

Définitions, enjeux et défis de l'IA appliquée à la médecine

Pr Fabrice DENIS



Histoire

- Années 40-50: concept de machine intelligente
 - Alan Turing, Machine Enigma, Test de Turing
- 1956 Terme "intelligence artificielle" à la conférence de Dartmouth: science et ingénierie de la fabrication de machines intelligentes, en particulier de programmes informatiques intelligents
- 1980-1990: Réseaux de neurones: traitent des informations non structurées et apprennent de nouvelles connaissances à partir de grandes quantités de données.
- 2020 et +: IA générative

Quelques définitions

- **IA (Intelligence Artificielle)** : L'IA fait référence aux systèmes ou machines qui miment l'intelligence humaine pour effectuer des tâches et qui peuvent s'améliorer progressivement en fonction des informations recueillies.
- **IA Faible** : Aussi appelée ANI (Artificial Narrow Intelligence), l'IA faible est conçue pour effectuer une tâche spécifique, comme la reconnaissance vocale ou le jeu d'échecs.
- **IA Symbolique** : L'IA symbolique, aussi appelée IA à base de connaissances, est une approche de l'IA qui utilise des symboles pour représenter et manipuler des connaissances. Elle se base sur des règles et des logiques prédéfinies.
- **IA Générative** : L'IA générative est une sous-catégorie de l'intelligence artificielle qui se concentre sur la création de contenu réaliste et original, comme des images, des vidéos, du texte ou de la musique.

Quelques définitions

- **Deep Learning** : Le Deep Learning est une sous-catégorie de l'apprentissage automatique (Machine Learning) qui utilise des réseaux de neurones artificiels pour apprendre et faire des prédictions à partir de grandes quantités de données. Il permet de découvrir des schémas complexes dans les données.
- **Large Language Model (LLM)** : Un Large Language Model est un modèle d'apprentissage automatique conçu pour comprendre et générer du langage naturel. Il est entraîné sur un vaste corpus de textes et est capable de prédire la probabilité d'un mot ou d'une phrase donnée en fonction du contexte
- **Boîte Noire** : Une boîte noire en IA fait référence à un système dont les résultats sont observables, mais dont les processus internes sont inconnus ou incompréhensibles pour les humains.
- **Boîte Blanche** : Une boîte blanche en IA est un système dont les processus internes sont entièrement compréhensibles et explicables pour les humains.
- **Boîte Grise** : Une boîte grise en IA est un système dont certains processus internes sont compréhensibles, mais d'autres non. Elle se situe donc entre la boîte noire et la boîte blanche

Forces incontestables de l'IA

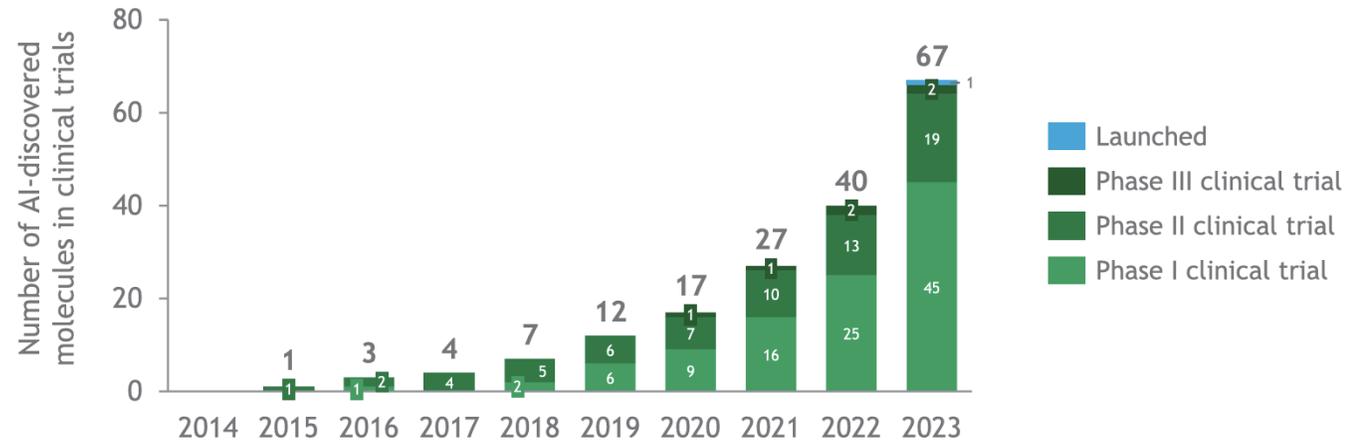
- **Automatisation** de tâches « simples »: plannings, interprétation, contourage, statistiques
- **Rapidité**
- **Génération** de comptes-rendus, rapport, courriers...
- **Triage** de données
- **L'Humain en a besoin... médecin augmenté**

