IA et santé

ChatGPT va-t-il bientôt nous soigner ?

Muriel Salle

Maîtresse de conférences · Université Claude Bernard Lyon 1 · Laboratoire S2HEP

AMPLEUR DES USAGES

L’IA générative est déjà là

9 sur 10

étudiant·es en santé ont déjà essayé un outil d’IA générative.

Un usage encore occasionnel, et une forte demande de formation.

Enquête du Pr Fabrice Denis, Institut de la e-Santé (INeS), 2025 + Le Quotidien du médecin, 13 mars 2025

1 sur 5

des médecins généralistes britanniques l’utilisent dans leur pratique clinique.

Sans cadre ni recommandations officielles.

Blease C. R. et al., « Generative artificial intelligence in primary care: an online survey of UK general practitioners », *BMJ Health & Care Informatics*, 2024, 31(1):e101102 (≈ 1 000 répondants)

Trois temps

1

Promesse

Quelles IA, pour quels usages ? Et que valent les chiffres qui circulent ?

2

Transformations

Ce que l'IA fait à la relation de soin, aux savoirs, à la responsabilité.

3

Vigilance critique

Biais, limites, et la place des soignant·es face aux outils.

Le programme de la matinée

8h00 – 8h45	**Partie 1**	Promesse : quelles IA, et que valent les chiffres ?
8h45 – 9h00	**Pause**	
9h00 – 10h30	**Parties 2 et 3**	Transformations, puis vigilance critique
10h30 – 10h45	**Échanges**	Vos questions
10h45 – 11h	**Pause**	
11h00 – 12h00	Chercheur·es invité·es : Thomas Guyet	

1

Promesse

Ce que l’IA promet à la médecine, et ce que valent les preuves.

DÉFINIR

Une intelligence… artificielle ?

La science qui fait faire aux machines des choses qui demanderaient de l'intelligence si elles étaient faites par des humains.

La définition, celle de Marvin Minsky, ne dit pas ce qu'est l'intelligence. Elle désigne des tâches.

D'après Marvin Minsky, cité par Laurent Ségalat, *L'intelligence artificielle*, 1985

DEUX FAMILLES

Des règles, ou des exemples

IA symbolique

« making a mind »

Des règles logiques écrites par des expert·es. Ce sont les systèmes experts des années 1970, comme MYCIN pour les infections bactériennes.

IA connexionniste

« modeling the brain »

Des réseaux de neurones formels qui apprennent à partir d'exemples. C'est l'apprentissage profond, et aujourd'hui l'IA générative.

Plassard, Gramond, Quillard et Levacher, « Médecine et intelligence artificielle », Le carnet de la MRSH, 2021

UN SYSTÈME EXPERT

Comment raisonne MYCIN



« Pourquoi cette question ? » MYCIN répond en affichant la règle qu'il teste.

E. H. Shortliffe, MYCIN, Stanford, 1974-1976 ; Yu V. L. et al., JAMA, 1979 ; règle résumée et traduite

DEUX FAMILLES

Des règles, ou des exemples

IA symbolique

« making a mind »

Des règles logiques écrites par des expert·es. Ce sont les systèmes experts des années 1970, comme MYCIN pour les infections bactériennes.

IA connexionniste

« modeling the brain »

