

REVUE DU MOIS

The
American Journal of
Emergency Medicine

Assessing the precision of artificial intelligence in emergency department triage decisions: Insights from a study with ChatGPT

Sinan Paslı, Abdul Samet Şahin, *et al.*

FEVRIER 2024



Board*Innovation*

OBJECTIF



Examiner la capacité de ChatGPT-4 à prédire les résultats du triage en fonction des règles locales des services d'urgence



Comparaison performances équipe de tri infirmier, chatGPT4 et médecin urgentiste (gold standard)



Board*Innovation*

METHODE



Etude observationnelle prospective monocentrique



Population: ensemble des patients se présentant aux urgences pendant 72h



Randomisation: patients triés selon équipe de triage, chatGPT-4 et l'urgentiste simultanément



Critère de jugement: analyse de la précision par courbes ROC



Board*Innovation*

RESULTATS



758 patients



Accord parfait entre l'équipe de triage et le gold standard (Cohen's Kappa 0.893; $p < 0.001$)



Accord parfait entre GPT4 et le gold standard (Cohen's Kappa 0.899; $p < 0.001$)



Board*Innovation*

LIMITES



Aucune comparaison face à autres IAs



L'incapacité de GPT-4 de garder un historique de saisie et de discussion conséquent



Oubli des règles de triage après un certain nombre d'itérations



Gold-standard peu reproductible (variabilité inter-opérateur)



Board*Innovation*

TAKE HOME MESSAGES



GPT-4, outil possible de triage



Appui décisionnel et non remplacement humain



Limites nombreuses nécessitant nouvelles études



Board*Innovation*