Faiblesses incontestables de l'IA

- **Données sources:** qualité, mise à jour
- **Interprétabilité:** obligation d'information: risque médico-légal+++
- Collecte de données **sensibles**
- **Biais algorithmiques:** sur groupes sous représentés
- **Hallucinations** dans l'IA générative
- **A besoin de l'humain... ne remplace pas le médecin**

En matière de recherche fondamentale

- Mise en évidence de **cibles thérapeutiques** moléculaires par IA
- Analyse conformationnelle des **thérapies** ciblées
- Meilleure efficacité
- Réussite x2 versus sans IA : 85% phase I, 40% phase 2
- Moindre toxicité
- Moindre coût de développement pré-clinique
- Augmentation exponentielle



En matière de recherche clinique

- Création de **jumeaux numériques**, y compris radiologiques
- Calcul des **effectifs de patients** optimisé
- **Moins de patients nécessaires** dans les essais randomisés dans le bras contrôle
- Effectif global de patients plus faible
- **Durée des inclusions plus faibles** car moins d'évènements nécessaires
- Coût et **durée de développement moindre**
- Exemple Novadiscovery: **simulation in silico** vs **étude réelle**

	Official FLAURA2 Phase III results (Investigator)	Official FLAURA2 Phase III results (BICR)	Predicted by jinkō with publicly available data
HR	0.62 (0.49-0.79)	0.62 (0.48-0.80)	0.602 (0.49-0.74)
PFS/TTP (experimental arm, osimertinib combination treatment)	PFS: 25.5 (24.7-NC) months	PFS: 29.4 (25.1-NC) months	TTP: 25.9 (25.1-27.1) months
PFS/TTP (comparator arm, single-agent osimertinib)	PFS: 16.7 (14.1-21.3) months	PFS: 19.9 (16.6-25.3) months	TTP: 17.3 (16.8-18.0) months

En matière de formation: forces

- **Pour les enseignants:**

- Gain de temps et d'exhaustivité de contenus
- Créations de QCM
- Mise en place de scénarios pédagogiques

- **Pour les apprenants:**

- Apprentissage adaptatif: **retour sur erreur immédiat** par l'IA générative
- Validation des acquis **en temps réel**
- **Apprenant + actif** qu'en cours magistral (Q&R immédiates sans retenues)
- Cas cliniques plus réalistes et **plus nombreux** (chirurgie, radiologie...)

En matière de formation: faiblesses

- **Pour les enseignants:**

- **Hallucinations+++**
- **Base de données sources:** exemple des études non publiées, des études en cours

- **Pour les apprenants:**

- Trop s'appuyer sur les outils d'interprétation automatique sans apprendre la **séméiologie radiologique**
- **Manque de notions de radio-anatomie**
- Perte de l'**expertise en contourage** pour les oncos-radiothérapeutes
- **Baisse de niveau** de certaines acquisitions (anatomie, orientations diagnostiques...)

En matière de financement

- 80% des startups en manque de trésorerie à 10 mois
- Levée de fond difficiles, longues, faibles: 1,4 MD 2022 à 200ME en 2023 !
- Coûts des développements cliniques médian: 1 ME
- Coûts du domaine réglementaire médian: 1 ME
- Modèle économique B to C... Remboursement... quasi impossible
- Modèle B to B: rentable si gain de temps médecin ou si acte CNAM (télésurveillance)

En matière d'évaluation des solutions

- Plusieurs milliers de solutions disponibles
- Comment choisir la plus pertinente, performante, mature ?
- Score MDS...

