b>Techniques utilisées</b>	<ul style="list-style-type: none"> <li>– Procédures de clustering (partitionnement) probabilistes</li> <li>– Partitionnement en k moyennes (<i>k-means clustering</i>)<sup>18</sup></li> <li>– Analyse hiérarchique des clusters (groupes)<sup>21</sup></li> </ul> <p>(Liste non exhaustive)</p>
<b>Avantages</b>	<i>Reconnaissance des formes et clustering des données</i> Des données similaires sont attribuées à différents clusters <sup>16</sup> .
<b>Domaines d'application</b>	<p><i>Reconnaissance des formes et clustering des données</i></p> <ul style="list-style-type: none"> <li>– Reconnaissance de différents mécanismes biologiques d'un tableau clinique par ailleurs « homogène »</li> <li>– Identification de schémas cachés dans les données génotypiques ou phénotypiques<sup>17, 19</sup></li> </ul> <p><i>Réduction de la dimensionnalité des données</i> Représentation d'ensembles de données complexes au moyen de leur appartenance à un cluster: accélération d'autres algorithmes d'apprentissage automatique par la réduction du nombre de calculs<sup>16, 17</sup>.</p>
<b>Limites</b>	La validation des clusters soulève certaines difficultés. Afin d'y parvenir, on peut notamment mesurer la stabilité des clusters dans des cohortes indépendantes <sup>19</sup> , ou annoter des éléments de données similaires et mesurer leur appartenance à un cluster <sup>21</sup> .

Fig. 2c Approche d'apprentissage en machine learning: apprentissage par renforcement  
Reinforcement learning



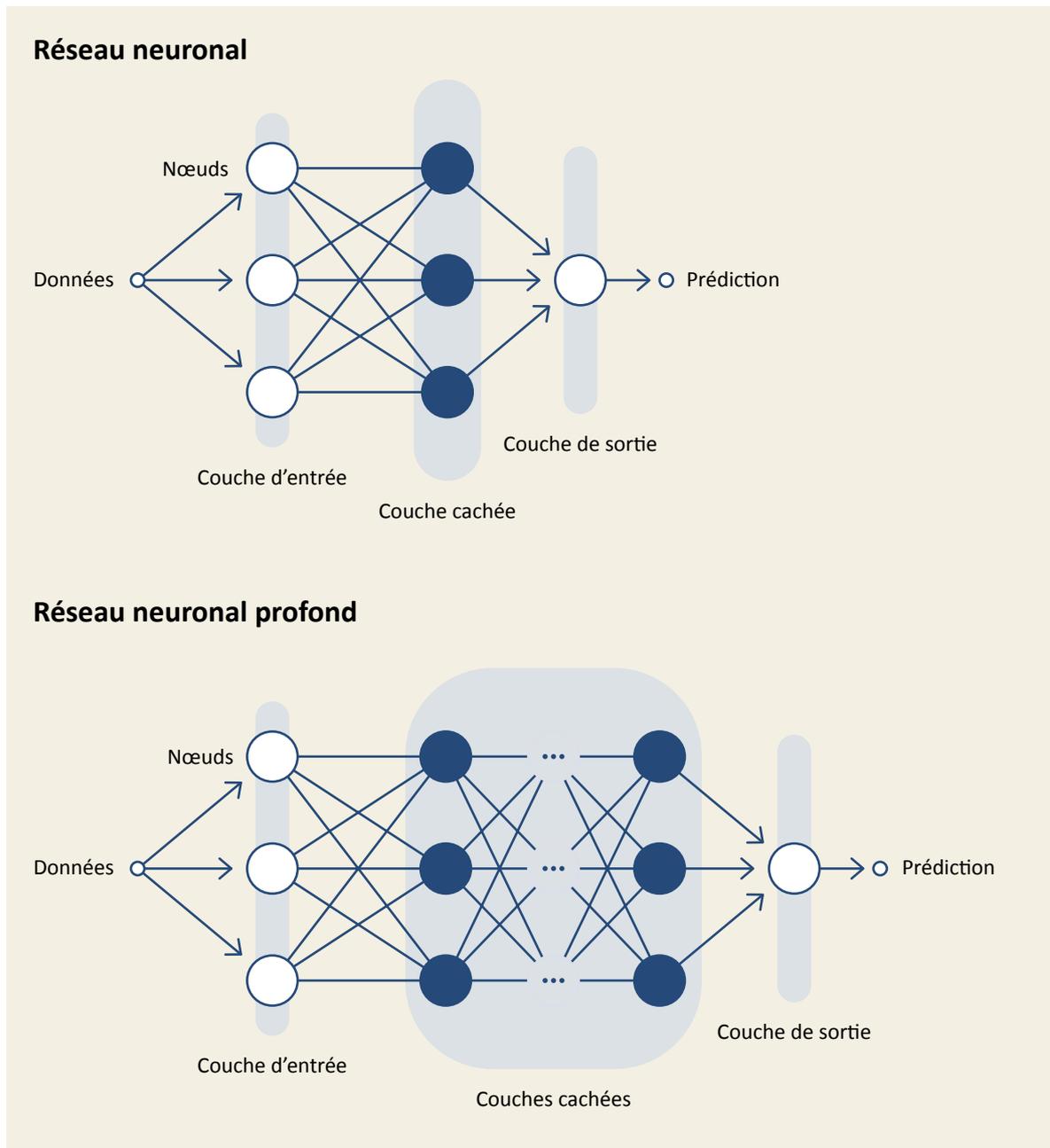
<b>Caractéristiques</b>	L'apprentissage par renforcement regroupe les algorithmes d'IA qui apprennent par leur interaction avec l'environnement.
	Un tel algorithme essaie des solutions de manière autonome, et son environnement lui renvoie un feedback (rétroaction). Ce retour d'information sur la manière dont il a trouvé la solution constitue un renforcement soit positif soit négatif. Grâce à la répétition de l'action et de la rétroaction correspondante, l'algorithme apprend à reconnaître l'effet particulier d'une certaine action dans un certain contexte <sup>14</sup> .
<b>Données utilisées</b>	<b>Données annotées</b> D'une manière typique, l'apprentissage par renforcement se base sur les données annotées qui sont d'abord mises à la disposition de l'algorithme d'IA.  Dans un deuxième temps, l'algorithme d'IA génère des données supplémentaires à partir de l'interaction avec son environnement (données d'observation) <sup>15</sup> .
<b>Techniques utilisées</b>	Les mêmes techniques que pour l'apprentissage supervisé.  L'entraînement de l'algorithme s'effectue en maximisant la rétroaction positive (« récompense ») : il apprend ainsi à sélectionner les décisions optimales <sup>15</sup> .
<b>Avantages</b>	Informations sur le fonctionnement de l'environnement <sup>11</sup>
<b>Domaines d'application</b>	<ul style="list-style-type: none"> <li>– Stratégies de traitement (p. ex. stratégie optimale pour le traitement de la septicémie aux soins intensifs<sup>22</sup>)</li> <li>– Contrôle robotique</li> <li>– Jeux qui fonctionnent au moyen d'IA<sup>14</sup></li> </ul>
<b>Limites</b>	L'apprentissage par renforcement requiert une mémoire volumineuse <sup>19</sup> .  Pour des raisons de sécurité, l'optimisation ne peut pas être réalisée par une interaction directe avec l'environnement (p. ex. le médecin), en particulier dans le domaine médical <sup>22</sup> .

**Deep learning : apprentissage automatique avec intégration de réseaux neuronaux profonds**

En lien avec l'apprentissage automatique, le deep learning suscite de plus en plus d'intérêt depuis 2012<sup>15,24</sup>. Le deep learning se base sur la technique des réseaux neuronaux profonds (*deep neural networks*, DNN), qui sont eux-mêmes issus des réseaux neuronaux (NN) (Fig. 3).

Les réseaux neuronaux existent depuis les années 1950<sup>25</sup>. Leur développement a évolué en réseaux neuronaux profonds (« *Networks went deep* ») à partir de 2006<sup>15,26</sup>. L'une des conditions de base pour cela a été *le big data* : à cette époque, on a disposé d'ensembles de données beaucoup plus importants qu'auparavant. En outre, le matériel informatique avait gagné en efficacité – en particulier les processeurs, devenus beaucoup plus rapides<sup>15</sup>.

Fig. 3 Caractéristiques du réseau neuronal et du réseau neuronal profond



Type de réseau	Caractéristiques
<b>Réseau neuronal (NN)</b>	<p>Les réseaux neuronaux se composent d'une couche d'entrée (<i>input layer</i>), d'une couche cachée (<i>hidden layer</i>), ainsi que d'une couche de sortie (<i>output layer</i>). Chacun de ces niveaux est constitué de nœuds (<i>nodes</i>). Un réseau neuronal peut contenir des millions de nœuds.</p> <p>Les nœuds des différents niveaux évoquent les neurones du cerveau humain<sup>27,28</sup>. Chaque nœud est pondéré et possède une valeur seuil, et chaque nœud du niveau suivant lui est relié. Si le signal de sortie (<i>output</i>) d'un nœud est supérieur à sa valeur seuil, ce nœud envoie son information aux nœuds du niveau suivant: l'architecture informatique apprend par le biais de ses essais et de ses erreurs<sup>24</sup>.</p>
<b>Réseau neuronal profond (DNN)</b>	<p>C'est le nombre de niveaux qui distingue les réseaux neuronaux profonds des réseaux neuronaux. Niveaux d'entrée et de sortie inclus, les réseaux neuronaux ne comprennent que trois niveaux. En revanche, les réseaux neuronaux profonds comprennent <u>deux ou plusieurs couches cachées</u> (<i>hidden layers</i>) en plus de la couche d'entrée (<i>input layer</i>) et de la couche de sortie (<i>output layer</i>)<sup>28</sup>.</p> <p>Les réseaux neuronaux profonds peuvent contenir des milliards de nœuds répartis sur des dizaines de niveaux.</p> <p>Les variantes du réseau neuronal profond comprennent les réseaux neuronaux convolutifs (<i>convolutional neural networks</i>), les réseaux neuronaux récurrents (<i>recurrent neural networks</i>) et les réseaux antagonistes génératifs (<i>generative adversarial networks</i>).</p> <ul style="list-style-type: none"> <li>– Les réseaux neuronaux convolutifs (CNN) sont utilisés dans la reconnaissance d'images, l'identification des personnes, la robotique autonome, les véhicules autonomes pour la conduite automatisée et la reconnaissance vocale.</li> <li>– Les réseaux neuronaux récurrents (RNN) sont utilisés pour la reconnaissance vocale et la traduction.</li> <li>– Les réseaux antagonistes génératifs (GAN) sont utilisés notamment pour la génération de nouvelles images et photos: c'est cette technique qui a donné lieu au phénomène des <i>deep fakes</i><sup>14</sup>.</li> </ul>

### Différence entre deep learning et apprentissage automatique classique

En quoi le deep learning se distingue-t-il de l'apprentissage automatique classique? Les algorithmes classiques d'apprentissage automatique utilisent typiquement des ensembles de données structurés pour réaliser leurs prédictions. Des données non structurées peuvent également être utilisées, à condition d'être manuellement préparées (*pre-processing*) de manière à correspondre au format structuré qui permet de les utiliser.

Le deep learning supprime une partie de cette préparation manuelle des données. Les algorithmes d'apprentissage profond peuvent travailler avec des données non structurées (p. ex. images, texte) sans que celles-ci aient été structurées au préalable. Dans le deep learning, les algorithmes extraient de manière autonome les caractéristiques essentielles des données non structurées de sorte à les différencier les unes des autres<sup>29</sup>. En médecine, le deep learning est notamment employé pour identifier et interpréter des schémas qui se retrouvent dans l'imagerie médicale, les préparations pathologiques, les photographies de lésion cutanée ou de fond d'œil, les électrocardiogrammes ou encore dans les paramètres vitaux<sup>3</sup>.

# Dans quels domaines de la médecine peut-on utiliser l'intelligence artificielle ?

On peut imaginer que les utilisations de l'intelligence artificielle accompagnent l'humain dans tous les aspects de sa santé, depuis sa conception jusqu'à son décès. Ce chapitre présente des exemples d'applications concrètes de chaque domaine du parcours du patient, afin d'illustrer les possibilités d'utilisation de l'IA en médecine.

## **L'IA dans les domaines de la prévention et du pronostic: l'exemple de la fibrillation auriculaire**

Principal facteur de risque cardio-embolique, la fibrillation auriculaire est un trouble du rythme cardiaque qui s'accompagne d'une augmentation de la morbidité et de la mortalité. Sa détection, cruciale pour la prévention des accidents vasculaires cérébraux, est désormais possible à un stade précoce grâce aux *wearables*. Il s'agit d'appareils comme une *smartwatch* ou un smartphone, portés directement sur le corps, qui surveillent le rythme cardiaque en continu. Ils fonctionnent au moyen de capteurs optiques qui enregistrent les variations du volume sanguin au sein du tissu microvasculaire.

Cette technologie porte le nom de photopléthysmographie (PPG). Elle requiert une source de lumière et un photodétecteur. Le signal ainsi généré se compose de deux éléments. D'une part, une composante de courant continu représente la lumière réfléchie, ou transmise par le sang artériel au repos, le sang veineux, la peau et les tissus. D'autre part, une composante de courant alternatif pulsé résulte de la variation de l'absorption de la lumière, provoquée par les changements du volume sanguin artériel. En outre, ces appareils portables sont pourvus de capteurs d'accélération, qui mesurent les forces d'accélération dans l'espace tridimensionnel, de sorte à se prémunir contre les artefacts qui peuvent survenir, en cas de mouvement, dans les signaux générés.

La PPG permet de déduire des paramètres physiologiques tels que la fréquence cardiaque, la pression artérielle, le débit cardiaque et la saturation en oxygène<sup>30</sup>. Au repos comme lors l'activité physique, le *wearable* mesure le signal en continu et apprend à reconnaître la fréquence cardiaque de la personne qui le porte. Un écart significatif dans les valeurs reconnues comme normales déclenche une alarme. De plus, l'appareil portable peut recommander, par exemple, d'effectuer un ECG, et, le cas échéant, il peut aussi le réaliser directement<sup>3</sup>. Ces algorithmes emploient une variété de méthodes: approches statistiques (p. ex. régression logistique, modèle de Markov), deep learning (p. ex. réseau neuronal

convolutif, CNN), apprentissage automatique (p. ex. machine à vecteurs de support, arbres de décision, *k-nearest neighbors*)<sup>30</sup>.

## **L'IA dans le domaine du diagnostic: l'exemple de la rétinopathie diabétique**

Si elle n'est pas diagnostiquée et traitée à temps, la rétinopathie diabétique (RD) peut entraîner une baisse importante de la vision, voire la cécité<sup>31</sup>. La photographie du fond de l'œil permet d'obtenir des images de la rétine, de la papille optique et de la macula. De telles images permettent de détecter et de surveiller des maladies telles que la RD, le glaucome, les néoplasies de la rétine et la dégénérescence maculaire liée à l'âge<sup>32</sup>. Ces dernières années, la tomographie en cohérence optique s'est imposée en tant que méthode de référence pour la détection de l'œdème maculaire diabétique<sup>31</sup>.