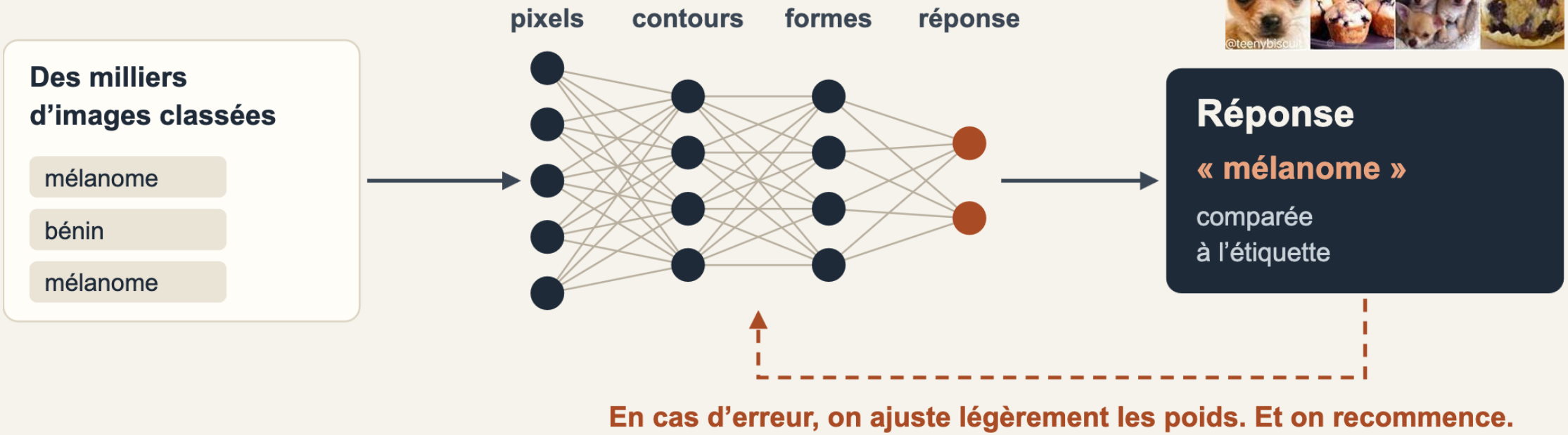
Des réseaux de neurones formels qui apprennent à partir d'exemples. C'est l'apprentissage profond, et aujourd'hui l'IA générative.

Plassard, Gramond, Quillard et Levacher, « Médecine et intelligence artificielle », Le carnet de la MRSH, 2021

APPRENDRE PAR L'EXEMPLE

Comment apprend un réseau de neurones





Aucune règle écrite : le savoir est dispersé dans des millions de poids.

Schéma simplifié. Krizhevsky, Sutskever et Hinton (AlexNet), 2012 ; Esteva A. et al., Nature, 2017

TYPOLOGIE

Quatre familles d’usages

Lire des images

Dermatologie, sénologie, ophtalmologie, pneumologie

Aider à décider

Diagnostic, choix du traitement, alertes

Produire du texte

Comptes rendus, courriers, synthèses de dossiers

Prédire

Pronostic, simulation, jumeaux numériques

D’après N. Ayache (Inria) et Académie nationale de médecine, « Systèmes d’IA générative en santé », mars 2024

AU CABINET

Ce que les généralistes utilisent vraiment

ANTIBIOCLIC
Antibiothérapie rationnelle en soins primaires

GESTACLIC

Pedia doc

Diversiclic

Des outils d'aide à la décision fondés sur les recommandations nationales et des sociétés savantes.

Antibioclic croise cinq variables : adulte, enfant, grossesse, allaitement, insuffisance rénale.

On est plus près de l'arbre de décision que de ChatGPT.

Santé Académie, « Médecin généraliste : la cartographie des applications et des outils numériques », juillet 2024 ; ANS

LA PROMESSE

Une médecine « 4P »

Personnalisée	Précise	Prédictive	Préventive

« L'intelligence artificielle permet d'augmenter les capacités de raisonnement des médecins. »

Nicholas Ayache (Inria), « Intelligence artificielle et médecine », [date à préciser]

LU SUR LINKEDIN

« La machine fait mieux que les médecins »

Stanford

« GPT-4 seul a battu 50 médecins. »

Microsoft

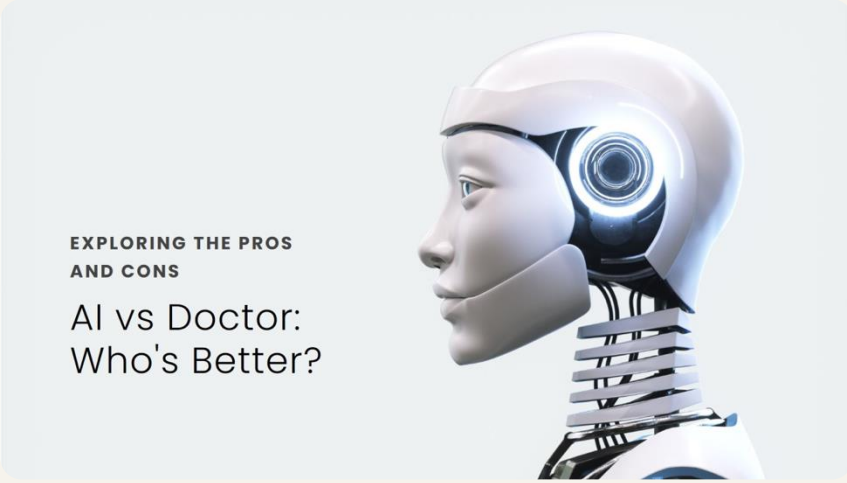
« Le bon diagnostic 85 fois sur 100. Les médecins, 20 fois sur 100. »

