№ 099

CHATBOT APPLICATION VERSUS CONVENTIONAL TRIAGING IN THE PREDICTION OF PATIENT'S DISPOSITION IN EMERGENCY DEPARTMENT

Authors: Shubhash Vijaya Kumar, Hashim Embong
Affiliation: Department of Emergency Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia Medical Centre, Jalan Yaacob Latif, 56000 Bandar Tun Razak, Cheras, Kuala Lumpur, Malaysia

Introduction:

Effective triage is critical to ensure timely care in emergency departments (EDs). This study assessed the performance of ChatGPT-4.0 in assigning triage categories and predicting patient disposition, compared to the conventional method.

Methods:

This retrospective study analysed 499 adult ED cases (June 2023–May 2024) using data from the hospital information system (C-HetS), Hospital Canselor Tunku Muhriz (HCTM). Patients' demographics, presenting complaints, and vital signs were manually entered into ChatGPT-4.0, and the model was prompted to assign a Malaysian Triage System (MTS) category, generate an emergency severity index (ESI) score mapped to MTS zones, and predict disposition (discharge, ward, ICU, mortuary). Outputs were compared to conventional triage and real disposition using cross-tabulation, Cohen's Kappa.
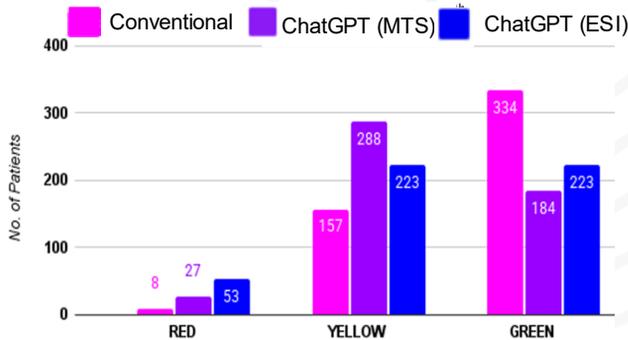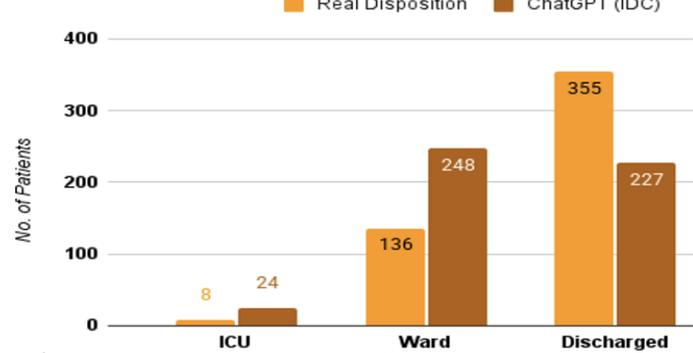
Results:

1) Demographics
Mean age: 54.5 +/- 19.8 y.o
Sex: ♂52.3%, ♀47.7%
Triage: 🟥 2%, 🟨 31%, 🟩 67%

2) Comparison of triage categories

3) Comparison of dispositions

4) Kappa score agreements

Comparison	Cohen's Kappa	SE	95% CI	P value	Interpretation
Conventional vs ChatGPT (MTS)	0.305	0.032	0.242-0.368	<0.001	Fair agreement
Conventional vs ChatGPT (ESI)	0.329	0.033	0.264-0.394	<0.001	Fair agreement
Real Disposition vs ChatGPT	0.288	0.035	0.219 – 0.357	<0.001	Fair agreement

Discussion:

ChatGPT-4.0 showed fair agreement with conventional triage (MTS and ESI) and disposition. It tended to up-triage patients, likely due to algorithmic interpretation of vitals and symptoms without contextual nuance. Agreement was slightly better with ESI, possibly due to its structured 5-level format. Disposition predictions were conservative, favouring ward/ICU admissions over discharge.

Conclusion:

ChatGPT-4.0 showed fair agreement with conventional triage and disposition. It may serve as a decision support tool to enhance consistency, but prospective real-time studies are needed to confirm its safety and effectiveness in live clinical settings.

References:
1. Sarbay, I., et.al. 2023. Performance of emergency triage prediction of an open-access natural language processing-based chatbot application (ChatGPT): A preliminary, scenario-based cross-sectional study. *Turkish Journal of Emergency Medicine* 23(3): 156-161.
2. Paslı, S., et. al. 2024. Assessing the precision of artificial intelligence in emergency department triage decisions: Insights from a study with ChatGPT. *The American Journal of emergency medicine* 78: 170-175.
3. Franc, J. M., et. al. 2024. Repeatability, reproducibility, and diagnostic accuracy of a commercial large language model (ChatGPT) to perform emergency department triage using the Canadian triage and acuity scale. *Canadian Journal of Emergency Medicine* 26(1): 40-46.