Dans le cadre d'une étude prospective menée dans dix cliniques de soins primaires, des personnes diabétiques, sans rétinopathie connue, ont été examinées à l'aide d'un appareil à empreinte numérique propriétaire, assisté par IA. Des photographies du fond d'œil et des tomographies de cohérence optique de la rétine ont été interprétées par comparaison avec des images rétiniques de référence. Pour cela, le système d'IA a employé un algorithme de deep learning de type réseau neuronal convolutif (CNN). Les personnes qui ont participé à l'étude ont auparavant suivi un protocole standardisé pour s'entraîner à utiliser l'appareil d'imagerie. La FDA a approuvé cet appareil pour une utilisation professionnelle dans le diagnostic de l'œdème maculaire diabétique et de la RD modérée à sévère<sup>31,33</sup>.

**Les applications de l'intelligence artificielle existent dans presque tous les domaines de la médecine. L'IA peut également apporter un soutien lors d'interventions chirurgicales extrêmement délicates.**

### **L'IA dans le domaine du traitement:**

#### **l'exemple des interactions médicamenteuses**

Le terme d'interactions médicamenteuses (*drug-drug interaction, DDI*) désigne les effets indésirables résultant de la prise concomitante de deux ou plusieurs médicaments. Cette influence réciproque des médicaments peut provoquer des troubles supplémentaires (*adverse drug reaction, ADR*), et, dans le pire des cas, entraîner la mort<sup>34</sup>. Généralement âgées, les personnes touchées sont plus enclines à présenter une, voire de multiples morbidités<sup>35</sup>.

Le deep learning est utilisé pour anticiper de nouvelles interactions médicamenteuses. Un exemple d'une telle prévision de nouvelles interactions médicamenteuses est la méthode NDD: *neural network-based method for drug-drug interaction prediction*. Dans ce processus constitué de plusieurs étapes, les informations sur les propriétés chimiques des médicaments et sur les effets indésirables des médicaments sont exploitées dans un réseau neuronal<sup>34</sup>.

### **L'IA dans le domaine du suivi de la maladie:**

#### **l'exemple de la prévision de l'évolution**

L'IA peut aider à surveiller l'état de santé des patient-e-s à différents stades de leur maladie. Elle utilise à cette fin des données longitudinales. Il s'agit de données collectées sur une certaine période (p. ex. au moyen de capteurs) et qui permettent d'anticiper sur l'évolution de la maladie. Les capteurs multimodaux utilisés en soins intensifs illustrent bien cette technologie; l'IA se base sur leurs données afin de pouvoir annoncer une défaillance circulatoire imminente<sup>36</sup>. L'enregistrement des signaux ECG à l'aide de *wearables* en constitue un autre exemple; on s'en sert pour prédire l'aggravation d'une insuffisance cardiaque<sup>37</sup>.

### **L'IA dans le domaine de l'administration:**

#### **l'exemple de la documentation**

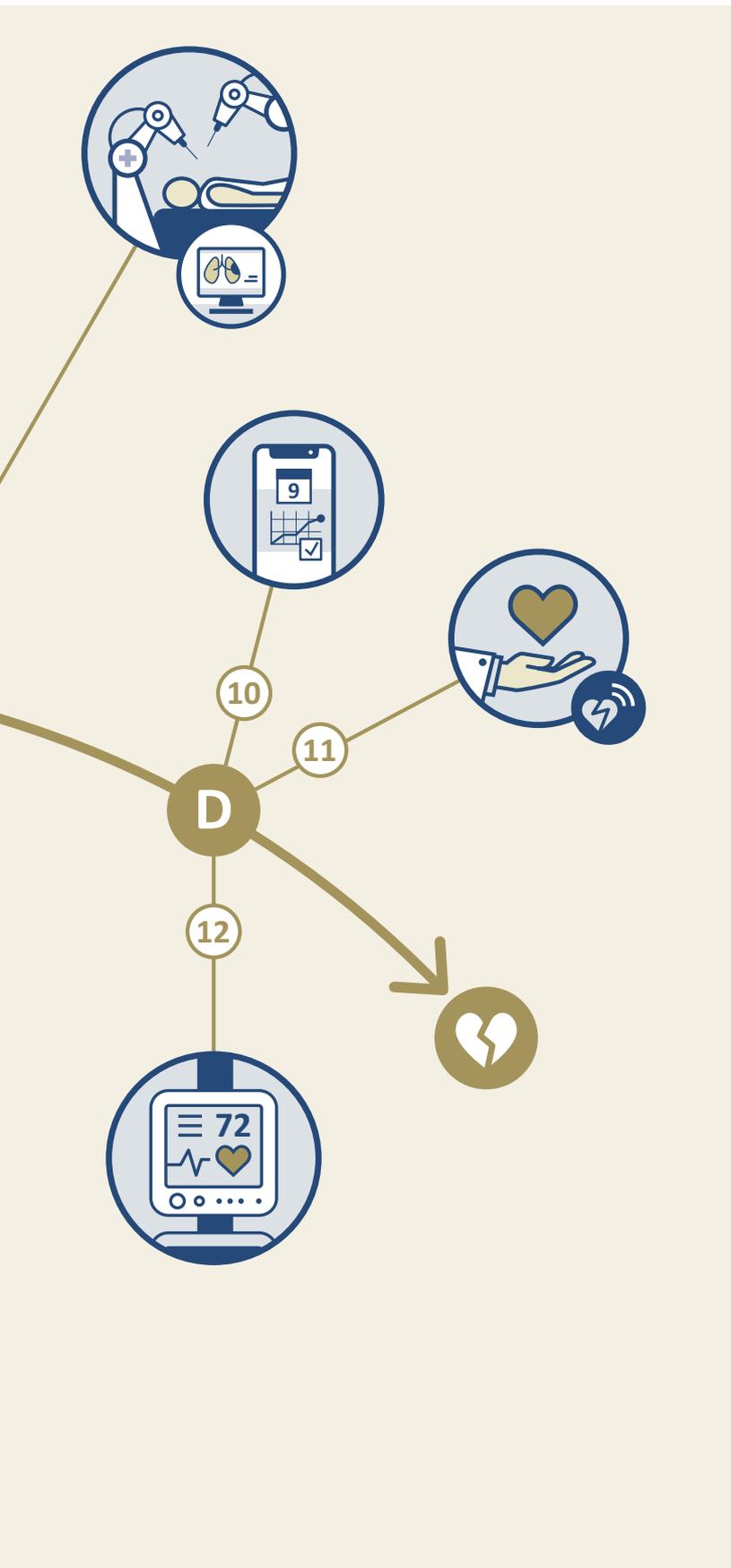
Omniprésente au long du parcours du patient, la documentation médicale fait partie intégrante de la routine quotidienne du personnel. En Suisse, les médecins qui travaillent dans le domaine des soins somatiques aigus consacrent chaque jour 20 % de leur temps de travail à des tâches liées à la documentation des dossiers médicaux. Dans ce domaine, ce chiffre atteint même 27 % en ce qui concerne les médecins en formation postgraduée. Ceux-ci consacrent ainsi presque autant de temps aux tâches administratives qu'aux activités en présence des patient-e-s<sup>38</sup>.

L'utilisation d'un dossier électronique de santé (EHR) augmente la disponibilité des données médicales<sup>39</sup>. Toutefois, un EHR peu convivial est également susceptible d'entraîner une augmentation du temps imparti à la documentation, au détriment du temps effectif consacré à la prise en charge des patient-e-s. La satisfaction du corps médical s'en trouve alors entamée. C'est pourquoi, aux États-Unis, des équipes médicales disposent de « scribes cliniques » pour rédiger la documentation médicale, ce qui permet de réduire la charge administrative et d'éviter un bon nombre d'erreurs. Ces scribes sont des personnes formées à la collaboration avec les médecins cliniciens. Elles sont employées dans le cadre des consultations médicales pour consigner les informations de manière pertinente et précise. Cela permet ainsi à l'équipe médicale de mieux se concentrer sur les aspects cliniques de la consultation. On utilise désormais également des « scribes numériques ». À l'aide d'IA, ces appareils utilisent les progrès de la reconnaissance vocale et du traitement automatique du langage naturel (TALN) afin de simplifier le travail de documentation des médecins<sup>40</sup>. La reconnaissance vocale permet aux médecins de dicter des notes, ou de documenter automatiquement les entretiens menés auprès de leurs patient-e-s<sup>41</sup>. Les conversations ne sont pas seulement transcrites mot à mot, mais elles peuvent également être résumées, si nécessaire. En outre, des rapports cliniques compréhensibles peuvent être générés au moyen d'IA, directement à partir de données primaires, telles que des clichés de radiologie ou de pathologie<sup>42</sup>.

### **De nombreux domaines d'utilisation**

L'utilisation de l'IA concerne presque toutes les disciplines médicales: radiologie, pathologie, dermatologie, ophtalmologie, cardiologie, gastroentérologie, psychiatrie/psychothérapie, chirurgie, etc. Même des opérations microchirurgicales extrêmement délicates, telles que des interventions intraoculaires<sup>43</sup>, peuvent être réalisées avec des systèmes robotiques assistés par intelligence artificielle.





### A Prévention et pronostic

1. Sélection d'embryons en vue d'une fécondation in vitro
2. Analyse et interprétation génomique chez le nouveau-né
3. Mesure de la fréquence cardiaque via *smartwatch*

### B Diagnostic

4. Diagnostic d'infarctus du myocarde ou d'accident vasculaire cérébral
5. Prévention de la cécité en cas de rétinopathie diabétique
6. Interprétation d'imageries médicales

### C Traitement

7. Identification des effets indésirables des interactions entre différents médicaments
8. Traitement d'appoint de la dépression ou de l'anxiété via des chatbots ou des applications
9. Exécution extrêmement précise d'opérations par des robots

### D Surveillance des patient-e-s / suivi

10. Check-up quotidien, p. ex. après une maladie grave, et documentation des progrès thérapeutiques au moyen d'applications
11. Amélioration de la sécurité des patient-e-s: avertissement des médecins et du personnel soignant lorsque les valeurs mesurées se situent en dehors de la norme
12. Prédiction de la date/de l'heure du décès à l'hôpital dans le but d'adapter les soins palliatifs en conséquence

# Quels sont les défis que l'intelligence artificielle présente en médecine ?

L'utilisation de l'intelligence artificielle en médecine ouvre de nouvelles possibilités, mais comporte également des défis, voire des risques<sup>44</sup>. On peut les décrire en se référant aux six dimensions qui déterminent la fiabilité : équité, autonomie et contrôle, transparence, fiabilité, sécurité et protection des données.

Les six dimensions de la fiabilité s'inspirent des exigences essentielles du Cadre pour une IA digne de confiance, formulé pour l'Europe par le Groupe d'experts de haut niveau sur l'intelligence artificielle (HLEG)<sup>45</sup>. En outre, elles se basent sur le catalogue de vérification de l'IA (*KI-Prüfkatalog*) de l'Institut Fraunhofer<sup>46</sup>.

## 1. Équité

Le principe d'équité (*fairness*) interdit toute inégalité de traitement injustifiée et donc toute discrimination illicite<sup>47</sup>. Toute activité humaine est également influencée par des biais involontaires et inconscients<sup>39</sup>. Comme leur apprentissage se base sur des données existantes (p. ex. des données médicales) imprégnées de leur propre historicité, les systèmes d'IA ne sont pas non plus à l'abri de distorsions méthodologiques.

De tels biais peuvent se répercuter sur la santé des patient-e-s sous la forme d'erreurs médicales et/ou de déséquilibres socio-économiques en matière de santé. Ce phénomène affecte notamment des groupes « protégés » (*protected patients*), lorsqu'un système d'IA se sert de données provenant de dossiers électroniques de santé (EHR) incomplets. Ces biais surviennent en raison de données manquantes (*missing data bias*), de tailles d'échantillons trop petites (*minority bias*), ou d'erreurs de mesure et de classification (*informativeness bias*)<sup>48,49</sup>.

### Des biais de multiples origines

L'intelligence artificielle ne peut formuler des prédictions solides que concernant une population de patient-e-s avec un certain nombre de caractéristiques, et au sujet desquelles le système a été entraîné, testé et validé. Si le système d'IA est utilisé sur une population de patient-e-s pour laquelle il n'a pas été entraîné, testé et validé, il peut en résulter des erreurs de prédictions (*training-serving skew*)<sup>49</sup>.

À l'origine des biais, on peut trouver non seulement les jeux de données utilisés pour l'apprentissage, mais également la conception du système d'IA, ainsi que les interactions du système d'IA tant avec les médecins qu'avec les patient-e-s.

### Biais liés à la conception

Dans la conception du système d'IA, la présence de biais peut être imputée à l'utilisation de variables peu significatives. C'est le cas du biais d'étiquetage (*label bias*), qui survient lorsqu'une variable décrit plutôt les inégalités en matière de soins de santé au lieu d'une vérité médicale. La conception peut également comporter un biais dû à la cohorte utilisée (*cohort bias*), si les données fournies au système d'IA ne concernent que des populations de patient-e-s faciles à mesurer. De même, l'omission d'autres populations qui seraient importantes pour la généralisation médicale des prédictions à réaliser peut provoquer un biais, par exemple, dans la répartition par sexe<sup>49</sup>.

### Biais liés à l'interaction avec les médecins

Le manque d'information se trouve à l'origine des biais liés à l'interaction du corps médical avec le système d'IA, par exemple :

- lorsque le médecin ignore que le système d'IA élabore des prédictions moins précises pour un certain groupe de patient-e-s (*automation bias*)<sup>50</sup>, ou
- lorsque le médecin reprend des recommandations émises par IA dans un cas particulier où elles n'étaient pas justifiées mais que le système IA apprend de cette décision erronée (*feedback loop bias*)<sup>49</sup>, ou
- lorsque le médecin néglige, volontairement comme involontairement, les avertissements du système IA concernant des groupes protégés de patient-e-s (*dismissal bias*)<sup>49</sup>.

### Biais liés aux patient-e-s

Des biais imputables aux patient-e-s peuvent également apparaître, notamment lorsqu'une partie plus défavorisée de la population n'est pas en mesure d'employer un système d'IA, en raison d'obstacles technologiques ou sociaux (*privilege bias*). Il est également possible que certaines personnes qui ont vécu de mauvaises expériences en lien avec la technologie fassent preuve d'une méfiance systématique à son encontre. De tels groupes de patient-e-s refusent ainsi de se faire soigner de cette manière et ne peuvent en conséquence pas bénéficier de soins de santé appropriés (*informed mistrust*). Il en va de même lorsqu'un groupe de patient-e-s particulier ne dispose pas de «voix», et n'a pas son mot à dire en ce qui concerne le développement, l'utilisation et l'évaluation des systèmes d'IA (*agency bias*)<sup>49</sup>.

### Procédures de test utiles

Des outils tels que FairTest permettent de tester l'équité des prédictions des algorithmes d'IA et de leurs associations avec des caractéristiques sensibles comme le sexe, l'ethnie, etc. FairTest comprend également des fonctions de débogage, qui peuvent être programmées pour éliminer les déclencheurs des effets inéquitables identifiés<sup>51,52</sup>.

## 2. Autonomie et contrôle

Dans ce contexte, le terme d'autonomie se réfère à deux aspects: d'une part, l'autodétermination de l'être humain (p. ex. la personne qui se sert de l'intelligence artificielle); d'autre part, l'autonomie

ou plus exactement la marge de manœuvre du système. Il s'agit de définir, pour chaque système d'IA, la répartition appropriée et responsable des tâches entre l'intelligence artificielle et l'humain qui en fait usage<sup>46,47</sup>.