Kenya

« 16 % d’erreurs de diagnostic en moins. »

Suède

« 29 % de cancers en plus, 12 % de cancers ratés en moins. »



Extraits d’un post LinkedIn récent vantant l’IA en santé, appuyé sur quatre études

GRILLE DE LECTURE

Cinq questions à poser à toute étude

1. Qui l’a produite ?
2. Quel type de publication : revue à comité de lecture, prépublication, communiqué ?
3. Qu’est-ce qui est mesuré, exactement ?
4. Dans quelles conditions : cas d’école ou consultations réelles ?
5. Qui finance, et qui a intérêt au résultat ?

1 · STANFORD

Donner l’outil ne rend pas meilleur

74 %

médecins sans GPT-4

76 %

médecins avec GPT-4 : écart non significatif

+16 pts

GPT-4 seul, en analyse secondaire

Essai randomisé, 50 médecins, vignettes cliniques. Revue à comité de lecture.

Goh E. et al., JAMA Network Open, 2024;7(10):e2440969

2 · MICROSOFT

Des médecins privés de leurs ressources

85,5 %

le système MAI-DxO, réglé pour la précision maximale

20 %

21 médecins, sans collègues, sans manuels, sans IA

304 cas du New England Journal of Medicine, rares et choisis pour leur difficulté. **Prépublication, produite par l'entreprise qui vend l'outil.**

Nori H. et al., « Sequential Diagnosis with Language Models », arXiv:2506.22405, juin 2025

3 · KENYA

Enfin des consultations réelles

39 849

consultations, 15 cliniques de Nairobi

−16 %

d’erreurs diagnostiques, en valeur relative

Un copilote qui alerte le clinicien sans décider à sa place. **Prépublication, coproduite par OpenAI. Effet sur la santé des patients non mesuré.**

Korom R. et al. (OpenAI et Penda Health), arXiv:2507.16947, juillet 2025

4 · SUÈDE

« Non inférieur » ne veut pas dire « meilleur »

+29 %

de cancers détectés au dépistage

1,55 ‰

cancers d’intervalle avec IA, contre 1,76 ‰ sans IA

Essai randomisé, 105 915 femmes. L’intervalle de confiance du ratio (0,65 à 1,18) inclut 1 : **la baisse de 12 % n’est pas démontrée.**

Essai MASAI. Hernström V. et al., Lancet Digital Health, 2025 ; Gommers J. et al., The Lancet, 2026;407:505-514

BILAN

Les quatre études, relues

Étude	Publication	Produite par	Ce qu'elle montre
Stanford	Revue à comité de lecture	Universités	L'outil n'aide pas les médecins
Microsoft	Prépublication	Microsoft	Cas d'école, médecins sans ressources
Kenya	Prépublication	OpenAI	Moins d'erreurs, effet patient inconnu
Suède	Revue à comité de lecture	Universités	Plus de détection ; baisse des cancers ratés non démontrée

UNE AUTRE ÉTUDE · OXFORD

Et quand un humain parle à la machine ?

95 %

l'IA seule, à qui l'on donne le cas complet

≤ 35 %

des personnes qui dialoguent avec un chatbot

47 %

des personnes sans IA, avec leurs connaissances ou un moteur de recherche

Part des cas où une pathologie pertinente est identifiée. Essai randomisé, 1 298 personnes, trois modèles (dont GPT-4o). **Revue à comité de lecture, universitaires.**

Bean A. M., Mahdi A. et al., Nature Medicine, 2026, doi:10.1038/s41591-025-04074-y ; relayé sur LinkedIn par Océane Sorel