Information sur le score MDS

OUI : un(e)

NON

Présence d'une certification ISO de la société ? (une seule réponse)

OUI

NON

Evaluation de la sécurité de l'éditeur (une seule réponse) :

Alerte de sécurité ANSM avec perte de chance

Alerte de sécurité sans perte de chance

Pas d'alerte de sécurité

Dans le business model : Ambition ou nécessité de remboursement ? (une seule réponse)

OUI (sauf si "dispositif à usage collectif hospitalier")

NON : pas besoin dans business model (vente de data, abonnement utilisateur, ect...)

NON : dispositif à usage collectif hospitalier

Probabilité de remboursement calculée : FAIBLE

VALIDER LE SCORE

Score INES
APPLI XYYYXX TEST

Information sur l'application	Evaluation de la cible
11.2 /20 C	13.0 /20 B
Evaluation clinique	Evaluation de l'éditeur
8.0 /20 D	2.5 /20 E

SCORE TOTAL
8.7/20 **D**

Echelle des quintiles -]8 : E, [8-10[: D, [10-13[: C, [13-15[: B, [15 : A
Wagneur et al. JMIR 2022

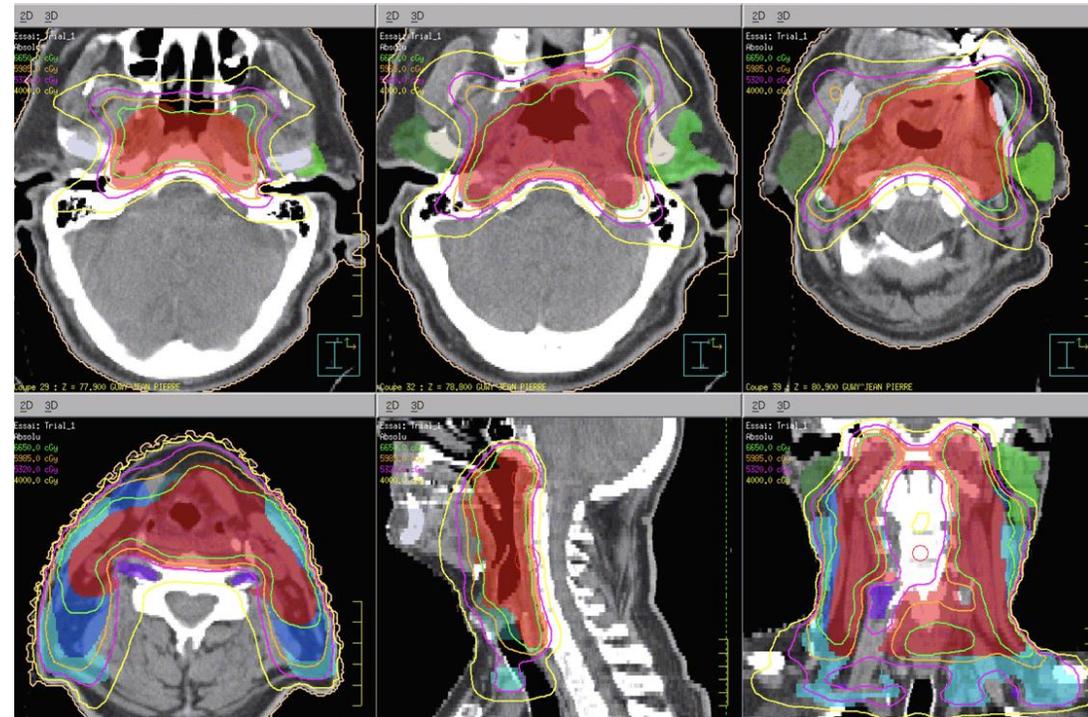
Liste de vos items retirants des points

Exemples d'applications



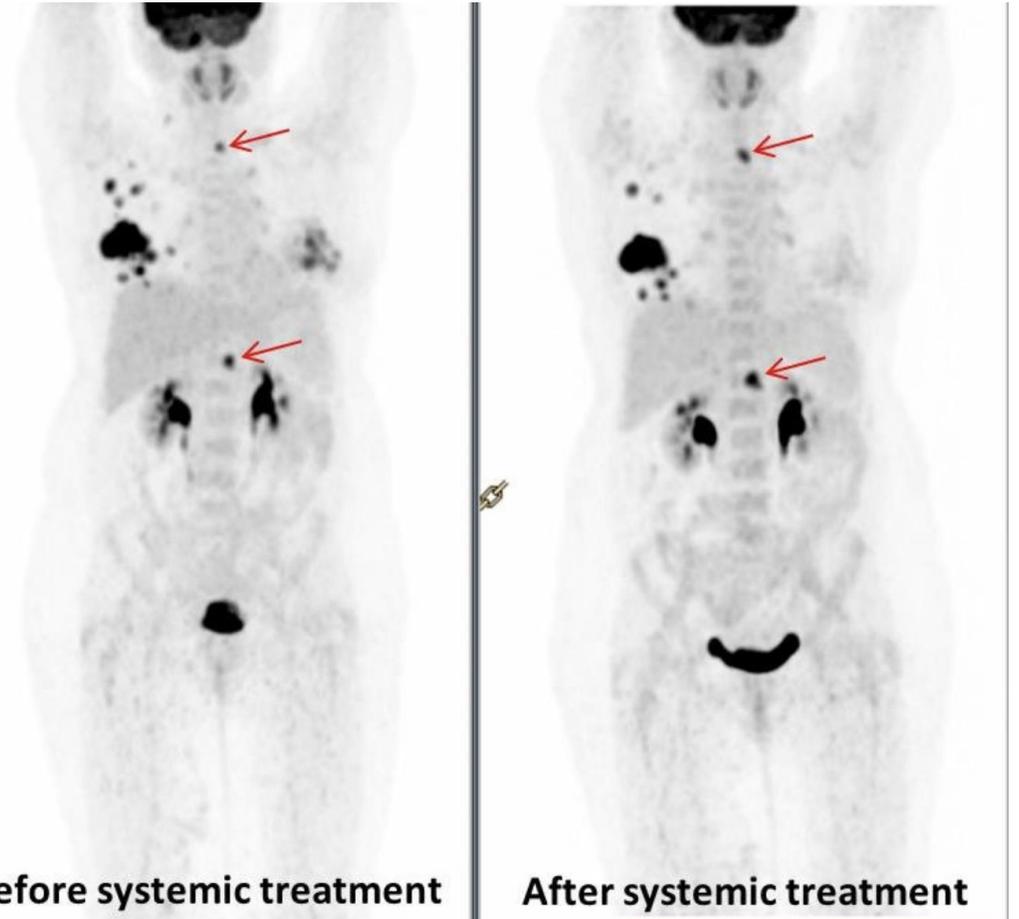
AIDE AU CONTOURAGE

- Organes à risque en radiothérapie +/- volumes cibles
- Excellent !
- Rapidité: temps divisé par 10 !
- Précision
- Réduction forte des écarts inter-observateurs
- Standard...



Réévaluation tumorale

- Analyse imageries avant et après traitement
- Caractérise et quantifie la réponse tumorale
- Simplifie travail du radiologue
- ... et de l'oncologue...