La collaboration entre l'humain et la machine peut s'envisager sur les modes suivants (Fig. 5):

### «In the Loop»

L'humain est indispensable au processus de travail. Cela signifie que la machine ne peut pas terminer une tâche sans la confirmation d'un humain. De plus, celui-ci peut corriger, remplacer et compenser les décisions que le système d'IA a prises de manière autonome<sup>46,53</sup>.

### «On the Loop»

La machine accomplit la majeure partie du travail. En situation ordinaire, l'humain n'est que peu – voire pas du tout – impliqué dans les décisions que prend le système d'IA. Il représente ici l'instance de contrôle, dont le rôle consiste à s'assurer que le processus de travail se déroule normalement et sans erreur. En outre, l'humain peut ultérieurement corriger, remplacer et compenser les décisions que le système d'IA a prises de manière autonome<sup>46,53</sup>.

### «Out of the Loop»

L'intervention humaine n'est pas requise. Les machines sont suffisamment précises et autonomes pour effectuer le travail de leur propre chef. La seule décision qui revient à l'humain consiste à choisir d'utiliser ou non le système d'IA<sup>46,53</sup>.

Fig. 5 Modes de collaboration envisageables entre l'humain et la machine

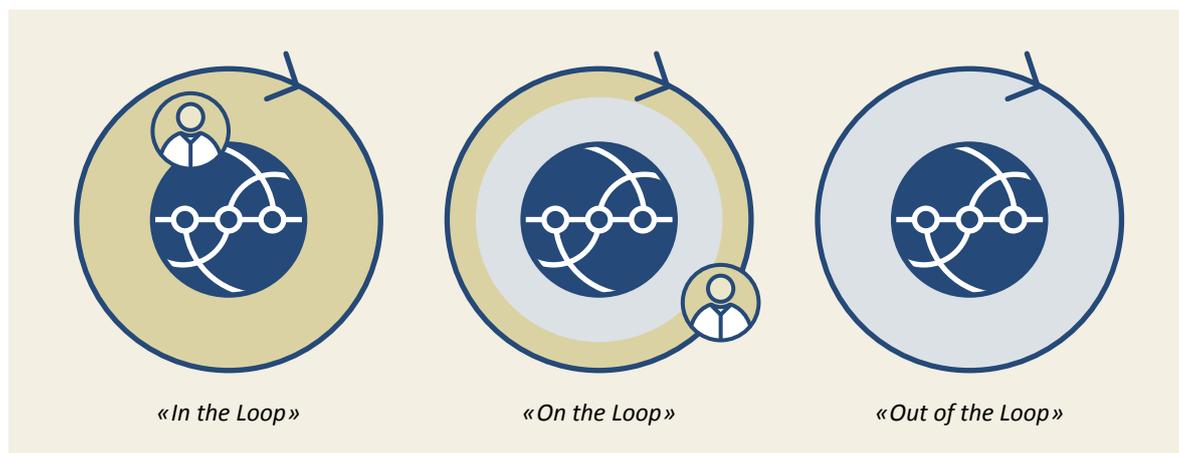
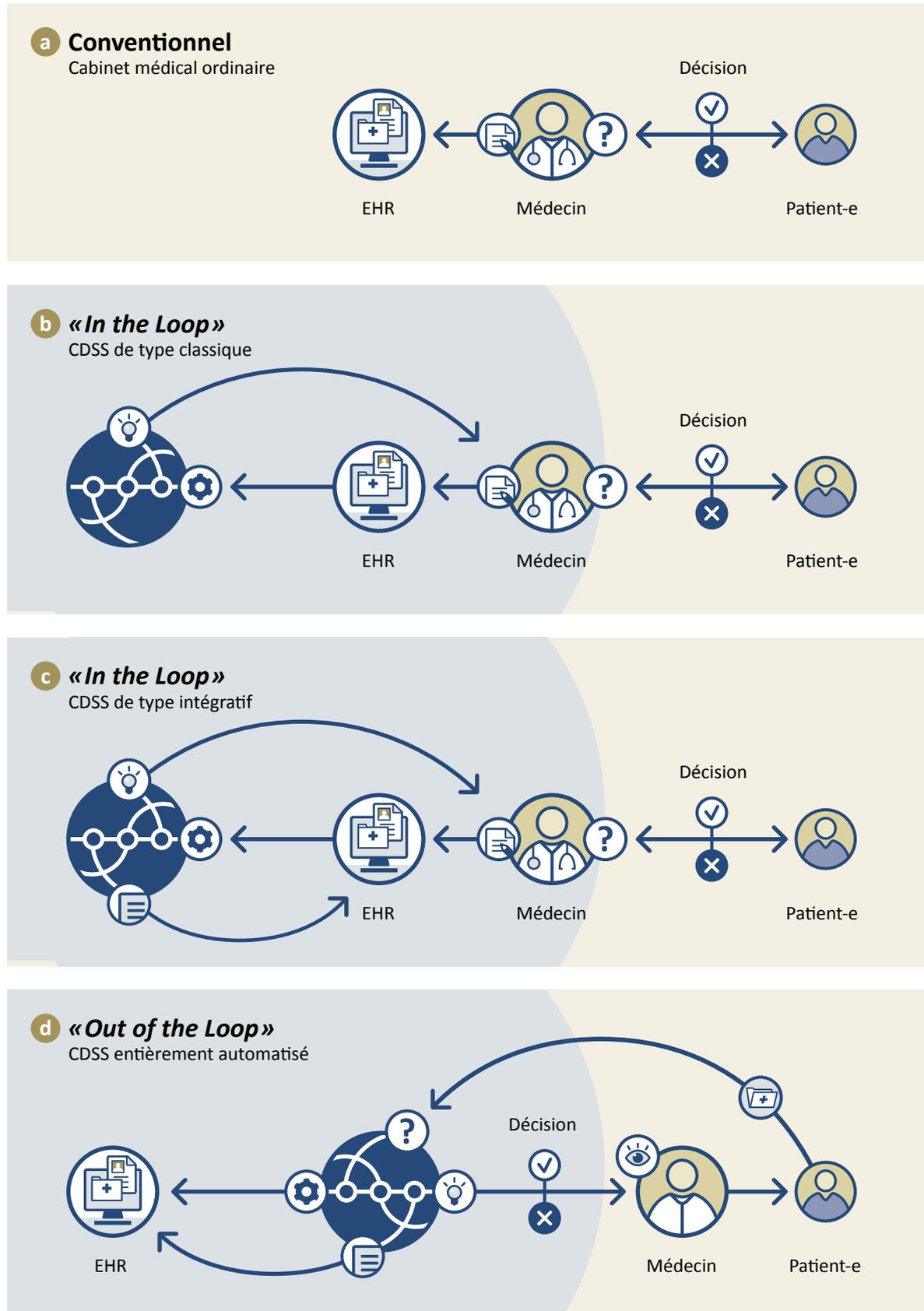


Fig. 6 Modes de collaboration envisageables pour les CDSS «*In the Loop*» et «*Out of the Loop*» avec les médecins



- ❓ Problème ou question pour le médecin (6a-c) ou le système d'IA (6d)
- 📄 Documentation du diagnostic et traitement par le médecin
- ⚙️ Traitement des données dans le système d'IA
- 💡 Solution proposée par le système d'IA en réponse à la question ou au problème posé
- 📄 Documentation des solutions proposées par le système d'IA
- 📄 Collecte d'informations médicales sur les patient-e-s par le système d'IA
- 👁️ Vérification par le médecin de la solution proposée par l'IA
- ✅/❌ Acceptation/refus par le médecin (6a-c) ou par le système d'IA (6d) de la solution proposée par l'IA

En ce qui concerne l'utilisation de l'IA en médecine, on peut concevoir les modes de collaboration «*In the Loop*» et «*Out of the Loop*» entre l'humain et la machine de la manière suivante (ici d'après l'exemple des CDSS Fig. 6):

#### Scénario conventionnel (Fig. 6a)

Dans un cabinet médical ordinaire, le médecin saisit les informations médicales relatives aux patient-e-s, prend des décisions cliniques conjointement avec les personnes concernées, et documente le diagnostic et le traitement à l'aide du dossier médical.

#### «*In the Loop*» (Fig. 6b et 6c)

Les systèmes d'aide à la décision clinique de type classique collectent des informations à partir du dossier électronique de santé (EHR). Après les avoir soumises à des algorithmes basés sur des règles ou sur l'apprentissage automatique, les CDSS présentent au médecin des suggestions diagnostiques et/ou des mesures thérapeutiques. La décision finale incombe toutefois au médecin (6b)<sup>32</sup>.

Dans le cas de CDSS de type intégratif, l'algorithme d'IA peut accéder de manière autonome au dossier électronique de santé, formuler des propositions à l'attention des médecins et les inscrire de manière autonome dans le dossier médical. C'est ici toujours au médecin qu'incombe la décision finale (6c)<sup>32</sup>.

#### «*Out of the Loop*» (Fig. 6d)

Dans les CDSS entièrement automatisés, l'algorithme d'IA recueille des informations médicales relatives aux patient-e-s, prend des décisions médicales de manière autonome et les documente dans le dossier électronique de santé (EHR)<sup>32</sup>.

#### Utilisation croissante de systèmes «*on the loop*»

À l'heure actuelle, il existe peu de systèmes «*Out of the Loop*» comme ceux qui permettent, sans supervision médicale, le diagnostic automatique de la rétinopathie diabétique<sup>31</sup>. Cette rareté est due aux nombreux obstacles techniques et réglementaires qui freinent le déploiement à grande échelle de systèmes d'IA entièrement autonomes. En revanche, on peut s'attendre à ce que les systèmes «*On of the Loop*» parviennent de plus en plus à s'imposer. Une étude a notamment attesté que l'IA peut contribuer au diagnostic pathologique du cancer de la prostate de telle sorte que seulement 25% des cas positifs requièrent une évaluation approfondie. Avec une grande précision, le système d'IA a pu automatiquement éliminer les résultats négatifs, soit 75% des cas étudiés<sup>54</sup>.

Pour le corps médical, les systèmes disponibles dans le commerce font apparaître en outre de nouveaux types d'interactions avec l'intelligence artificielle, également susceptibles d'affecter la situation du médecin en termes d'autonomie et de contrôle<sup>55</sup>. Tandis que la plupart des systèmes disponibles du type «*In the Loop*» assistent directement les médecins, il existe des systèmes apparentés qui fournissent un soutien d'arrière-plan, plutôt indirect.

- Un système d'IA qui permet au médecin de repérer plus facilement une fracture sur une radiographie, en la mettant en évidence, correspond au soutien de type direct<sup>56</sup>.
- En revanche, un système d'IA qui trie les images radiologiques que le médecin va examiner constitue un exemple de soutien de type indirect. Un tel classement préalable peut être effectué en fonction du degré de gravité, d'urgence ou de douleur ... C'est le cas d'un système d'IA qui envoie automatiquement au médecin les clichés sur lesquels un pneumothorax a été détecté, afin qu'ils soient interprétés en priorité<sup>57</sup>.

#### Soutien par IA pour le personnel non médical

D'autres systèmes d'IA permettent de confier des processus médicaux à des tiers. C'est notamment le cas d'un système qui aide le personnel paramédical à réaliser des examens complexes d'échocardiographie en toute autonomie<sup>58</sup>. Les cardiologues peuvent établir un diagnostic sur la base des données ainsi recueillies après qu'elles leur ont été transmises. Cette technologie permet à la fois un gain de temps pour les médecins et un élargissement des compétences pour le personnel de santé non-médecin.

### L'acceptation et le savoir-faire du corps médical en matière d'IA sont indispensables

Pour l'intégration de systèmes dans le quotidien professionnel, l'acceptation de l'intelligence artificielle par le corps médical s'avère essentielle, à l'hôpital comme au cabinet du médecin de famille ou du médecin spécialiste. Il lui incombe de déterminer la manière la plus efficace de collaborer avec la machine intelligente, d'où importance de l'impliquer dans le développement des systèmes d'IA utilisés en médecine.

En plus de la répartition des tâches le système d'IA et l'humain qui l'utilise, il est indispensable de définir les connaissances préalables requises pour l'utilisation correcte du système d'IA, son contrôle et, le cas échéant, sa surveillance. Cela implique que les médecins soient suffisamment informés en ce qui concerne le but, l'utilité, la manipulation (utilisation, contrôle, surveillance) et les limites et risques du système employé. L'acquisition de la capacité nécessaire à l'utilisation de l'IA doit refléter le devoir de diligence des médecins. Ce savoir-faire contribue, entre autres, à éviter le biais d'automatisation, décrit précédemment (voir la dimension «Équité») <sup>46</sup>.

### 3. Transparence

Pertinente pour l'acceptation de l'IA, la notion de transparence comprend des aspects de traçabilité, d'explicabilité, d'interprétabilité et de reproductibilité. On parle de transparence ou de traçabilité lorsqu'un humain est en mesure de comprendre entièrement ou partiellement le fonctionnement et les décisions d'un système d'IA. Lorsqu'il existe une explication sur la manière dont le système d'IA réalise ses prédictions, on parle d'explicabilité. L'interprétabilité du modèle décrit la transparence de la méthode qu'un système d'IA utilise pour élaborer ses prédictions <sup>46,47</sup>.

Par exemple, des modèles basés sur des règles sont qualifiés d'interprétables en raison de leur transparence. En revanche, quoique généralement plus robustes, les modèles de type boîte noire (*black box*) sont à peine interprétables <sup>47</sup>. S'intéressant particulièrement à cet aspect, les spécialistes exigent qu'il soit possible d'expliquer et de comprendre comment les systèmes d'IA parviennent à leurs prédictions. Cette question se révèle délicate dans le contexte des réseaux neuronaux profonds (DNN) <sup>3</sup>. C'est pourquoi la recherche se poursuit dans ce domaine <sup>59</sup>.

### Recommandations et initiatives internationales

Plusieurs initiatives internationales ont abordé les thèmes de la transparence, de l'explicabilité et de la traçabilité. En voici quelques exemples :

- TRIPOD (*Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis*) publie des recommandations en matière de documentation pour la conception d'études et pour le développement, la validation et la mise à jour des modèles qui réalisent des prédictions pronostiques ou diagnostiques au moyen d'IA <sup>60</sup>.
- La checklist MI-CLAIM (*Minimum Information about CLinical Artificial Intelligence Modelling*) doit permettre de documenter de manière transparente les algorithmes d'intelligence artificielle utilisés en médecine. D'une part, cela consiste à pouvoir évaluer leur efficacité clinique, y compris l'équité et les biais. D'autre part, cette liste de vérification doit contribuer à rendre le processus de conception technique des systèmes d'IA cliniques rapidement compréhensible <sup>61,62</sup>.
- CONSORT-AI (*CONsolidated Standards Of Reporting Trials-Artificial Intelligence*) tient lieu de nouvelle ligne directrice pour la documentation des essais cliniques qui évaluent les interventions médicales qui intègrent l'intelligence artificielle <sup>63</sup>.
- SPIRIT-AI (*Standard Protocol Items: Recommendations for Interventional Trials-Artificial Intelligence*), constitue une autre ligne directrice pour les protocoles d'études cliniques, développée en parallèle <sup>64</sup>.

Il s'agit, en fin de compte, de trouver un équilibre satisfaisant entre, d'une part, la robustesse ou fiabilité d'un système d'IA, et, d'autre part, la traçabilité ou transparence de son fonctionnement. En effet, une transparence suffisante doit permettre aux médecins d'utiliser correctement les prédictions des systèmes, compte tenu de leur qualification préalable <sup>46</sup>. Par ailleurs, il convient d'éviter que la documentation de la transparence ne conduise à une surcharge de travail due au cumul de détails sans pertinence <sup>47</sup>.