UNE ÉTUDE, DEUX LECTURES

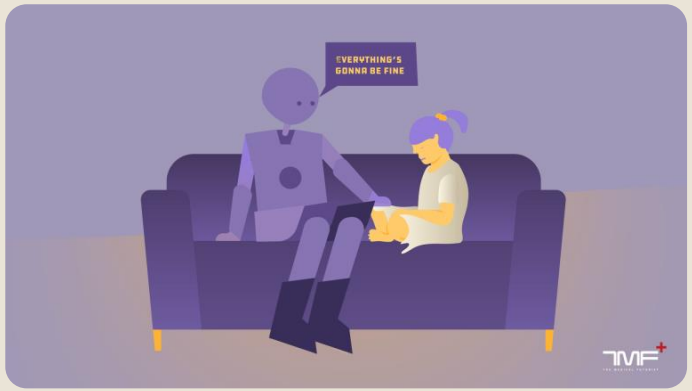
ChatGPT plus empathique que les médecins ?

78,6 %

des réponses de ChatGPT préférées à celles des médecins

52 / 211

mots en moyenne : médecins, puis ChatGPT



MACSF Prévention médicale, 2024

Des résultats « impressionnants » : la téléconsultation par IA n'est plus très loin.

Shafi Ahmed, LinkedIn

L'exactitude des réponses n'a pas été évaluée ; les critères sont subjectifs.

Ayers J. W. et al., JAMA Internal Medicine, 2023;183(6):589-596 ; 195 échanges tirés d'un forum Reddit ; illustration : The Medical Futurist

L’AUTRE PROMESSE

« L’IA ne remplacera pas les médecins »

Shafi Ahmed

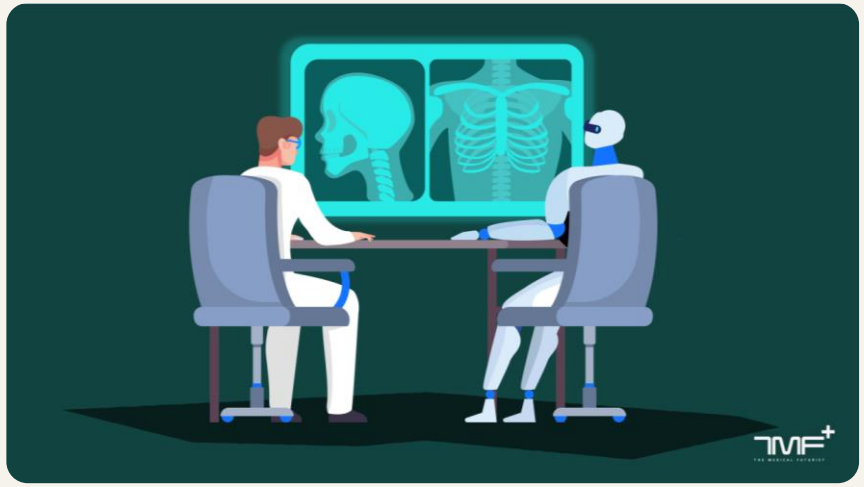
Chirurgien et « futurologue »

L’IA comme « stagiaire numérique ». Cite un produit commercial nommé comme une révolution.

Bertalan Meskó

« The Medical Futurist », conférencier

Humain et machine, plus forts ensemble. À l’appui : un épisode de *Dr House*.



Mais l’association ne va pas de soi : à Stanford, les médecins avec GPT-4 ne font pas mieux ; à Oxford, les personnes aidées d’un chatbot font moins bien.

GÉNÉALOGIE D’UNE PROMESSE

Une phrase qui a cinquante ans

Années 1970

MYCIN

jugé supérieur aux médecins généralistes

2016

Watson, IBM

« fait déjà mieux que les médecins »

2022

IBM

revend sa division Watson Health

2024-2025

GPT-4, MAI-DxO

annoncés quatre fois plus précis que des généralistes

Ségalat, 1985 ; Massie, Quaderni, 1992 ; Caducee.net, 2016, cité par Mathieu-Fritz, Sciences sociales et santé, 2021 ; Microsoft, 2025

Une promesse est un discours. Elle a des auteurs et des intérêts.

Lire une étude, c’est d’abord se demander qui parle.

