Potentiel et Limites de l'IA Dans Les Maladies Auto-Inflammatoires : Une étude de Non-Infériorité Basée Sur Des Vignettes Cliniques

1er Auteur : Oumaya, BEN HAMDA, Médecine interne / Unité de Recherche Maladies Auto-Immunes UR17DN02, Hôpital Militaire Principal d'Instruction de Tunis, Tunis, Tunisie
Autres auteurs, équipe:
- Nour houda, GUEDICHE, Médecine interne / Unité de Recherche Maladies Auto-Immunes UR17DN02, Hôpital Militaire Principal d'Instruction de Tunis, Tunis, Tunisie
- Yasmin KHRIFECH, Médecine interne / Unité de Recherche Maladies Auto-Immunes UR17DN02, Hôpital Militaire Principal d'Instruction de Tunis, Tunis, Tunisie
- Bilel, ARFAOUI, Médecine Interne, Hôpital militaire de Bizerte, Tunis, Tunisie
- Sameh, SAYHI, Médecine interne / Unité de Recherche Maladies Auto-Immunes UR17DN02, Hôpital Militaire Principal d'Instruction de Tunis, Tunis, Tunisie
- Faida, AJILI, Médecine interne / Unité de Recherche Maladies Auto-Immunes UR17DN02, Hôpital Militaire Principal d'Instruction de Tunis, Tunis, Tunisie
- Nadia, BEN ABDELHAFIDH, Médecine interne / Unité de Recherche Maladies Auto-Immunes UR17DN02, Hôpital Militaire Principal d'Instruction de Tunis, Tunis, Tunisie

Introduction

L'intelligence artificielle (IA) s'impose progressivement dans le domaine médical, offrant des perspectives intéressantes pour l'aide au diagnostic et à la prise en charge. Son intérêt est particulièrement attendu dans les maladies rares, complexes et parfois méconnues, notamment les maladies auto-inflammatoires (MAI).

➔ L'objectif de cette étude est de comparer les performances de l'IA à celles de médecins spécialistes en médecine interne à travers des vignettes cliniques portant sur trois MAI : la fièvre méditerranéenne familiale (FMF), la maladie de Stil de l'adulte (MSA), et la maladie de VEXAS.

Matériels et méthodes

- Nous avons mené une étude observationnelle transversale à l'aide de trois vignettes représentatives de la FMF, de la MSA et de la maladie de VEXAS, qui ont été élaborées et validées par deux experts seniors de référence. Chaque vignette comportait des questions portant sur l'interprétation du bilan biologique, le diagnostic positif, les outils de confirmation diagnostique et la conduite thérapeutique.
- Les questionnaires ont été diffusés à un groupe de médecins internistes. En parallèle, les mêmes vignettes ont été soumises à 6 modèles d'IA disponibles en accès libres (ChatGpt modèle-5, Gemini, Perplexity, Deepseek, Claude-4 et Mistral).
- Les réponses ont été confrontées au référentiel fourni par les experts, puis codées de manière binaire (réponse juste ou fausse).
- La non-infériorité des performances du groupe IA par rapport au groupe des internistes, a été évaluée par la comparaison des moyennes des taux de réponses correctes entre les deux groupes, en utilisant une marge d'erreur delta fixée à 0.5, seuil jugé acceptable par nos experts.

Résultats

- Vingt-trois médecins internistes ont participé à l'étude. La moyenne d'ancienneté en exercice était de 6,2 ans [2-15].
- **Le taux global de réponses** correctes était 94,6% dans le groupe des internistes contre 66,7% dans le groupe IA. La comparaison de performances entre les 2 groupes a été décrite dans le tableau 1.

Tableau 1 : Comparaison de performances entre les internistes et l'IA

	INTERNISTES	IA	MOYENNE DES DIFFERENES (IA-INTERNISTE)
Interprétation des bilans biologiques	100%	77,8%	-0,33 IC 95% [-0,54 ; -0,12]
Etablissement du diagnostic positif	94,2%	66,7%	-0,83 IC 95% [-1,24 ; -0,41]
Identification des outils de confirmation diagnostique	95,7%	55,6%	-1,20 IC 95% [-1,56 ; -0,84]
Prise en charge thérapeutique	88,4%	55,7%	-0,99 IC 95% [-1,81 ; -0,16]

- **La comparaison des moyennes (IA – internistes)** a montré que l'IA respectait la marge de non-infériorité uniquement pour **l'interprétation du bilan biologique**. En revanche, pour le diagnostic principal les outils de confirmation diagnostique, et la prise en charge médicamenteuse, la marge n'était pas respectée, indiquant une performance inférieure de l'IA par rapport aux internistes.
- **L'analyse indépendante de chaque vignette** a montré que, pour la FMF et la MSA, les performances de l'IA étaient comparables à ce les des internistes sur l'ensemble des items. En revanche, pour la maladie de VEXAS, plus rare, l'IA a obtenu des scores inférieurs, avec une différence moyenne de -2,74 (IC 95% [-3,47 ; -2,01]).
- **L'analyse globale**, regroupant l'ensemble des variables et des vignettes, retrouvait une différence moyenne de -3,35 (IC 95 % [-4,92 ; -1,78]), confirmant que la **non-infériorité de l'IA n'était pas établie sur l'ensemble des tâches.**

Conclusion

Nos résultats montrent que l'IA ne remplace pas les médecins internistes dans le diagnostic et la prise en charge des MAI, en particulier pour les entités rares telles que la maladie de VEXAS. Toutefois, elle montre une performance comparable pour l'interprétation des bilans biologiques, suggérant un rôle potentiel comme outil complémentaire d'aide à la décision clinique.