### Une situation complexe en matière de responsabilité

Les aspects de traçabilité et de reproductibilité des prédictions basées sur l'IA s'avèrent également pertinents pour les questions juridiques. C'est notamment le cas de question de la responsabilité civile et pénale. Tandis qu'une erreur médicale isolée peut nuire à un individu, une erreur commise invariablement par un système d'IA chargé de prédictions diagnostiques ou thérapeutiques relatives à une population entière peut nuire à un grand nombre

de patient-e-s. Indispensable, la prévention d'un tel risque passe par l'élimination systématique des erreurs, les audits, les simulations et validations approfondies, ainsi que les validations prospectives à l'hôpital comme au cabinet médical. Elle doit avoir pour objectif de maintenir le risque d'erreur d'un système d'IA à un niveau très faible dans la pratique clinique<sup>3</sup>.

Qui doit être tenu pour responsable des dommages causés aux patient-e-s lors de l'utilisation d'un système d'IA – le médecin, l'hôpital ou le fabricant? Voici l'état actuel de ce dilemme: en cas de problème en lien avec l'IA dans le cadre d'un traitement médical, l'attribution de la responsabilité entre le médecin traitant et le fabricant (ou tout autre responsable du système d'IA) relève de la répartition entre leurs domaines de compétences<sup>65,66</sup>. Le médecin doit assumer la responsabilité des actions et omissions que ses compétences professionnelles lui permettent d'identifier, d'influencer et de maîtriser<sup>b</sup>. Cela comprend également la décision même de s'appuyer sur des systèmes d'IA dans le cadre du diagnostic et du traitement. En contrepartie, il en va de même de la décision de renoncer à leur utilisation dans le cadre d'un traitement où l'IA correspond aux standards de bonne pratique<sup>c</sup>.

#### **Discussions en cours sur les adaptations possibles du partage des responsabilités**

Un médecin peut également engager sa responsabilité en acceptant une tâche au-delà de ses propres qualifications professionnelles (faute par acceptation de prise en charge)<sup>67</sup>. Les devoirs de diligence des médecins en rapport avec l'IA peuvent varier selon le cas d'utilisation et devraient être définis en conséquence. Ce sont les bases juridiques actuelles qui déterminent l'évaluation de la responsabilité du médecin<sup>68</sup>. En cas d'instauration de nouvelles lois sur l'utilisation de l'IA en médecine, celles-ci devraient être conçues de manière à tenir compte des circonstances relatives à son utilisation. En plus des prétentions en dommages-intérêts ou en réparation du tort moral, il convient d'examiner la responsabilité pénale si, lors de l'utilisation de l'IA, une violation du devoir de diligence par le médecin, intentionnelle ou par négligence, a entraîné un préjudice pour un-e patient-e. On envisage actuellement une révision des bases de la responsabilité, et ses différents aspects font l'objet d'une discussion à l'échelle internationale.

<sup>b</sup> ATF 133 III 121; ATF 130 IV 7, consid. 3.3.

<sup>c</sup> Comparaisons concernant la diligence: ATF 134 IV 175, consid. 3.2.; ATF 130 IV 7, consid. 3.3.

## **4. Fiabilité**

Le terme de fiabilité englobe des aspects en lien avec la qualité du système d'IA. Ils comprennent la performance, la robustesse, l'interception des erreurs et l'évaluation de l'incertitude en fonction du modèle d'IA employé. Il s'agit là de tester le risque d'erreur du système d'IA. Cette vérification clarifie notamment l'exactitude des prédictions du système d'IA:

- lorsqu'il est utilisé de manière conforme à son objectif et à ses exigences;
- lorsqu'il reçoit, afin de mettre sa robustesse à l'épreuve, des données susceptibles de poser problème (p. ex. données à la limite du domaine d'application) ou des données d'entrée manipulées (p. ex. attaques adverses)<sup>46</sup>.

Un système d'IA ne peut élaborer des prédictions correctes qu'au sein du spectre dans lequel il a été entraîné<sup>46</sup>. On trouve souvent en médecine des données non structurées, par exemple sous la forme d'annotations formulées librement par le médecin dans un dossier médical ou dans une base de données<sup>3</sup>. Afin d'entraîner les systèmes d'IA, il faut disposer de jeux de données de haute qualité, représentatifs et diversifiés, en lien avec la population de patient-e-s concernée. Ce n'est qu'à cette condition que ces systèmes peuvent apprendre à parvenir aux bonnes conclusions pour leurs prédictions<sup>39</sup>.

#### **Les systèmes d'IA ne sont que partiellement transposables à d'autres populations de patient-e-s**

Un système d'IA qui a été entraîné, testé et validé dans une clinique sur une certaine population de patient-e-s ne peut pas être automatiquement transféré (généralisabilité ou portabilité) à une autre population dans une autre clinique. Il faut donc effectuer un test avant de pouvoir utiliser un système d'IA pour une nouvelle population de patient-e-s, afin de permettre de déterminer clairement si ses prédictions concernant cette nouvelle population restent bien correctes. Le cas contraire requiert une adaptation<sup>69</sup>. Il convient de vérifier à cet égard s'il existe un surapprentissage (*overfitting*) pour un algorithme: par exemple, si l'échantillon (taille de la population de patient-e-s) s'avère trop petit dans le contexte d'un certain questionnement médical, l'algorithme d'IA (apprentissage automatique) peut y reconnaître des modèles spécifiques (c'est-à-dire des modèles qui ne sont pas universellement valables). Sur la base de ces modèles, l'algorithme risque alors d'élaborer des prédictions trop restreintes qu'il généralisera à l'ensemble des données de manière erronée<sup>3</sup>. La faculté de généralisation de l'algorithme est limitée.

### Les données d'entraînement en tant qu'important facteur de qualité

La qualité et la distribution des données d'entraînement se révèlent également déterminantes pour l'entraînement d'un algorithme d'IA fiable. Or, dans le système de santé, ce sont surtout les données des patient-e-s qui sont collectées, alors que jusqu'à présent, peu de données sont collectées sur les personnes en bonne santé.

## De grands ensembles de données médicales annotées et de haute qualité sont nécessaires afin de former, tester et valider les systèmes d'IA.

Par exemple, l'étude des fonctions cérébrales et de leur altération à l'aide d'imagerie médicale telle que l'IRM nécessite un ensemble de données de référence illustrant des cerveaux «sains», en plus des données pathologiques des patient-e-s. Ces images médicales peuvent être ainsi utilisées pour entraîner et tester les systèmes d'IA utilisés à cette fin. Par ailleurs, l'identification de variantes de gènes potentiellement impliquées dans des maladies requiert également des données fondamentales concernant la population en bonne santé, de sorte à décrire comment est distribuée la fréquence des variations génétiques<sup>70</sup>. Compte tenu de la finalité d'un système d'IA, il s'agit également de définir des critères de qualité, tels que la sensibilité et la spécificité, ainsi que les valeurs à atteindre<sup>3,71</sup>. Ainsi, il est possible de tolérer un taux plus ou moins élevé de faux positifs, en fonction l'objectif poursuivi.

### Validation à l'aide de jeux de données indépendants

La validation des systèmes d'IA requiert des jeux de données indépendants servant à vérifier leur transférabilité à une nouvelle population de patient-e-s (portabilité), ou leur généralisation, leur robustesse, ou encore leur performance. En théorie, outre les données provenant des hôpitaux et des cabinets médicaux, on peut aussi utiliser à cette fin des cohortes ou des données issues de registres cliniques et épidémiologiques. La reproductibilité des performances du modèle est examinée en le validant au moyen d'ensembles de données indépendants, dont les caractéristiques ressemblent à celles des données utilisées pour les tests d'entraînement.

Ce qui permet de distinguer ces ensembles de données dits indépendants comprend notamment leurs divergences en termes de caractéristiques des patient-e-s, de biais potentiels ou de facteurs perturbateurs. Ils peuvent également contenir des données manipulées. La validation montre à quel point le modèle est transférable (généralisation) et robuste. L'identification des risques tout au long du processus de développement du système d'IA en constitue la condition préalable. Il faut également avoir validé les différentes étapes, à savoir la collecte, l'annotation, le test et l'entraînement des données pertinentes pour le système d'IA. Mentionnées dans la section précédente, c'est pour cette validation que se révèlent utiles les normes internationales relatives aux conceptions d'étude assistées par IA, telles que le plan d'analyse pour l'apprentissage automatique<sup>72</sup>.

### La performance n'est pas synonyme d'efficacité clinique

Dans ce contexte, il est également important de clarifier certaines questions. Pourquoi a-t-on utilisé une méthode d'IA particulière plutôt qu'une autre ? Et comment a-t-on comparé la performance de cette méthode d'IA avec la meilleure alternative possible ? À cela s'ajoute le fait que l'efficacité clinique d'un système d'IA (*efficacy*) n'est pas garantie par la seule validation de ces performances. En effet, elle ne considère pas seulement la précision d'un système (*accuracy*), mais également son influence sur le succès du traitement (*patient outcome*)<sup>73,74</sup>. Cela implique qu'un système d'IA doit pouvoir démontrer qu'il améliore les résultats cliniques. C'est pourquoi sa validation dans un environnement clinique réel s'avère indispensable<sup>3</sup>. Enfin, il faut évidemment tenir compte de son économie, c'est-à-dire du rapport coûts/bénéfices et de la rentabilité.

## 5. Sécurité

Le terme de sécurité comprend la protection contre les attaques externes (*security*) et la protection contre les dangers émanant du système d'IA lui-même (*safety*)<sup>47</sup>. En matière de sécurité informatique (*security*), on considère les domaines de l'intégrité et de la disponibilité. On dit que l'intégrité du système d'IA est affectée lors d'une attaque externe, si sa fonctionnalité s'en trouve modifiée ou altérée. C'est notamment le cas de modifications apportées de manière non autorisées ou involontaires, comme lors d'attaque antagoniste (*adversarial attack*) ou d'empoisonnement des données (*data poisoning*)<sup>46</sup>.

Dans le cas d'une attaque antagoniste, une personne malveillante (*hacker*) détermine à sa guise la prédiction du système d'IA, par exemple en manipulant des données. Quant à l'empoisonnement des données, il décrit la manipulation ciblée des données d'entraînement dans l'intention de provoquer des modifications fonctionnelles néfastes. On dit que la disponibilité du système d'IA est affectée lors d'une attaque externe, si celui-ci ne réagit plus ou pas suffisamment<sup>46</sup>.

Le domaine de la sécurité touche aux risques de préjudices matériels ou corporels causés par un dysfonctionnement ou une panne dans un système d'IA. Une conception insuffisante de l'intégration du système d'IA dans son environnement en constitue, entre autres, une origine possible<sup>46</sup>.

## 6. Protection des données

Lors du développement du système d'IA et lors de son exploitation, la protection des données sensibles revêt une grande importance. La catégorie des données sensibles désigne, dans ce contexte, les données personnelles, de même que d'autres données non personnelles (p. ex. secrets commerciaux)<sup>46</sup>. Du point de vue de la protection des données, la responsabilité, la transparence et l'implication de la personne concernée constituent des points particulièrement importants lors de l'utilisation de l'IA. Il convient de clarifier si l'utilisation de l'IA se fait sous la responsabilité individuelle ou dans le cadre d'un traitement de données sur commande (p. ex. sur le cloud) (art. 9 de la nouvelle loi sur la protection des données, LPD).

### Protection obligatoire des systèmes IA contre les risques de sécurité

Pour le médecin, la responsabilité en matière de protection des données est inchangée. En cas d'utilisation de l'IA, le traitement de données sur commande requiert toutefois des dispositions contractuelles spécifiques, notamment relatives à l'utilisation des données par des mandataires pour le développement de systèmes d'IA (art. 9, al. 1, let. A, nLPD). Il convient également de noter qu'un système d'IA doit être protégé de manière adéquate contre les risques liés à la sécurité, par des

mesures techniques et organisationnelles appropriées (art. 8 nLPD). En outre, lorsqu'un traitement des données peut entraîner un risque élevé pour la personnalité ou les droits fondamentaux de la personne concernée, une évaluation des risques (analyse d'impact sur la protection des données) doit en principe être effectuée au préalable (art. 22, al. 1, nLPD).

Les facteurs de risques se situent au niveau de l'empoisonnement des données (*data poisoning*) et de la synthèse de la parole (*speech synthesis*). Il s'agit de manipulation des données ou d'imitation – donc de modification – des entrées linguistiques (voir la dimension «Sécurité»). Les prédictions des systèmes d'IA s'en trouvent modifiées en conséquence. Le danger provient également du piratage automatisé (*hacking*), qui permet de divulguer des données sensibles protégées, telles que des informations médicales personnelles<sup>75</sup>. Il convient de garder à l'esprit que des procédés d'IA spéciaux ou la mise en relation des données générées par le système d'IA avec d'autres données (*data linkage*) sont susceptibles de créer une occasion de dévoiler des identités anonymisées. Ceci peut conduire, pour les personnes concernées, à une violation des droits de la personnalité. C'est pourquoi il est nécessaire de protéger efficacement les données consultées, traitées ou stockées par le système d'IA, lors de son développement et de son exploitation<sup>46</sup>.

**Les données consultées, traitées ou stockées par le système d'IA doivent être protégées de manière efficace.**

### Des exigences plus strictes dans la nouvelle loi sur la protection des données (nLPD)

Dans le contexte d'un traitement réalisé à l'aide d'IA, l'utilisation de données personnelles doit s'accompagner d'un motif justificatif, à savoir un consentement (art. 31 nLPD). L'alternative consiste à travailler avec des données anonymisées. De manière générale, le traitement des données doit être reconnaissable pour les personnes concernées, même en cas d'utilisation de l'IA (art. 6, al. 3, nLPD; art. 19 nLPD). L'utilisation de l'IA peut faire l'objet d'un consentement explicite. En l'absence d'une telle spécification, l'automatisation d'un processus décisionnel

individuel doit être connue de la personne concernée, pour autant qu'elle implique une conséquence juridique ou affecte la personne de manière significative, de même que si la décision est fondée exclusivement sur l'IA (art. 21 nLPD). La personne concernée est également en droit d'obtenir des informations sur la logique qui sous-tend la décision automatisée (art. 25, al. 2, let. F, nLPD).

#### **Nouvelles approches techniques pour le traitement de grands ensembles de données dans le respect de la protection des données**

Appelé *big data*, le regroupement de données personnelles en grands ensembles de données représente un défi en matière de protection des données qui met en évidence les limites du partage de données. Cependant, l'entraînement des modèles requiert que les systèmes d'IA disposent de grands ensembles de données pertinents pour la population de patient-e-s.

C'est pourquoi de nouvelles approches techniques permettent aux systèmes d'IA d'accéder à des ensembles de données distribués sans avoir à les regrouper physiquement. L'apprentissage fédéré illustre cette technologie<sup>76</sup>. Sa première étape consiste à entraîner les modèles d'IA au moyen de différentes données hospitalières locales. Dans une deuxième étape, la division physique des paramètres du modèle permet de centraliser les paramètres locaux du modèle dans un modèle global. Il s'est avéré que les modèles globaux ainsi entraînés fournissent des performances similaires à celles des modèles qui ont un accès direct à toutes les données (c'est-à-dire à un ensemble *big data*).