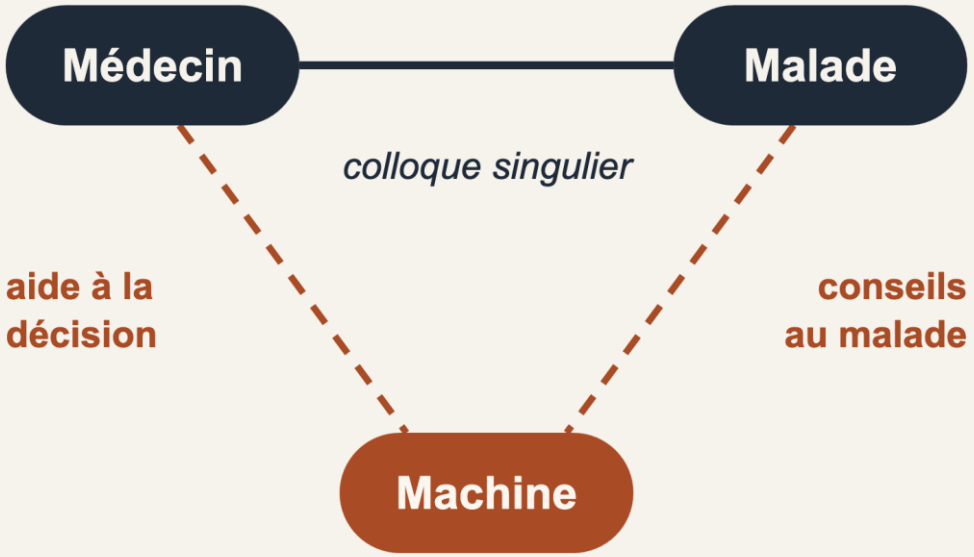
2

Transformations

Ce que le recours à l’IA fait à la relation de soin, aux savoirs et à la responsabilité.

RELATION DE SOIN

Médecin, malade… et machine



La relation de soin devient tripartite.

Ce qu’on espère

Libérer le médecin des tâches administratives pour lui rendre du temps clinique.

Ce qu’on risque

Un intermédiaire de plus entre médecin et patient : une médecine plus efficace, une relation appauvrie.

Mittelstadt B., rapport pour le Conseil de l’Europe, 2021 ; Zeinatou, 2019

DEUX RISQUES

Trop croire la machine, et désapprendre

Biais d’automatisation

On suit la recommandation de l’outil, même quand elle contredit ce qu’on voit.

Perte de compétences

Ce qu’on délègue à la machine, on cesse de le pratiquer, puis de savoir le faire.

Rappel de Stanford : un outil performant ne rend pas, à lui seul, le médecin meilleur. Tout se joue dans la manière de s’en servir.

Mittelstadt, Conseil de l’Europe, 2021 ; Goh et al., JAMA Network Open, 2024

SAVOIRS

Ce que la machine ne voit pas

Une enquête en radiologie

« La reconnaissance des anomalies par les radiologues ne se résume pas à un simple repérage de lésions. »

Anichini et Geffroy, Sciences sociales et santé, 2021

Les savoirs tacites

« Tout ce qui est nécessaire pour agir efficacement [...] mais qui ne peut pas être entièrement formulé, écrit ou transmis. »

Harry Collins, 1974

L'IA oblige à dire plus clairement que jamais en quoi consiste l'expertise médicale.

HARRY COLLINS

Trois sortes de savoirs tacites

Relationnel

Ce qui pourrait se dire, mais ne se dit pas.

L'astuce de l'interne pour installer le patient avant une ponction lombaire.

Somatique

Ce que le corps apprend par la répétition.

Sentir sous les doigts un ventre qui se défend ; reconnaître des crépitants.

Collectif

Ce que l'on apprend en vivant parmi les autres.

Entendre ce que cache le « j'ai un peu mal » d'une patiente qui ne se plaint jamais.

Pour Collins, le savoir tacite collectif est celui qu'aucune machine ne peut acquérir.

H. Collins, Tacit and Explicit Knowledge, 2010 ; Artificial Experts, 1990 ; M. Polanyi, The Tacit Dimension, 1966

RESPONSABILITÉ ET EXPLICABILITÉ

Trois questions face à une boîte noire

1. Sur quelles informations la recommandation repose-t-elle ?
2. Pourquoi cette recommandation a-t-elle été produite ?
3. Qui est en mesure d’en répondre si elle s’avère erronée ?