Dépistage

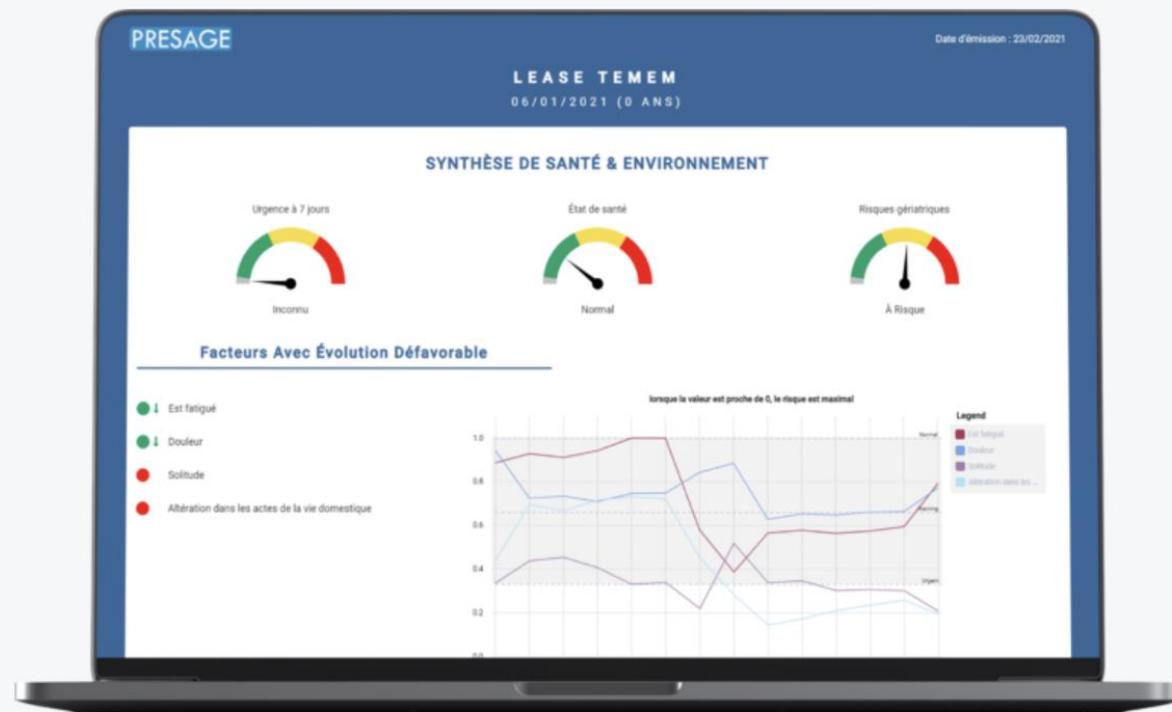
- Indication dans le dépistage des cancers du poumon chez le fumeur* Se=94%, Sp 95%
- Dépistage du cancer du sein: idem Humain mais -44% temps**
- À venir: dépistage automatique d'incidentalomes... sur imagerie corps entier
- À évaluer car faux positifs=gestes invasifs...

Facilitations pour le médecin

- Outils de rédaction automatique de compte-rendu de Cs (Nabla...)
- Outils de vulgarisation de compte-rendus anapath/radio (Vulgaroo...)
- Outils d'analyse de bilan biologique (Kiro...)
- Outils de prédiction d'hospitalisation de personnes âgées: Présage



Application mobile **PRESAGE**



Console d'analyse clinique

Prévention d'hospitalisations aux urgences

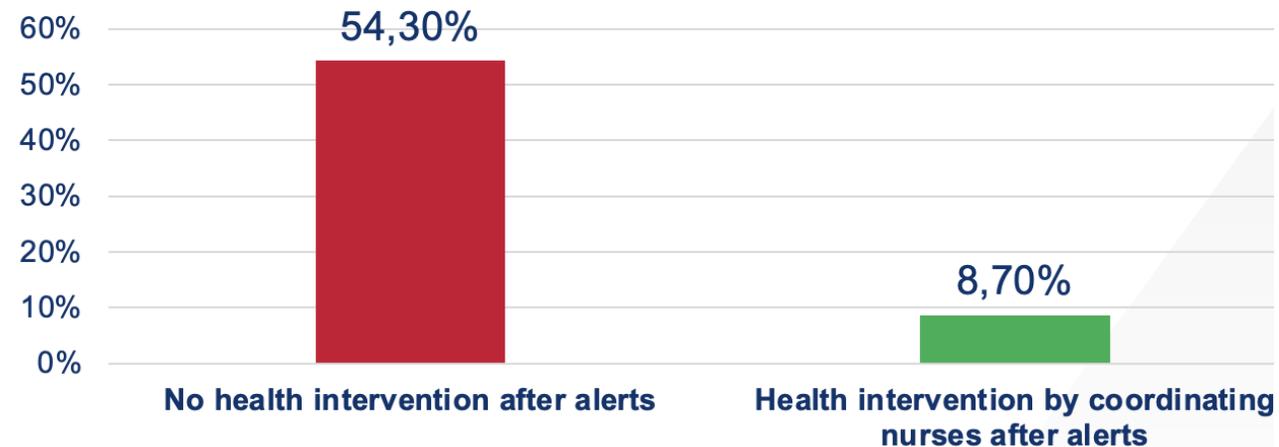
Détection précoce du risque d'hospitalisation par IA

- Application smartphone DM
- Estimation de l'état d'autonomie d'un sénior
- Prédit son risque d'hospitalisation dans les 2 prochaines semaines.
- Par auxiliaire de vie/proche → IA → Alerte infirmière/médecin EHPAD → mesures préventives
- 2 études cliniques sur 500000 données de vie réelles