L'apprentissage en essaim (*swarm learning*) représente une évolution de cette technologie. Il utilise également des ensembles de données distribués pour l'entraînement des modèles d'IA, mais ne requiert pas de modèle centralisé<sup>77</sup>.

Pour ces deux technologies, les dispositions relatives à la protection des données n'ont qu'une portée limitée, car les données ne quittent pas les hôpitaux. En outre, le partage de données ne concerne que des informations abstraites et anonymes (les paramètres du modèle).

### **Information, consentement éclairé et documentation**

---

- **Information et consentement** Il faut s'attendre à ce que l'utilisation de l'IA entraîne à l'avenir des exigences accrues en matière d'obligation d'information concernant les « nouveaux risques inconnus ». Le médecin doit fournir aux patient-e-s une information suffisante, qui leur permette d'exercer leur droit à l'autodétermination et de comprendre l'intervention médicale en question, y compris la contribution de l'IA. Le consentement des patient-e-s ne peut être valable qu'en lien avec une telle information<sup>d,e</sup> le devoir d'information fait l'objet de l'art. 10 du code de déontologie de la FMH<sup>83</sup>).

- **Documentation** Les étapes du traitement, y compris l'utilisation de l'IA, doivent être documentées et inscrites dans le dossier médical, de même que les entretiens d'information. La documentation médicale sert à assurer la thérapie comme à conserver des justificatifs. Le dossier médical doit être rédigé de manière à documenter de manière complète et compréhensible le déroulement du traitement médical. Par conséquent, les faits médicaux essentiels au traitement sont soumis à l'obligation de consigner<sup>f</sup> (voir aussi l'art. 12 du code de déontologie de la FMH<sup>83</sup>).

<sup>d</sup> ATF 133 III 121

<sup>e</sup> ATF 117 Ib 197

<sup>f</sup> ATF 141 III 363

### **La propriété intellectuelle en matière de modèles d'IA**

La protection des données peut également concerner des données qui ne sont pas relatives à des personnes, telles que toutes celles sur lesquelles des tiers disposent de droits (p. ex. les données liées à une licence). En fonction de sa pertinence pour l'entreprise, même un modèle d'intelligence artificielle au sein d'un système peut valoir d'être protégé<sup>46</sup>. La propriété intellectuelle (PI) assure les innovations et les inventions humaines. Or, les systèmes d'IA recèlent également le potentiel de créer de la nouveauté, ce qui soulève des questions en matière de propriété intellectuelle. Comment évaluer les innovations et inventions que l'humain a créées, en comparaison à celles que des systèmes d'IA ont générées? Les directives existantes en matière de propriété intellectuelle doivent-elles être adaptées en ce qui concerne les systèmes d'IA, notamment afin de protéger les inventions provenant de l'IA, protéger l'intelligence artificielle elle-même, ainsi que les données qu'elle utilise<sup>78</sup>?

### **Situation juridique actuelle en matière de propriété intellectuelle**

Conformément au «principe du créateur» généralement applicable dans le domaine de la PI, les droits de propriété intellectuelle reviennent, en l'absence de toute autre base légale, au créateur du bien immatériel concerné. En tant que créateur, seule une personne physique peut être prise en compte. Par conséquent, les biens immatériels créés par des systèmes d'IA sont traditionnellement attribués à des personnes physiques. Pour le droit d'auteur, l'existence d'une création intellectuelle, c'est-à-dire d'une expression de l'esprit humain, constitue même une condition explicite pour la protection. L'accent devrait être mis sur les personnes physiques qui ont programmé le système d'IA et qui l'ont notamment paramétré à l'aide de données d'entraînement et d'autres intrants et consignes. En revanche, selon la situation juridique actuelle, si aucune personne physique ne peut être identifiée comme créateur en raison de la large autonomie du système d'IA, la création ne peut en général pas être soumise à une protection<sup>79</sup>. En raison des incertitudes juridiques actuelles, il est recommandé de veiller à ce que les parties concernées concluent des accords contractuels clairs sur l'attribution des éventuels droits de propriété intellectuelle qui en résultent.

## **Conclusion intermédiaire**

### **Quelle partie prenante humaine doit veiller à ce qu'un système d'IA se montre et demeure digne de confiance?**

Beauchamp et Childress ont défini quatre principes fondamentaux d'éthique médicale: le respect de l'autonomie des patient-e-s, la non-malfaisance, la bienfaisance et la justice<sup>80</sup>. L'Union européenne (UE) a également développé des Lignes directrices en matière d'éthique pour une IA digne de confiance. Celles-ci s'appuient sur quatre principes éthiques: le respect de l'autonomie humaine, la prévention de toute atteinte, l'équité et l'explicabilité<sup>45</sup>. Les principes et les directives poursuivent l'objectif qu'un système d'IA fasse, pour des groupes de patient-e-s comparables, des prédictions identiques (généralisabilité), avec une performance comparable et robuste (précision, sensibilité, spécificité). Un système d'IA doit également attribuer les interventions médicales de manière équitable, indépendamment de facteurs socio-économiques<sup>49</sup>. Même sous une forme telle qu'une personne électronique (*e-person*<sup>81,82</sup>), les systèmes d'IA ne possèdent pas de personnalité au sens juridique. Tant que ceci reste le cas, la responsabilité de leurs décisions revient aux humains qui les conçoivent et les utilisent de manière professionnelle, y compris les professionnel-le-s de la santé. Il leur incombe de veiller à ce que les systèmes d'IA n'aillent pas à l'encontre des principes de l'action humaine (voir aussi les art. 2-4 du code de déontologie de la FMH<sup>83</sup>).

# Comment l'intelligence artificielle est-elle réglementée en médecine?

Jusqu'à présent, il n'existe pas de lois ou de directives internationales uniformes pour réglementer l'utilisation de l'IA en médecine.

Diverses autorités et organisations sont toutefois en train d'élaborer des réglementations et des prescriptions correspondantes, ou d'adapter les lois et les normes existantes.

En 2019, la FDA a publié un document de discussion dans lequel elle présente un cadre réglementaire pour les logiciels basés sur l'intelligence artificielle et l'apprentissage automatique (ML) dans les dispositifs médicaux<sup>84</sup>. Un plan d'action concernant ces applications a été formulé sur la base des retours d'information<sup>85</sup>. Il thématise notamment la manière dont les systèmes d'IA pourraient être gérés s'ils poursuivaient sans cesse leur apprentissage. Tous les dispositifs médicaux à base d'IA qui ont été jusqu'à présent approuvés par la FDA utilisent des algorithmes verrouillés (*locked algorithms*), qui n'apprennent pas indéfiniment<sup>84</sup>. Toutes les technologies à base d'IA/ML approuvées par la FDA se trouvent dans une base de données en libre accès (*open access*)<sup>55</sup>. En 2021, la FDA a publié dix principes directeurs pour le développement de dispositifs médicaux basés sur l'apprentissage automatique<sup>86</sup>.

L'UE a mis en consultation, en avril 2021, une proposition d'harmonisation des prescriptions concernant l'IA (législation sur l'intelligence artificielle)<sup>87</sup>. En novembre 2021, la Commission a adopté cette proposition avec des modifications<sup>88</sup>.

## Le cadre légal et réglementaire relatif à l'utilisation de l'intelligence artificielle n'en est qu'à ses débuts.

Le document définit l'intelligence artificielle, ainsi que les possibles méthodologies sous-jacentes (art. 3 et annexe I<sup>87</sup>). L'UE classe les systèmes d'IA en fonction de leur risque (inacceptable, élevé, faible, minimal). Ses exigences sont définies différemment en fonction du groupe de risque. Les systèmes médicaux basés sur l'IA sont classés dans le groupe à haut risque: s'ils correspondent à des dispositifs médicaux des classes Im, Is, IIa, IIb et II (seule la classe I n'en fait pas partie) conformément au RDM<sup>89</sup>; s'ils relèvent du règlement sur les dispositifs médicaux de diagnostic in vitro (art. 6, annexe II, RDIV<sup>87</sup>); s'ils sont utilisés pour le triage des services d'urgence et la priorisation des secours (annexe III<sup>87</sup>).

La Commission européenne a ajouté un nouvel article (art. 52a) et un nouveau considérant (lié à l'art. 70a) servant à préciser que les systèmes d'IA à usage général ne doivent pas être tenus pour conformes à leur destination au sens du présent règlement<sup>88</sup>. La proposition de l'UE prévoit également la création d'un Comité européen pour l'intelligence artificielle. Celui-ci sera composé de personnes représentant les autorités de contrôle nationales compétentes en matière d'IA, ainsi que du Contrôleur européen de la protection des données et de la Commission. Ce Comité garantira une mise en œuvre harmonieuse, efficace et cohérente du règlement sur l'IA<sup>90</sup>. En 2021, l'UE a également publié une étude comparative des lois relatives à la responsabilité civile en cas d'utilisation de l'IA<sup>91</sup>.

Il existe par ailleurs un certain nombre d'autres recommandations et normes internationales concernant différents aspects de l'utilisation des systèmes d'IA et des systèmes spécialement conçus pour l'apprentissage automatique<sup>92</sup>. En outre, une équipe interdisciplinaire a élaboré un guide à l'intention des fabricants de systèmes d'IA et des organismes d'audit chargés d'évaluer les systèmes d'IA en tant que dispositifs médicaux. Ce guide présente des instructions utiles pour développer des systèmes d'IA en conformité avec les règlements et obtenir leur homologation en tant que dispositifs médicaux au sein de l'espace européen. De plus, il propose des listes de vérification pour contrôler la conformité réglementaire des systèmes d'IA<sup>93</sup>.

### **L'ordre juridique également contraignant en cas d'utilisation de l'IA**

La Coalition internationale des autorités de réglementation des médicaments (ICMRA) a élaboré deux études de cas hypothétiques en matière d'IA, afin d'illustrer certains des défis que pose l'utilisation de l'IA pour la réglementation mondiale des médicaments. Sur la base de ces études, elle a examiné si les réglementations existantes concernant l'IA devaient être adaptées, et, le cas échéant, dans quelle mesure<sup>94</sup>. Les mêmes conclusions s'appliquent pour la Suisse: l'utilisation de l'IA doit respecter, dans tous ses aspects, le droit en vigueur. Il convient d'adapter et/ou de compléter l'ordre juridique là où il ne répond plus aux exigences des nouveaux développements technologiques<sup>85</sup>. C'est notamment le cas dans utilisation de l'IA sur le mode «*out of the loop*», ou lorsque les résultats issus de l'IA ne peuvent plus être attribués à un humain. En Suisse, le débat actuel pose la question des approches juridiques à suivre pour réglementer les systèmes d'IA<sup>95,96,97</sup>. Les limitations possibles, voire nécessaires, entre ce qui est techniquement faisable et les aspects éthiques ou financiers devraient alimenter les discussions à l'échelle sociétale.

# Quelles sont les exigences de la FMH en matière d'intelligence artificielle dans le domaine médical ?

L'arrivée de l'intelligence artificielle dans la médecine ouvre des perspectives intéressantes pour l'activité du corps médical. Toutefois, l'utilisation de l'IA dans la pratique médicale quotidienne soulève également de nombreuses questions. C'est pourquoi la FMH a établi un catalogue d'exigences qui doit contribuer à une utilisation judicieuse et adéquate de l'IA.

Dans les années à venir, l'intelligence artificielle devrait profondément modifier la façon de penser et d'agir des médecins. L'IA doit permettre de développer et d'améliorer de manière judicieuse et adéquate tous les processus qui jalonnent le parcours du patient. Afin de mieux maîtriser les défis qui en découlent, la FMH a formulé **dix exigences**, pertinentes du point de vue du corps médical.

Ces **dix exigences** concernent les systèmes qui sont employés à des fins diagnostiques ou thérapeutiques, étant donné qu'ils détiennent le potentiel de remplacer les prestations médicales de base. L'accent est mis sur les systèmes utilisant l'intelligence artificielle qui, selon le Règlement relatif aux dispositifs médicaux (RDM), se réfèrent à la classe IIa, ou à une classe supérieure.

Les lois, réglementations et directives suivantes, existantes ou proposées, ont été prises en compte : RDM<sup>89</sup>, ODim<sup>98</sup>, RDIV<sup>99</sup>, législation sur l'IA<sup>87</sup>, *Guideline for AI for medical products*<sup>93</sup>, *Guiding principles for good machine learning practice for medical device development*<sup>86</sup>.

**1. Le système d'IA doit renforcer, et non remplacer, l'intelligence humaine et la relation entre les médecins et leurs patient-e-s**<sup>100</sup>.

**2. Le système d'IA s'inspire des principes de la médecine fondée sur les preuves (evidence-based medicine, EbM)**<sup>101</sup>.

**3. Les prestations d'un système d'IA doivent être efficaces et adéquates. Elles visent à simplifier les processus administratifs et à en décharger les médecins.**

**4. Le système d'IA doit être contrôlé périodiquement, et des mesures correctives doivent être apportées immédiatement si le contrôle en démontre la nécessité.**

Un suivi régulier du système d'IA doit être effectué (surveillance post-marketing), notamment afin d'identifier de nouveaux risques. Les performances et le rapport bénéfices/risques peuvent ainsi être vérifiés dans l'environnement clinique réel, et le suivi permet de continuellement les améliorer, si nécessaire. En effet, un système d'IA ne peut notamment élaborer de solides prédictions que pour une population de patient-e-s présentant un certain nombre de caractéristiques pour lesquelles il a été entraîné, testé et validé.

**5. Les médecins restent, en accord avec leurs patient-e-s, les décideurs en ce qui concerne l'application et la gestion des propositions générées par le système d'IA.**

Même sous une forme telle qu'une personne électronique (*e-person*<sup>81,82</sup>), les systèmes d'IA ne possèdent pas de personnalité au sens juridique. Tant que ceci reste le cas, la responsabilité de leurs décisions revient aux humains qui les conçoivent et les utilisent de manière professionnelle, y compris les professionnel-le-s de la santé.

Même dans le cas des systèmes «*out of the loop*», où l'intervention humaine n'est plus nécessaire, le médecin peut encore décider d'utiliser ou non le système d'IA, même s'il ne peut plus influencer directement les prédictions et les décisions que ce système d'IA accomplit de manière autonome.

**6. Chaque système d'IA doit être accompagné d'un « mode d'emploi » destiné aux médecins.**

Ce mode d'emploi doit fournir toutes les informations nécessaires à l'utilisation conforme du système d'IA. Il doit permettre aux médecins de comprendre et de surveiller le fonctionnement du système d'IA et, si nécessaire, d'outrepasser ses décisions. Le fabricant doit veiller à ce que, lors de l'utilisation du système d'IA, les médecins soient en mesure de satisfaire à leur devoir d'information et de diligence.

**Le « mode d'emploi » doit indiquer de manière claire :**

*... le rôle attribué au médecin dans le fonctionnement du système.*

Il faut notamment définir leurs compétences en matière de décision selon les systèmes « *in the loop* » ou « *on the loop* ».

*... les objectifs ou tableaux cliniques dans lesquels le système d'IA peut être utilisé, ainsi que les cas dans lesquels il ne doit pas l'être.*

*... comment utiliser le système d'IA.*

Le mode d'emploi doit notamment indiquer la manière de saisir les données (et lesquelles), de reconnaître les dysfonctionnements et incidents particuliers ou de savoir si le système fonctionne correctement, ainsi que la manière de refuser et remplacer, si nécessaire, les décisions du système.

*... la méthodologie sur laquelle se base le système d'IA.*

Les principes de base et les normes d'excellence de la compétence en matière de données doivent être respectés et mis en œuvre.