Comment justifier une décision dont on ne comprend pas le raisonnement ? Et comment informer le patient pour obtenir son consentement éclairé ?

Mittelstadt, 2021 ; Plassard et al., 2021 ; Académie nationale de médecine, 2024

EFFICACITÉ

Une décision optimale est-elle forcément une bonne décision de soin ?

Mortalité, morbidité, coûts… mais aussi qualité de vie, relation, éthique, justice.

OPTIMAL OU BON ?

Trois cas où l’optimal déraille

Une alerte qui sonne trop

Un modèle de détection du sepsis très diffusé manque 2 cas sur 3 et alerte pour 18 % des patients hospitalisés.

Performant sur le papier, épuisant au chevet.

Des règles qui s’additionnent

79 ans, cinq maladies chroniques, toutes les recommandations suivies : 12 médicaments, 19 prises par jour.

Chaque règle est juste ; leur somme ne l’est pas.

Ce que veut le malade

Cancer du poumon métastatique : avec des soins palliatifs précoces, meilleure qualité de vie et 11,6 mois de survie médiane, contre 8,9.

Traiter plus n’est pas mieux traiter.

Qu’optimise-t-on, au nom de qui, et qui décide de ce qui compte ?

Wong A. et al., JAMA Intern. Med., 2021 ; Boyd C. M. et al., JAMA, 2005 ; Temel J. S. et al., NEJM, 2010

3

Vigilance critique

Biais, limites, et la place des soignant·es face aux outils.

BIAIS

« Il ne faut pas que la machine, pour auto-apprendre, soit perturbée par trop de singularité, de différences ethniques, de conditions sociales [...], en un mot, par la vie. »

L'IA fonctionne par moyenne et par écart-type. Ce qui s'écarte de la moyenne disparaît.

Didier Sicard, L'éthique médicale et la bioéthique, PUF

UN ALGORITHME, 200 MILLIONS DE PERSONNES

Prédire des coûts n’est pas repérer des besoins

Le but : repérer les patient·es qui ont besoin d’un suivi renforcé

→

Ce qu’il mesure : les dépenses de santé passées

→

Le piège : à besoin égal, on dépense moins pour les patient·es noir·es

17,7 %

de patient·es noir·es orienté·es vers le suivi par l’algorithme

46,5 %

si l’on corrige le biais

L’algorithme a appris les inégalités d’accès aux soins.

Obermeyer Z. et al., « Dissecting racial bias in an algorithm used to manage the health of populations », Science, 2019

GENRE

L’IA hérite des angles morts de la médecine

Les données

Un savoir médical construit sur un corps masculin par défaut. L’outil apprend ce qu’on lui montre, angles morts compris.

Les concepteurs

Qui écrit les algorithmes ? Les femmes restent minoritaires en informatique, là où se décide ce qu’on demande aux machines.

Exemple : le genre du patient et le diagnostic du syndrome coronarien aigu aux urgences (thèse de Laura Fanfano, 2026).

Fanfano L., thèse, avril 2026 ; Favre C., HDR, Université Lumière Lyon 2, 2024

POSITIONNEMENT

Ce que recommande l'Académie de médecine

Recommandation 2

« Il serait contraire à l'éthique de se passer de l'aide de ces outils. »

Recommandation 4

« Toute utilisation [...] doit faire l'objet d'une supervision humaine. »

Une institution reprend la promesse à son compte. Reste à savoir qui supervise, avec quelle formation et quel temps.

Académie nationale de médecine, rapport « Systèmes d'IA générative en santé », mars 2024

CE QUI RESTE AU MÉDECIN

Les « 4C »

Compréhension globale

Esprit critique

Compassion

Conscience professionnelle

Selon Ayache, ces qualités restent « pour l'instant l'apanage de l'intelligence humaine ».

Nicholas Ayache (Inria), « Intelligence artificielle et médecine »

ChatGPT va-t-il bientôt nous soigner ?

Sur des cas d'école, il diagnostique parfois mieux que nous.

Entre les mains d'un médecin, il ne le rend pas forcément meilleur.

Soigner reste une relation et une responsabilité, que la machine n'endosse pas.

11H–12H

Cas concrets

Des chercheur·es présentent leurs travaux en cours.

[Noms des intervenant·es et titres des présentations]