Résultats de l'étude en vie réelle

- 3 centres
- 206 patients de plus de 85 ans (âge moyen)
- Sensibilité à 14 jours: 83%, Spécificité 86%
- Alerte suivie d'intervention: 8,7%
d'hospitalisations
- Alerte non suivie d'intervention: 54,3%
($p < 0.0001$)

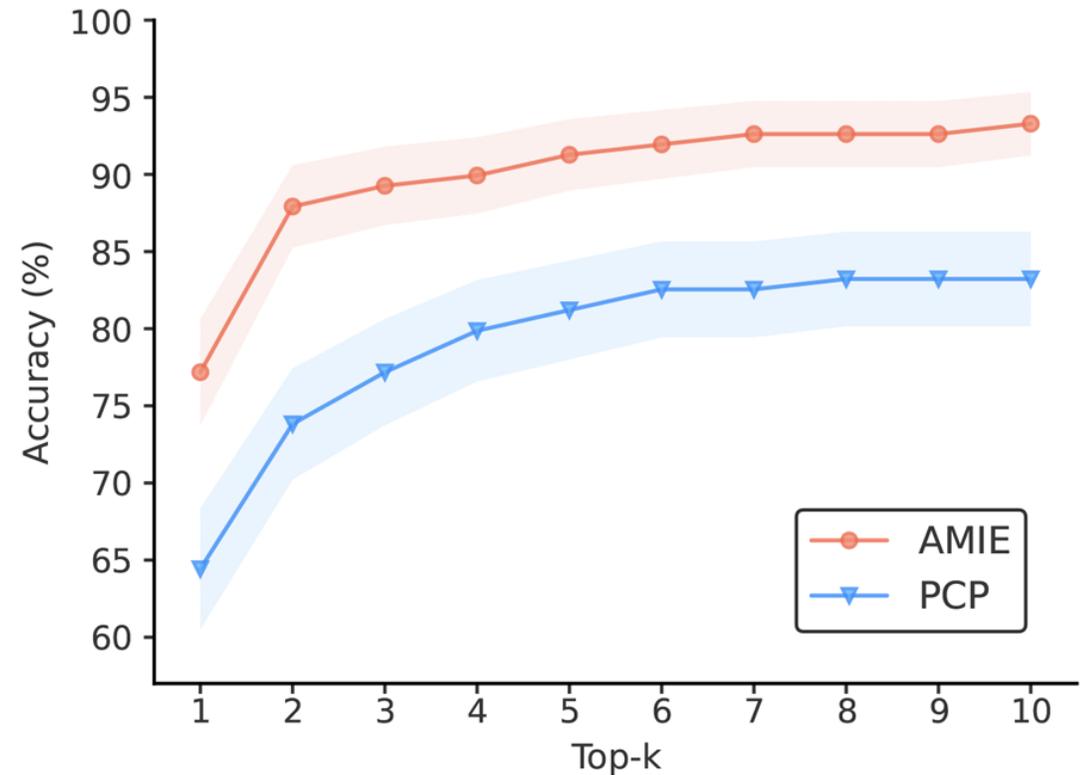
**Comparison of hospitalization rate
after PRESAGE CARE predictive alerts** :**