Cela comprend notamment un concept de données judicieux, déclaré de manière transparente et compréhensible, ainsi qu'une collecte et un traitement des données qui correspondent à ce concept.

*... comment le système d'IA a été entraîné, testé et validé.*

Le mode d'emploi doit notamment indiquer les ensembles de données indépendants utilisés et leurs caractéristiques (taille, répartition par sexe, par âge, etc.). Il doit en outre décrire leur adéquation (au niveau statistique et médical/clinique), y compris la description des tests relatifs à d'éventuels biais et erreurs. De plus, il doit mentionner les chiffres clés en ce qui concerne la précision de la performance et de l'efficacité clinique du système d'IA, également en comparaison avec les meilleures méthodes alternatives et les meilleurs modèles d'IA possibles.

*... la manière dont le système d'IA garantit la protection des données et la sécurité.*

En particulier, le mode d'emploi doit préciser où et sous quelle forme les données (données personnelles et métadonnées) sont stockées, comment ces informations sont protégées de manière adéquate contre les risques de sécurité, et comment est assurée la sécurité du système d'IA.

*... auprès de quel point de contact le médecin peut s'adresser.*

Il faut s'assurer que l'utilisatrice ou l'utilisateur reçoive une aide compétente en cas de questions ou de problèmes.

**7. En vue de l'application et de la gestion des systèmes d'IA, les médecins doivent disposer d'une offre appropriée de formation initiale, postgraduée et continue.**

Le système de formation médicale doit donner aux futurs médecins et aux médecins en exercice les outils leur permettant d'utiliser l'IA et les techniques modernes de manière pertinente et compétente dans leur quotidien médical. Cela inclut la confrontation avec les aspects juridiques et éthiques<sup>102</sup>, de même que la capacité à utiliser les données de manière compétente, ainsi qu'à comprendre et à accompagner les différents aspects de la médecine basée sur les données<sup>103,104,105,106</sup>. L'intégration des notions de base de l'analyse des données, y compris concernant l'apprentissage automatique, se révèle indispensable pour pouvoir utiliser l'IA en connaissance de cause.

**8. Les médecins et les patient-e-s doivent être associés au développement des systèmes d'IA afin de participer à la définition des exigences concernant leur utilisation.**

Pour que cette utilisation puisse atteindre ses objectifs en matière d'efficacité, il faudra notamment déterminer quelles questions médicales doivent être résolues par le système d'IA et dans quel contexte le système AI doit être utilisé.

**9. Les médecins doivent être informés de l'utilisation de systèmes d'IA qui influencent indirectement leur travail.**

C'est notamment le cas des systèmes qui peuvent surveiller la qualité du travail des médecins en arrière-plan.

**10. Une infrastructure de données doit être créée à l'échelle nationale, afin de permettre d'entraîner et de tester les systèmes d'IA au moyen de vastes ensembles de données médicales annotées de haute qualité, et afin d'assurer la reproductibilité des résultats par une validation robuste.**

Les systèmes d'IA doivent être adaptés à la situation locale des données (voir l'exigence 4). Pour ce faire, il est nécessaire de disposer d'ensembles de données représentatifs de la population locale de patient-e-s à tester. Ceux-ci doivent être collectés, traités et stockés dans le respect de la protection et de la sécurité des données.

Dans ce contexte, il est important de disposer d'une gestion de la qualité certifiée, ainsi que d'une annotation certifiée des données utilisées et collectées.

La Suisse a en outre besoin d'accéder aux plateformes de données internationales<sup>h</sup>. Le corps médical suisse peut également s'appuyer ici sur de nouvelles solutions, telles que l'apprentissage fédéré, permettant un apprentissage décentralisé des données.

La FMH n'est pas seule à avoir formulé des exigences. Le Comité permanent des médecins européens (CPME) et l'Association médicale américaine (AMA) ont également défini des exigences minimales pour les systèmes d'IA<sup>108,109</sup>.

En Suisse, des ateliers thématiques sur les revendications en matière d'IA ont également eu lieu, notamment intitulés «Régulation des algorithmes décisionnels pilotés par IA» (*IRGC, International risk governance center*) et «Artificial Intelligence in medicine – can it be trusted?» (ASSM, Académie suisse des sciences médicales; ASST, Académie suisse des sciences techniques)<sup>110,111</sup>.

<sup>h</sup> On peut citer l'exemple du Royaume-Uni, qui collecte des données auprès des médecins de famille dans le cadre du réseau *Clinical Practice Research Datalink (CPRD)*<sup>107</sup>

# Quelle sera l'évolution du profil professionnel du médecin ?

L'arrivée de l'intelligence artificielle en médecine provoque à moyen terme un changement profond du profil professionnel du médecin. Toutefois, il est peu probable qu'un jour l'intelligence artificielle assume pleinement ses rôles.

L'IA a le potentiel d'améliorer la qualité des soins médicaux. Elle peut à la fois réduire les erreurs humaines et minimiser la charge de travail des médecins pour les tâches de routine, notamment dans le domaine administratif. Ainsi, si l'on parvient à intégrer avec succès l'IA dans le quotidien clinique pour les tâches de routine administratives et médicales, cela pourrait permettre aux médecins de consacrer plus de temps aux activités exigeantes et au contact personnel avec leurs patient-e-s<sup>39</sup>. Le profil professionnel du médecin s'en trouverait effectivement fortement modifié. La question est de savoir dans quelle mesure.

Il convient tout d'abord de considérer les rôles et, par conséquent, les compétences qu'un médecin doit utiliser au quotidien ou acquérir au cours de sa formation afin d'exercer au mieux cette profession.

## **Définition des différents rôles du médecin d'après le référentiel CanMEDS**

Une définition de ces rôles a été fournie avec le Cadre de compétences des médecins intitulé *Canadian medical education directives for specialists* (CanMEDS, *Physician competency framework*<sup>112</sup>). Le référentiel CanMEDS décrit les rôles qu'un médecin devrait jouer et les compétences nécessaires pour les assumer. Le catalogue suisse des objectifs d'apprentissage pour les étudiant-e-s en médecine (PROFILES) a intégré sa structure de base.

D'après CanMEDS, le médecin peut incarner sept rôles: expert médical, communicateur, collaborateur, gestionnaire (*leader*), promoteur de la santé, érudit et professionnel. Le rôle de l'expert médical y représente l'intégration des six autres rôles.

Aucun de ces rôles n'est défini d'un point de vue exclusivement technique. Il est donc peu probable que des machines intelligentes puissent les assumer entièrement<sup>113</sup>. Toutefois, les machines intelligentes disposent d'une puissance de calcul et d'une capacité de stockage presque illimitées leur permettant de traiter extrêmement rapidement des ensembles de données multimodales extrêmement volumineux et d'appliquer des algorithmes complexes. Elles sont donc capables de tester des hypothèses complexes, et ce, à toute heure du jour ou de la nuit. Par rapport aux humains, elles peuvent traiter les informations et prendre des décisions de manière plus systématique, et réagir plus rapidement aux changements<sup>2</sup>.

**Il y a des choses que les machines intelligentes font mieux que l'humain. D'autres non. La technologie ne peut pas remplacer les relations humaines.**

Étant donné qu'une prise en charge complète des rôles médicaux par des machines intelligentes n'est guère réaliste, la question se pose de savoir quelles compétences quotidiennes des médecins pourraient être transférées aux systèmes d'IA et quelles en seraient les répercussions sur les rôles de la profession médicale.

### Définition des rôles d'après CanMEDS<sup>112</sup>

Terme	Définition
<b>Expert médical</b> ( <i>medical expert</i> )	L'expert médical tient un rôle central de la profession de médecin et se rapporte au domaine d'activité clinique. En tant que professionnel de la santé, le médecin combine tous les autres rôles de CanMEDS en appliquant ses connaissances médicales, ses compétences cliniques et ses valeurs professionnelles à la prestation de soins de qualité et sûrs, axés sur les patient-e-s.
<b>Communicateur</b> ( <i>communicator</i> )	En tant que communicateur, le médecin établit des relations avec les patient-e-s et leur entourage proche, tel que les familles. Cette relation facilite la collecte et l'échange d'informations importantes pour des soins de santé efficaces.
<b>Collaborateur</b> ( <i>collaborator</i> )	En tant que collaborateur, le médecin coopère efficacement avec d'autres professionnels-le-s de la santé, afin d'assurer la sécurité et la qualité de soins centrés sur les patient-e-s.
<b>Gestionnaire</b> ( <i>leader</i> )	En tant que gestionnaire, le médecin s'engage afin de promouvoir collectivement la vision d'un système de santé de haute qualité. Par son activité en tant que médecin, administrateur, érudit ou enseignant, le gestionnaire assume la responsabilité de fournir d'excellents soins aux patient-e-s.
<b>Promoteur de la santé</b> ( <i>health advocate</i> )	En tant que promoteur de la santé, le médecin apporte son expertise et son influence en coopérant avec d'autres parties prenantes ou groupes de patient-e-s dans l'objectif d'améliorer le système de santé. Il collabore avec les patient-e-s afin de déterminer et de comprendre leurs besoins. Le promoteur de la santé doit représenter d'autres personnes lorsque cela s'avère nécessaire, et il soutient la mobilisation des ressources visant à apporter des changements.
<b>Érudit</b> ( <i>scholar</i> )	En tant qu'érudit, le médecin fait preuve d'un engagement durant sa vie entière, de sorte à viser l'excellence pratique au moyen d'un apprentissage continu. Ce faisant, l'érudit instruit également d'autres médecins, évalue leurs connaissances et contribue à leur formation.
<b>Professionnel</b> ( <i>professional</i> )	En tant que professionnel de la santé, le médecin est tenu de veiller à la santé et au bien-être de chaque patient-e, comme de l'ensemble de la société. Il se fonde pour cela sur un comportement éthique, une conduite personnelle intègre et vertueuse, la responsabilité envers le corps médical et la société, une régulation au service de la santé et le maintien de sa propre santé.

### Le nouveau rôle du médecin ...

#### *... en tant qu'expert médical:*

Comme nous l'avons mentionné, la force des systèmes d'IA réside dans leur capacité à traiter de très grandes quantités de données d'une grande complexité. D'une part, ils font appel à l'expertise médicale; d'autre part, ils génèrent par leurs propres moyens des connaissances expertes destinées aux médecins. Ces deux points constituent des aspects partiels du rôle d'expert médical. Or, l'essence même de ce rôle n'est pas seulement d'assimiler des connaissances, tel qu'un système d'IA en est

partiellement capable, mais également d'intégrer tous les autres rôles professionnels du médecin (communicateur, collaborateur, leader, promoteur de la santé, érudit, professionnel). C'est pour cela que l'IA ne pourra probablement jamais assumer entièrement ce rôle.

#### *... en tant que communicateur et collaborateur:*

Le rôle du communicateur va gagner en importance. L'appréciation positive ou négative des patient-e-s en ce qui concerne l'utilisation de l'IA pour soutenir leur santé pourrait refléter les opinions les plus diverses. C'est pourquoi il incombe au médecin

de répondre aux questions individuelles, notamment en ce qui concerne la protection des données ou la traçabilité, ainsi que de fournir des informations de manière neutre et ouverte<sup>113</sup> (voir aussi « Information, consentement éclairé et documentation » dans le chapitre « Quels sont les défis que l'intelligence artificielle présente en médecine ? »).

Étant donné que la numérisation dans le domaine de la santé rend les connaissances médicales plus accessibles, l'expertise n'est plus exclusivement réservée aux médecins. En effet, les psychologues, le personnel infirmier et d'autres professionnel-le-s de la santé peuvent désormais également y avoir accès, de même que les patient-e-s. Les systèmes d'IA soutiennent de plus en plus la disponibilité des connaissances médicales pour ces groupes de personnes. C'est le cas notamment du système d'échographie piloté par IA ou de la montre intelligente (*smartwatch*) pour la détection des arythmies. Le personnel paramédical et les patient-e-s peuvent ainsi se forger des opinions indépendantes au sujet d'un état de santé. Cet aspect entraînera des conséquences importantes sur les rôles du médecin en tant que collaborateur et communicateur<sup>113</sup>.

#### *... en tant que gestionnaire et promoteur de la santé :*

En tant que gestionnaire, le médecin devrait collaborer avec les patient-e-s, dans l'objectif de promouvoir, pour leur bien, une utilisation transparente de l'IA en médecine. En tant que promoteur de la santé, le médecin peut contribuer à déterminer si les diverses utilisations médicales de l'IA réduisent ou renforcent les injustices sociales. Le cas échéant, il convient de défendre les intérêts des groupes de personnes défavorisées.

#### *... en tant qu'érudit :*

Une compétence élevée des médecins en matière de technologie numérique est un préalable indispensable à la perception du changement de tous ces rôles du médecin, c'est-à-dire à l'utilisation judicieuse de l'IA. Les médecins doivent avoir la capacité et la volonté de se former en permanence, au sujet de l'origine, du traitement, de l'intégration des données de santé et de leur gouvernance, ainsi que de la méthodologie de l'IA, son éthique et son évaluation et interprétation critiques<sup>113,4</sup>. Cette disposition à l'apprentissage représente une tâche essentielle dans le cadre de l'évolution du rôle du médecin en tant qu'érudit.

#### *... en tant que professionnel :*

En tant que professionnel de la santé, le médecin devrait dorénavant accepter le soutien de l'IA dans certains domaines partiels. Cela lui permet de se concentrer sur ses tâches principales ou sur le développement des compétences complémentaires qu'un système d'IA ne peut justement pas assumer – du moins dans un proche avenir<sup>113</sup>.

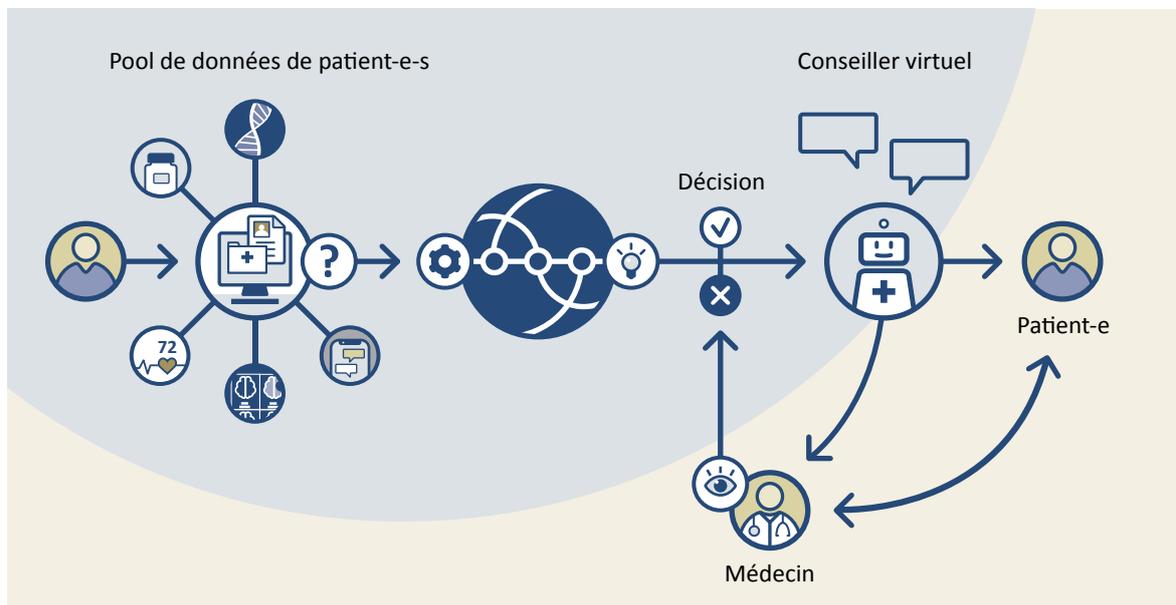
#### **Le médecin en tant que diffuseur d'informations, interprète et conseiller**

Pour le profil professionnel du médecin, l'arrivée de l'IA en médecine signifie que les soins médicaux deviendront de plus en plus une interaction entre médecins, professionnel-le-s de la santé, systèmes d'IA et patient-e-s. À l'avenir, c'est de plus en plus à l'endroit où se trouvent les patient-e-s que seront fournies les prestations médicales, c'est-à-dire dans de nombreux endroits différents. Le développement accéléré de la télémédecine durant la pandémie de Covid-19 illustre bien cette évolution.

La technologie médicale se déplacera avec les individus, voire en eux, fournissant ainsi des flux continus de données sur les patient-e-s qu'il s'agira de classer<sup>114</sup>. Cela donnera naissance à des disciplines médicales spécifiques qui, à leur tour, modifieront la composition de l'« équipe de soins » en termes de spécialisation du personnel. La conception compétente, appropriée et efficace de l'interaction entre l'humain et la machine intelligente apportera également sa contribution dans ce contexte. Dans un avenir proche, les médecins développeront leurs nouveaux rôles de sorte à compléter les domaines que les systèmes d'IA ne peuvent pas prendre en charge. Ils tiendront le rôle de pivot dans le traitement des informations médicales, techniques, réglementaires et éthiques, ainsi qu'un rôle d'interprète et de conseiller auprès des patient-e-s<sup>39</sup>.

On affirme également que, dans les vingt prochaines années, des conseillers médicaux virtuels permettront de fournir des conseils de santé personnalisés globaux et de prévenir les maladies. Ces machines pourront accéder aux données médicales et personnelles les plus diverses (telles que les analyses génomique ou microbiologique et les données sur l'alimentation et l'activité physique) et les évaluer à l'aide de l'IA (Fig. 7). Cette vision devrait devenir une réalité en combinant l'analyse du génome avec la médecine numérique, l'intelligence artificielle et la robotique<sup>3,4</sup>.

Fig. 7 Action du conseiller médical virtuel: saisie des données multimodales et leur exploitation avec l'aide de l'IA



- ❓ Problème ou question pour le système d'IA
- ⚙️ Traitement des données dans le système d'IA
- 💡 Solution proposée par le système d'IA en réponse à la question ou au problème posé
- 👁️ Vérification par le médecin de la solution proposée par l'IA
- ✔️ ✖️ Acceptation/refus par le système d'IA de la solution proposée par l'IA

### Réorientation de la formation des médecins

Dans tous les cas, le système de formation médicale doit fournir aux médecins en formation et à ceux qui exercent déjà leur activité les outils qui donneront leur pertinence à l'IA et aux techniques modernes. Cela inclut également le cadre juridique et éthique<sup>102</sup>. Actuellement, durant leurs études prégraduées, les médecins sont préparés aux rôles définis par le référentiel CanMEDs. Ceux-ci reflètent toutefois une pratique de la médecine largement antérieure à l'arrivée des systèmes d'IA. Un nouveau rôle apparaît et les médecins devront désormais également y être préparés: celui d'analyste de données. Ce rôle inclut la capacité de comprendre et d'accompagner les différents aspects d'une médecine guidée par les données. Il s'agit notamment d'assimiler les notions de base de l'analyse des données, telles que des informations sur l'apprentissage automatique, afin de pouvoir utiliser l'IA en connaissance de cause.

Des efforts sont en cours pour réformer les études de médecine dans ce sens. Le réseau suisse de formation médicale a ainsi défini, pour les études de médecine, des objectifs d'apprentissage en lien avec cette thématique<sup>104,103</sup>. En outre, certains

établissements suisses de formation médicale proposent déjà des offres spécifiques dans le cadre de la formation initiale, postgraduée et continue<sup>115,116</sup>. La FMH se penchera également sur le thème des «compétences numériques» (*Digital skills*) et mettra à disposition du corps médical du matériel d'information correspondant.

### En conclusion

L'image de la profession médicale va forcément se trouver modifiée par l'intelligence artificielle – reste à savoir à quelle vitesse se produira cette évolution. Les médecins devraient participer activement à ce développement et accompagner ce processus avec une attitude positive. Et ceci, selon la devise «**Intelligence artificielle for future? — Oui, mais avec notre contribution!**»

## Remerciements

La FMH remercie le Professeur Michael Krauthammer, la Professeure Nikola Biller-Andorno, et la Professeure Brigitte Tag pour leurs revues et leurs contributions extrêmement précieuses. Nous remercions le Professeur Jörg Goldhahn pour ses informations sur l'évolution du profil professionnel des médecins, ainsi que le Professeur Christian Johner pour la discussion sur les défis et les exigences de l'IA en médecine; leurs contributions nous ont été des plus utiles. Nous remercions le Professeur Roland Wiest pour ses précieuses contributions initiales. Un grand merci à Monsieur Walter Kern pour la simplification linguistique de la version allemande. Nous remercions chaleureusement l'Atelier Richner pour l'excellente réalisation graphique. Nous remercions également les groupes de travail de la FMH pour leurs précieuses indications et discussions sur ce thème. Nous remercions enfin les organisations de médecins affiliées à la FMH pour leurs précieuses réactions au chapitre «Quelles sont les exigences de la FMH en matière d'intelligence artificielle dans le domaine médical».

# Références bibliographiques

1. Densen, P. Challenges and opportunities facing medical education. *Trans. Am. Clin. Climatol. Assoc.* 122, (2011).
2. Liaw, W. & Kakadiaris, I. A. Artificial intelligence and family medicine: Better together. *Family Medicine* vol. 52 (2020).
3. Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine* vol. 25 44–56 (2019).
4. NHS. The Topol Review. Preparing the healthcare workforce to deliver the digital future. An independent report on behalf of the Secretary of State for Health and Social Care. (2019).
5. FMH. Digital Trends Survey 2021. [www.fmh.ch/digital-trends-survey-2021-fr](http://www.fmh.ch/digital-trends-survey-2021-fr) (2021).
6. OECD. Artificial Intelligence in Society. [www.oecd.org/publications/artificial-intelligence-in-society-ee-dfee77-en.htm](http://www.oecd.org/publications/artificial-intelligence-in-society-ee-dfee77-en.htm) (2019).
7. Mccarthy, J. What is artificial intelligence? [www-formal.stanford.edu/jmc](http://www-formal.stanford.edu/jmc) (2007).
8. Winston, P. Artificial Intelligence. *Addison-Wesley, Reading, MA* (1992).
9. Society for Automotive Engineers. SAE J3016-Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. *SAE International* (2021).
10. ADAC. So funktioniert ein automatisiertes Auto. [www.adac.de/rund-ums-fahrzeug/ausstattungs-technik-zubehoer/autonomes-fahren/grundlagen/auto-automatisiert-funktion](http://www.adac.de/rund-ums-fahrzeug/ausstattungs-technik-zubehoer/autonomes-fahren/grundlagen/auto-automatisiert-funktion) (2020).
11. Russell, S. & Norvig, P. Artificial Intelligence A Modern Approach SECOND EDITION Prentice Hall Series in Artificial Intelligence (2003).
12. Domingos, P., Kok, S., Poon, H., Richardson, M. & Singla, P. Unifying Logical and Statistical AI. [www.aaai.org/Papers/AAAI/2006/AAAI06-001.pdf](http://www.aaai.org/Papers/AAAI/2006/AAAI06-001.pdf)
13. Molnar, C. Interpretable machine learning. A Guide for Making Black Box Models Explainable. [christophm.github.io/interpretable-ml-book](http://christophm.github.io/interpretable-ml-book) (2019).
14. Blarer, A., Buffat, M., Busch, C., Egloff, D., Fanzun, J., Haefliger, G., Langer, P., Loison, B., Luder, T., Malz, A., Scheidegger, E., Schneider, T., Schöll, M., Schwaar, P., Stämpfli, M., Weber, V. Herausforderungen der künstlichen Intelligenz. Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat (2019).
15. Taddy, M. & Booth, C. The Technological Elements of Artificial Intelligence. [www.nber.org/papers/w24301](http://www.nber.org/papers/w24301) (2018).
16. Krzyk, K. Towards Data Science. Coding Deep Learning For Beginners. [towardsdatascience.com/coding-deep-learning-for-beginners-types-of-machine-learning-b9e651e1ed9d](https://towardsdatascience.com/coding-deep-learning-for-beginners-types-of-machine-learning-b9e651e1ed9d) (2018).
17. IBM Cloud Education. Supervised vs. Unsupervised Learning: What's the Difference? [www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning](http://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning) (2015).
18. IBM Cloud Education. Machine Learning. <https://www.ibm.com/cloud/learn/machine-learning> (2020).
19. Noorbakhsh-Sabet, N., Zand, R., Zhang, Y. & Abedi, V. Artificial Intelligence Transforms the Future of Health Care. *American Journal of Medicine* vol. 132 (2019).
20. Johner, C. Künstliche Intelligenz in der Medizin. [www.johner-institut.de/blog/regulatory-affairs/kuenstliche-intelligenz-in-der-medizin/#section\\_scroll2](http://www.johner-institut.de/blog/regulatory-affairs/kuenstliche-intelligenz-in-der-medizin/#section_scroll2) (2020).
21. Dalton, L., Ballarin, V. & Brun, M. Clustering Algorithms: On Learning, Validation, Performance, and Applications to Genomics. *Curr. Genomics* 10, (2009).
22. Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C. & Faisal, A. A. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat. Med.* 24 (2018).
23. Merriam-Webster. Algorithm. *Merriam-Webster.com dictionary*. [www.merriam-webster.com/dictionary/algorithm](http://www.merriam-webster.com/dictionary/algorithm).
24. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks (2017). <https://dl.acm.org/doi/10.1145/3065386>.
25. Rosenblatt, F. The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review* vol. 65 (1958).
26. Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput.* 2006 Jul;18(7):1527-54.
27. Merriam-Webster. Neural Network. <https://www.merriam-webster.com/dictionary/neural-network>.
28. IBM Cloud Education. Neural Networks. [www.ibm.com/cloud/learn/neural-networks#toc-neural-net-u3voPJVU](http://www.ibm.com/cloud/learn/neural-networks#toc-neural-net-u3voPJVU) (2020).
29. IBM Cloud Education. Deep Learning. <https://www.ibm.com/cloud/learn/deep-learning> (2020).
30. Pereira T, Tran N, Gadhoumi K, Pelter MM, Do DH, Lee RJ, Colorado R, Meisel K, Hu X. Photoplethysmography based atrial fibrillation detection: a review. *NPJ Digit Med.* 2020 Jan 10;3:3. doi: 10.1038/s41746-019-0207-9.
31. Abràmoff, M. D., Lavin, P. T., Birch, M., Shah, N. & Folk, J. C. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *Digit. Med.* 1, 39 (2018).
32. Yu Kun-Hsing, Beam Andrew L & Kohane Isaac S. Artificial intelligence in healthcare. *Nat. Biomed. Eng.* 2, 719–731 (2018).
33. Keane, P. A. & Topol, E. J. With an eye to AI and autonomous diagnosis. *npj Digit. Med.* 1, (2018).
34. Rohani, N. & Eslahchi, C. Drug-Drug Interaction Predicting by Neural Network Using Integrated Similarity. *Sci. Rep.* 9, (2019).
35. Raja, K., Patrick, M., Elder, J. T. & Tsoi, L. C. Machine learning workflow to enhance predictions of Adverse Drug Reactions (ADRs) through drug-gene interactions: Application to drugs for cutaneous diseases. *Sci. Rep.* 7, (2017).

36. Hyland SL, Faltys M, Hüser M, Lyu X, Gumbsch T, Esteban C, Bock C, Horn M, Moor M, Rieck B, Zimmermann M, Bodenham D, Borgwardt K, Rättsch G, Merz TM. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat. Med.* 2020 Mar;26(3):364-373. doi: 10.1038/s41591-020-0789-4.
37. Stehlik J, Schmalfluss C, Bozkurt B, Nativi-Nicolau J, Wohlfahrt P, Wegerich S, Rose K, Ray R, Schofield R, Deswal A, Sekaric J, Anand S, Richards D, Hanson H, Pipke M, Pham M. Continuous Wearable Monitoring Analytics Predict Heart Failure Hospitalization: The LINK-HF Multicenter Study. *Circ Heart Fail.* 2020 Mar;13(3):e006513. doi: 10.1161/CIRCHEARTFAILURE.119.006513.
38. Trezzini, B., Meyer, B., Ivankovic, M., Jans, C. & Golder, L. Der administrative Aufwand der Ärzteschaft nimmt weiter zu. *Schweizerische Ärztezeitung* (2020). doi: 10.4414/saez.2020.18482.
39. Rajkomar, A., Dean, J. & Kohane, I. Machine Learning in Medicine. *N. Engl. J. Med.* 380, (2019).
40. Coiera, E., Kocaballi, B., Halamaka, J. & Laranjo, L. The digital scribe. *npj Digit. Med.* 1, (2018).
41. Chiu, C.-C., Tripathi, A., Chou, K., Co, C., Jaitly, N., Jaunzeikare, D., Kannan, A., Nguyen, P., Sak, H., Sankar, A., Tansuwan, J., Wan, N., Wu, Y., Zhang, X. (2018) Speech Recognition for Medical Conversations. *Proc. Interspeech 2018*, 2972-2976, doi: 10.21437/Interspeech.2018-40.
42. Jing, B., Xie, P. & Xing, E. P. On the automatic generation of medical imaging reports. in *ACL 2018 - 56<sup>th</sup> Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers) vol. 1* (2018).
43. Gehlbach, P. L. Robotic surgery for the eye. *Nature Biomedical Engineering* vol. 2 (2018).
44. EU. Bericht über die Auswirkungen künstlicher Intelligenz, des Internets der Dinge und der Robotik in Hinblick auf Sicherheit und Haftung. (2020) doi: 10.1787/ab757416-en.
45. Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz eingesetzt von der Europäischen Kommission. Ethik-Leitlinien für eine vertrauenswürdige KI. [digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai](https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai) (2018).
46. Fraunhofer IAIS. Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz – KI Prüfkatalog (2021). [www.iais.fraunhofer.de/ki-pruefkatalog](https://www.iais.fraunhofer.de/ki-pruefkatalog).
47. Fraunhofer IAIS. Vertrauenswürdiger Einsatz von künstlicher Intelligenz. (2019).
48. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential Biases in Machine Learning Algorithms using Electronic Health Record Data. *JAMA Intern Med.* 2018 Nov 1;178(11):1544-1547. doi: 10.1001/jamainternmed.2018.3763.
49. Rajkomar, A., Hardt, M., Howell, M.D., Chin, M.H. & Author, C. Ensuring Fairness in Machine Learning to Advance Health Equity HHS Public Access. *Ann Intern Med* 169, 866–872 (2018).
50. Lyell, D. & Coiera, E. Automation bias and verification complexity: a systematic review (2016). doi: 10.1093/jamia/ocw105.
51. Tramèr F., Atlidakis V., Geambasu R., Hsu D., Hubaux J.P., Humbert M., Juels A., Lin H. FairTest: Discovering Unwarranted Associations in Data-Driven Applications. *Proceedings - 2nd IEEE European Symposium on Security and Privacy, EuroS and P* (2017). doi: 10.1109/EuroSP.2017.29.
52. FairTest. <https://github.com/columbia/fairtest>.
53. Nothwang, W.D., McCourt, M.J., Robinson, R.M., Burden, S. A. & Curtis, J. W. The human should be part of the control loop? in *Proceedings - 2016 Resilience Week, RWS* (2016). doi: 10.1109/RWEEK.2016.7573336.
54. Campanella G, Hanna MG, Geneslaw L, Mirafior A, Werneck Krauss Silva V, Busam KJ, Brogi E, Reuter VE, Klimstra DS, Fuchs TJ. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med.* 2019 Aug;25(8):1301-1309. doi: 10.1038/s41591-019-0508-1.
55. Benjamins, S., Dhunoo, P. & Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digit. Med.* 3, (2020).
56. FDA. Evaluation of automatic class III designation for Osteodetect. [www.fda.gov/about-fda/cdrh-transparency/evaluation-automatic-class-iii-designation-de-novo-summaries](https://www.fda.gov/about-fda/cdrh-transparency/evaluation-automatic-class-iii-designation-de-novo-summaries).
57. FDA. Critical Care Suite: Radiological Computer Assisted Triage and Notification Software. [www.accessdata.fda.gov/cdrh\\_docs/pdf18/K183182.pdf](https://www.accessdata.fda.gov/cdrh_docs/pdf18/K183182.pdf)
58. Narang A, Bae R, Hong H, Thomas Y, Surette S, Cadieu C, Chaudhry A, Martin RP, McCarthy PM, Rubenson DS, Goldstein S, Little SH, Lang RM, Weissman NJ, Thomas JD. Utility of a Deep-Learning Algorithm to Guide Novices to Acquire Echocardiograms for Limited Diagnostic Use. *JAMA Cardiol.* 2021 Jun 1;6(6):624-632. doi: 10.1001/jamacardio.2021.0185.
59. Yang, A., Wang, F. & Sheu, Y.-H. Illuminating the Black Box: Interpreting Deep Neural Network Models for Psychiatric Research. *Front. Psychiatry* [www.frontiersin.org](https://www.frontiersin.org) 11, 551299 (2020).
60. Collins, G.S., Reitsma, J.B., Altman, D.G. & Moons, K.G.M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD Statement. *Eur. Urol.* 67, 1142–1151 (2015).
61. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, Arnaout R, Kohane IS, Saria S, Topol E, Obermeyer Z, Yu B, Butte AJ. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med.* 2020 Sep;26(9):1320-1324. doi: 10.1038/s41591-020-1041-y.
62. Mongan, J., Moy, L. & Kahn, C. E. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiol. Artif. Intell.* 2, e200029 (2020).
63. Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J. & Denniston, A. K. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature medicine* vol. 26 (2020).

64. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group; SPIRIT-AI and CONSORT-AI Steering Group; SPIRIT-AI and CONSORT-AI Consensus Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med*. 2020 Sep;26(9):1351-1363. doi: 10.1038/s41591-020-1037-7.
65. Fedlex. Bundesgesetz über die Produktheftpflicht, SR 221.112.944. (2010).
66. Fedlex. Bundesgesetz über die Produktesicherheit, SR 930.11. (2010).
67. Landolt, H. & Herzog-Zwitter, I. *Arzt Haftungsrecht*. in 359ff (Dike Verlag, 2014).
68. Regina E. Aebi-Müller, Walter Fellmann, Thomas Gächter, Bernhard Rüttsche & Brigitte Tag. *Arztrecht*. in 267–328 (Stämpfli Bern, 2016).
69. Song X, Yu ASL, Kellum JA, Waitman LR, Matheny ME, Simpson SQ, Hu Y, Liu M. Cross-site transportability of an explainable artificial intelligence model for acute kidney injury prediction. *Nat Commun*. 2020 Nov 9;11(1):5668. doi: 10.1038/s41467-020-19551-w.
70. Lacaze P, Pinese M, Kaplan W, Stone A, Brion MJ, Woods RL, McNamara M, McNeil JJ, Dinger ME, Thomas DM. The Medical Genome Reference Bank: a whole-genome data resource of 4000 healthy elderly individuals. Rationale and cohort design. *Eur J Hum Genet*. 2019 Feb;27(2):308-316. doi: 10.1038/s41431-018-0279-z.
71. Rose, S. Machine Learning for Prediction in Electronic Health Data. *JAMA Network Open* vol. 1 (2018).
72. Nevin, L. Advancing the beneficial use of machine learning in health care and medicine: Toward a community understanding. *PLoS Medicine* vol. 15 (2018).
73. Fryback, D. G. & Thornbury, J. R. The efficacy of diagnostic imaging. *Med. Decis. Mak.* 11, (1991).
74. INFANT Collaborative Group. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet*. 2017 Apr 29;389(10080):1719-1729. doi: 10.1016/S0140-6736(17)30568-8.
75. Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. (2018). <https://doi.org/10.17863/CAM.22520>.
76. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, Bakas S, Galtier MN, Landman BA, Maier-Hein K, Ourselin S, Sheller M, Summers RM, Trask A, Xu D, Baust M, Cardoso MJ. The future of digital health with federated learning. *NPJ Digit Med*. 2020 Sep 14;3:119. doi: 10.1038/s41746-020-00323-1.
77. Warnat-Herresthal S, Schultze H, Shastry KL, Manmohan S, Mukherjee S, Garg V, Sarveswara R, Händler K, Pickkers P, Aziz NA, Ktena S, Tran F, Bitzer M, Ossowski S, Casadei N, Herr C, Petersheim D, Behrends U, Kern F, Fehlmann T, Schommers P, Lehmann C, Augustin M, Rybniker J, Altmüller J, Mishra N, Bernardes JP, Krämer B, Bonaguro L, Schulte-Schrepping J, De Domenico E, Siever C, Kraut M, Desai M, Monnet B, Saridaki M, Siegel CM, Drews A, Nuesch-Germano M, Theis H, Heyckendorf J, Schreiber S, Kim-Hellmuth S; COVID-19 Aachen Study (COVAS), Nattermann J, Skowasch D, Kurth I, Keller A, Bals R, Nürnberg P, Riess O, Rosenstiel P, Netea MG, Theis F, Mukherjee S, Backes M, Aschenbrenner AC, Ulas T; Deutsche COVID-19 Omics Initiative (DeCOI), Breteler MMB, Giamarellos-Bourboulis EJ, Kox M, Becker M, Cheran S, Woodacre MS, Goh EL, Schultze JL. Swarm Learning for decentralized and confidential clinical machine learning. *Nature*. 2021 Jun;594(7862):265-270. doi: 10.1038/s41586-021-03583-3.
78. WIPO. Artificial Intelligence and Intellectual Property. *World Intellectual Property Organization*. [www.wipo.int/about-ip/en/artificial\\_intelligence/policy.html](http://www.wipo.int/about-ip/en/artificial_intelligence/policy.html)
79. Hilty, R. M. *Urheberrecht*. (Stämpfli juristische Lehrbücher, 2020).
80. Vieth, A. Tom L. Beauchamp, James F. Childress (2009) *Principles of Biomedical Ethics*. 6. Auflage. *Ethik der Medizin* 22, (2010).
81. Europäisches Parlament. Regelung der zivilrechtlichen Haftung beim Einsatz künstlicher Intelligenz. [www.europarl.europa.eu/doceo/document/TA-9-2020-0276\\_DE.html](http://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_DE.html).
82. Gless, S. Von der Verantwortung einer E-Person. *Goltdammer's Arch. für Strafr.* 164, 324–329 (2017).
83. FMH. Standesordnung der FMH. [www.fmh.ch/fr/a-propos-de-la-fmh/statuts-autres-reglements.cfm](http://www.fmh.ch/fr/a-propos-de-la-fmh/statuts-autres-reglements.cfm) (2020).
84. FDA. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)-Discussion Paper and Request for Feedback. [www.fda.gov/media/122535/download](http://www.fda.gov/media/122535/download) (2019).
85. FDA. Software as a Medical Device (SaMD) Action Plan. [www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device](http://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device) (2021).
86. FDA. Good Machine Learning Practice for Medical Device Development: Guiding Principles. [www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles](http://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles) (2021).
87. EU. Verordnung des Europäischen Parlaments und Rates. Zur Festlegung harmonisierter Vorschriften für Künstliche Intelligenz (Gesetz über Künstliche Intelligenz) und Änderung bestimmter Rechtsakte der Union. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0206> (2021).

88. Council of the European Union. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts- Presidency compromise text, 14278-21. (2021).
89. EU. Verordnung (EU) 2017/745 des Europäischen Parlaments und Rates über Medizinprodukte, zur Änderung der Richtlinie 2001/83/EG, der Verordnung (EG) Nr. 178/2002 und der Verordnung (EG) Nr. 1223/2009 und zur Aufhebung der Richtlinien 90/385/EWG und 93/42/EWG des Rates. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A32017R0745> (2017).
90. EU. Neue Vorschriften für künstliche Intelligenz – Fragen und Antworten (2021). [https://ec.europa.eu/commission/presscorner/detail/de/QANDA\\_21\\_1683](https://ec.europa.eu/commission/presscorner/detail/de/QANDA_21_1683)
91. Karner, E., Koch, B. A. & Geistfeld, M. A. Comparative Law Study on Civil Liability for Artificial Intelligence Justice and Consumers. (2020).
92. Johner, C. Regulatorische Anforderungen an Medizinprodukte mit Machine Learning. [www.johner-institut.de/blog/iec-62304-medizinische-software/regulatorische-anforderungen-an-medizinprodukte-mit-machine-learning](http://www.johner-institut.de/blog/iec-62304-medizinische-software/regulatorische-anforderungen-an-medizinprodukte-mit-machine-learning) (2021).
93. Johner, C., Molnar, C., Purde, A., Rad, A., Dierks, C., Bunk, S., Piechottka, S. Leitfaden zur KI bei Medizinprodukten. [https://github.com/johner-institut/ai-guideline/blob/master/Guideline-AI-Medical-Devices\\_DE.md#leitfaden-zur-ki-bei-medizinprodukten](https://github.com/johner-institut/ai-guideline/blob/master/Guideline-AI-Medical-Devices_DE.md#leitfaden-zur-ki-bei-medizinprodukten).
94. ICMRA. Informal Innovation Network Horizon Scanning Assessment Report – Artificial Intelligence. [www.icmra.info/drupal/en/strategicinitiatives/innovation](http://www.icmra.info/drupal/en/strategicinitiatives/innovation) (2021).
95. Digital Society Initiative. Ein Rechtsrahmen für Künstliche Intelligenz. [www.dsi.uzh.ch/de/research/projects/strategy-lab/strategy-lab-21.html](http://www.dsi.uzh.ch/de/research/projects/strategy-lab/strategy-lab-21.html) (2021).
96. Graf, F. & Obrecht, L. Rechtliche Rahmenbedingungen für Künstliche Intelligenz in der Schweiz. *Jusletter* (2021) doi: 10.38023/431d79f8-7b98-4828-97beb346864f371d.
97. Fischer, J. & Fuchs, M. Brauchen wir eine Legaldefinition für künstliche Intelligenz? *Jusletter* (2021) doi: 10.38023/681145f8-fb8b-4354-90d7-79a7b5112fdc.
98. Fedlex. Medizinprodukteverordnung. SR 812.213. [www.fedlex.admin.ch/eli/cc/2020/552/fr](http://www.fedlex.admin.ch/eli/cc/2020/552/fr) (2020).
99. EU. Verordnung 2017/746 des Europäischen Parlaments und des Rates über In-vitro-Diagnostika und zur Aufhebung der Richtlinie 98/79/EG und des Beschlusses 2010/227/EU der Kommission. <https://eur-lex.europa.eu/legal-content/de/ALL/?uri=CELEX%3A32017R0746> (2017).
100. American Medical Association. Augmented Intelligence in Health Care policy report. (2019).
101. Crochane Deutschland. Evidenzbasierte Medizin. [www.cochrane.de/ueber-uns/evidenzbasierte-medizin](http://www.cochrane.de/ueber-uns/evidenzbasierte-medizin).
102. WHO. Ethics and governance of artificial intelligence for health. <https://www.who.int/publications/i/item/9789240029200>. (2021).
103. Bonvin R., Buhmann J., Cotrini Jimenez C., Egger M., Geissler A., Krauthammer M., Schirlo C., Spiess C., Steurer J., Vokinger K. N., Vogt J. Neue Lernziele für das Medizinstudium erarbeitet: Studierende auf den Einsatz von maschinellem Lernen vorbereiten. In: *Schweizerische Ärztezeitung* (2022), 4, S. 98-101.
104. Buhmann J., Felix J., Gächter T., Kowatsch T., Lehmann R., von Lutterotti N., Schedler K., Steurer J., Wolfrum C. (2018) Digitalisierung der Medizin: Konsequenzen für die Ausbildung. *Schweizerische Ärztezeitung*, 99 (42). 1441-1444.
105. Schüller, K., Busch, P. & Hindinger, C. Future Skills: Ein Framework für Data Literacy- Kompetenzrahmen und Forschungsbericht. (2019).
106. Misra, A., Jütting, J. & Kuonen, D. Advancing data literacy in the post-pandemic world. (2021).
107. CPRD. Clinical Practice Research Datalink. [www.cprd.com](http://www.cprd.com).
108. CPME. CPME Policy on AI in Health Care. [www.cpme.eu](http://www.cpme.eu) (2019).
109. American Medical Association. Augmented intelligence in health care: Policy. [www.ama-assn.org/system/files/2019-08/ai-2018-board-policy-summary.pdf](http://www.ama-assn.org/system/files/2019-08/ai-2018-board-policy-summary.pdf) (2018).
110. IRGC. The Governance of Decision-Making Algorithms. [www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/the-governance-of-decision-making-algorithms](http://www.epfl.ch/research/domains/irgc/specific-risk-domains/projects-cybersecurity/the-governance-of-decision-making-algorithms) (2018).
111. SAMW. Expert workshop report organised by SATW and SAMW on 'Artificial Intelligence in medicine - can it be trusted?' [www.assm.ch/fr/Projets/Apercu-des-projets/Intelligence-artificielle.html](http://www.assm.ch/fr/Projets/Apercu-des-projets/Intelligence-artificielle.html) (2018).
112. Frank, J.R., Snell, L., Sherbino, J. & Royal College of Physicians and Surgeons of Canada. CanMEDS 2015 physician competency framework (2015).
113. Rampton, V., Mittelman, M. & Goldhahn, J. Implications of artificial intelligence for medical education. *The Lancet Digital Health* vol.2 (2020).
114. Wartman, S.A. & Donald Combs, C. Medical education must move from the information age to the age of artificial intelligence. *Acad. Med.* 93, (2018).
115. Weissmann, Y., Wolfrum, C. & Goldhahn, J. Bachelor Humanmedizin an der ETH Zürich. *Praxis (Bern. 1994)*. 109, (2020).
116. CAIM. Master of Science Artificial Intelligence in Medicine. [www.aim.unibe.ch/education/msc\\_aim/index\\_eng.html](http://www.aim.unibe.ch/education/msc_aim/index_eng.html) (2021).

FMH – Fédération des  
médecins suisses, Berne  
Case postale  
3000 Berne 16

Tél. 031 359 11 11  
info@fmh.ch