À venir: outil de diagnostic

Comparaison avec médecin: 6 à 0 pour l'IA

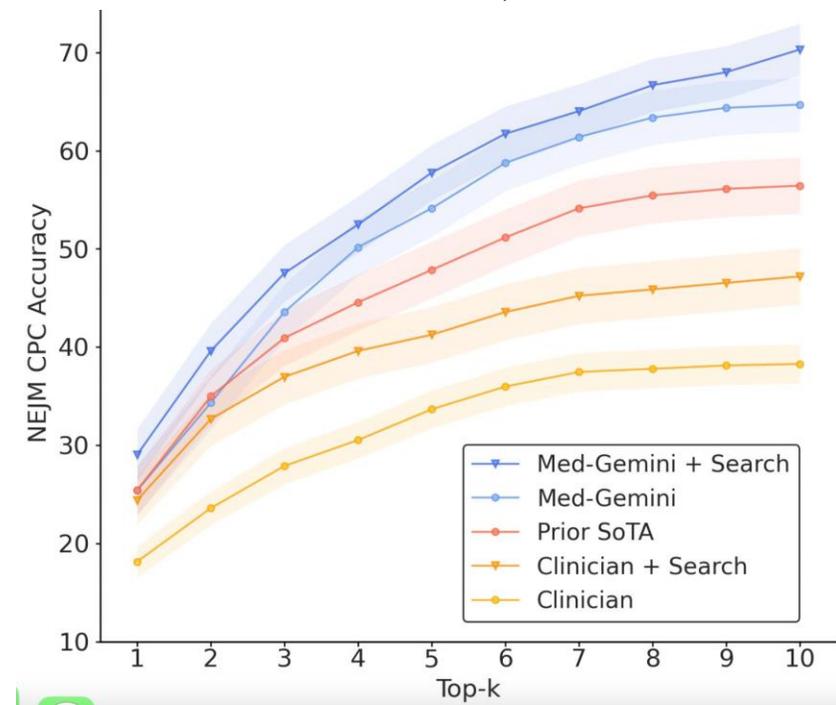
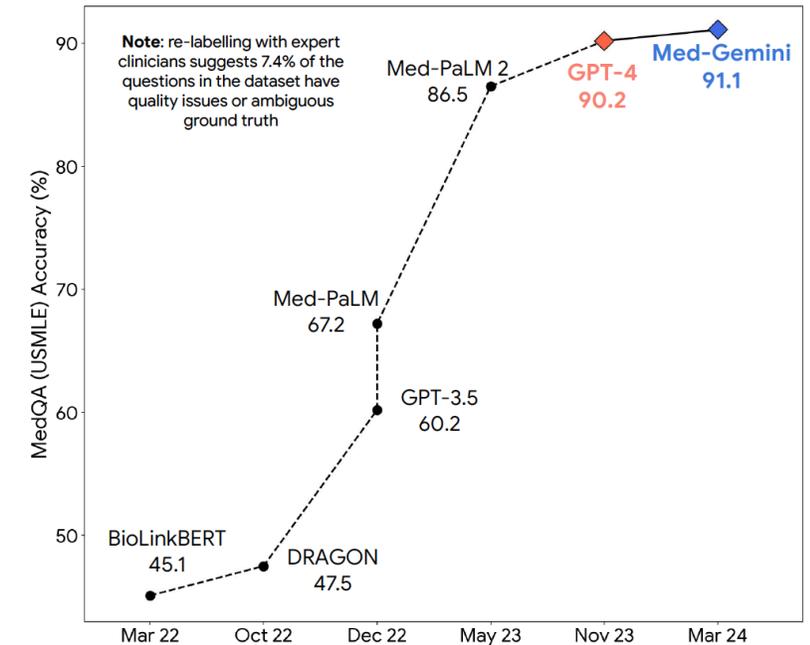
- 1) diagnostic 30 patients: 97% ChatGPT vs 87 MD (*Annals of Emergency Medicine 2023*)
- 2) ECG: IA supérieur ou égal à cardiologue (*JAMA Cardiology 2021*)
- 3) Sévérité Parkinson (*Nature DM 2023*)
- 4) Glaucome: IA mieux, Rétinopathie: idem MD (*JAMA Opht 2024*)
- 5) Réponses de l'IA près de quatre fois plus qualitatives et près de 10 fois plus empathiques que les réponses des médecins. (*JAMA Int Med 2023*)
- 6) Google AMIE: meilleur efficacité diagnostic que MD dans les 6 spécialités a surpassé les médecins dans 24 des 26 critères de qualité de la conversation, notamment la politesse, les explications, l'honnêteté et l'expression de l'attention et de l'engagement (*ArXIV 2024*)



À venir: outil de diagnostic: Gemini

- Capacités générales de raisonnement multimodal et à contexte long
- Intègre recherche Web
- Surpasse GPT4 (87% vs 65% médecin)
- 91% de précision
- Nettement supérieur au médecin (test sur cas complexes NEJM)
- Récupère info dans DPI, répond à des questions sur vidéo de chir, **imagerie**, dialogue médical, courriers...

SoTA on MedQA (USMLE)



Conclusion

- Bien se former à la santé numérique
- Outils d'aide et non de remplacement (examen clinique++)
- Gare à la baisse de niveau de formation
- Risque médico-légal si boîte noire aide à la décision: Intelligibilité des algos+++
- Bénéfices cliniques en augmentation
- Nécessité d'études cliniques pour toute indication