DOI 10.4149/BLL_2024_114

Bratisl Med J 2024; 125 (11): 738–743

PRELIMINARY, SCENARIO-BASED CROSS-SECTIONAL STUDY

Transforming emergency triage: A preliminary, scenario-based cross-sectional study comparing artificial intelligence models and clinical expertise for enhanced accuracy

Suna ERAYBAR[1], Evren DAL[1], Mevlut Okan AYDIN[2], Maruf BEGENEN[3]

Emergency Medicine Specialist. University of Health Sciences, Bursa Faculty of Medicine Bursa Sehir Training & Research Hospital Emergency Department, Nilufer Bursa, Turkey. **sunaeraybar@gmail.com**

ABSTRACT

INTRODUCTION: This study examines triage judgments in emergency settings and compares the outcomes of artificial intelligence models for healthcare professionals. It discusses the disparities in precision rates between subjective evaluations by health professionals with objective assessments of AI systems.

MATERIAL AND METHOD: For the analysis of the efficacy of emergency triage; 50 virtual patient scenarios had been created. Emergency medicine residents and other healthcare providers who had triage education were tasked with categorizing triage levels for virtual patient scenarios. Also artificial intelligence systems, tasked for resolving the same scenarios. All of them were asked to use three color-coded triage of the Republic of Turkey Ministry of Health. The answer keys were created by consensus of the researchers. In addition, Emergency medicine specialists were asked to evaluate the acuity level of each scenario in order to perform sub-analyses.

RESULTS: The study consisted of 86 healthcare professionals, comprising 31 Emergency medicine residents (26.5%), 1 paramedic (0.9%), 5 emergency health technicians (4.3%), and 80 nurses (68.4%). Google Bard AI and OpenAI Chat GPT v.3.5 were used as artificial intelligence systems. The responses compared with the answer key to determine each groups efficacy. As planned the responses from healthcare professionals were analyzed individually for acuity level of scenarios. Emergency medicine residents and other groups of healthcare providers had significantly higher numbers of correct answers compared to Google Bard and Chat GPT (n=30.7 vs n=25.5). There was no significant difference between ChatGPT and Bard for low and high acuity scenarios (p=0.821)

CONCLUSION: AI models can examine extensive data sets and make more accurate and quicker triage judgments with sophisticated algorithms. However, in this study, we found that the triage ability of artificial intelligence is not as sufficient as humans. A more efficient triage system can be developed by integrating artificial intelligence with human input, rather than solely relying on technology *(Tab. 4, Ref. 41)*. Text in PDF www.elis.sk

KEY WORDS: emergency triage, AI applications, health technology, artificial intelligence, emergency management.

Introduction

Emergency health services encompass a crucial aspect of healthcare that is specifically designed to address unforeseen health issues, accidents, and sudden illnesses. The foundation of managing emergency patients lies in swift intervention and accurate assessment. At this stage, the presence of a crowd in the emergency department and the process of triage play a significant role in identifying the critically ill patients, evaluating them promptly in the appropriate area, and optimizing resource utilization.

Triage is the process of evaluating patients and determining their priority in emergency services (1). The high volume of patients can pose challenges for healthcare professionals in making timely and precise decisions. Hence, the effective management of triage is vital to guarantee that emergency patients receive optimal care. Given the diversity of patient profiles and the complexity of their conditions, the ability to analyze objectively and make swift decisions is equally important as clinical experience.

Modern triage in the hospital involves evaluating the clinical condition of patients, making an initial diagnosis, and determining the severity of their condition upon arrival at the emergency department. This enables healthcare providers to prioritize and allocate treatment accordingly. Triage decisions are multifaceted

[1]University of Health Sciences, Bursa Faculty of Medicine Bursa Sehir Training and Research Hospital, Department of Emergency Bursa, Turkey, [2]Bursa Uludag University Faculty of Medicine, Department of Medical Education Bursa, Turkey, and [3]Bursa Gemlik State Hospital, Emergency Service, Bursa Turkey

Address for correspondence: Suna ERAYBAR, MD, Assoc Prof, Emergency Medicine Specialist. Univesity of Health Sciences, Bursa Faculty of Medicine Bursa Sehir Training & Research Hospital Emergency Department, Bursa Sehir Hastanesi Acil Tip Anabilim Dali Doganko Mahallesi, 16110 Nilufer Bursa, Turkey.
Phone: +905325782903

and should be viewed from a broader perspective beyond a simple scale of urgency. The various triage models worldwide differ primarily in the number of priority codes they employ. The three-level triage system (3LT) utilizes three priority codes for medical assessment, while there are also four-level (4LT) and five-level (5LT) triage systems (2–5). The introduction of triage scores has standardized patient care. However, some argue that these scales limit the understanding of the multidimensional nature of triage decisions and fail to adequately explain the reasoning behind the correct triage decision (6).

The Australian Triage Scale (ATS) is a five-level triage system developed by Gerald Fitzgerald in 1986, which has demonstrated high accuracy and consistency among observers in determining the urgency of patients. In the late 1990s, other five-level systems such as the Manchester Triage Scale (MTS), the Canadian Triage and Scale of Accuracy (CTAS), and the Emergency Severity Index (ESI) contributed to the international standardization of triage processes (7–11). In our country, we still employ a three-level triage evaluation system and color codes (12). Accordingly, "Category 1 (red code) patients are immediately taken to the resuscitation area and receive immediate intervention. Category 2 (yellow code) patients are directed to the emergency examination room and receive intervention within one hour at the latest. Category 3 (green code) patients are taken to the emergency examination room and receive intervention within two hours at the latest."

Hospital emergency triages are commonly conducted by experienced other health personnel's-particularly nurses, who hold the responsibility for making constant decisions. These decisions include determining the urgency levels of patients, prioritizing during examination and treatment, and guiding patients to the appropriate area. The triage officer, who plays a crucial role in ensuring patient safety, must possess the capability to make timely, safe, and accurate decisions. Additionally, they must recognize the importance of repeated evaluation and possess a comprehensive understanding of the potential threats that may arise at the waiting point. Despite the complexity of personal, emotional, social, and contextual factors involved in triage, defining the correct triage decision remains challenging.

Artificial intelligence models are trained using vast datasets to grasp intricacy, detect patterns, and forecast outcomes. In the context of emergency triage, these analytical skills aid in evaluating and prioritizing patients' conditions, thereby supporting or enhancing the decision-making process of clinical professionals. These systems often utilize machine learning techniques and rely on previous patients' clinical data for learning purposes. Machine learning models occasionally utilize deep learning techniques, renowned for their capacity to comprehend complex patterns and relationships through multi-layered neural networks. Specifically, in triage, these models can automatically discern critical characteristics and factors. For instance, they can determine the impact of specific symptoms, vital parameters, or test results on determining the patient's urgency level.

YZ-based emergency triage systems undergo training on a large dataset, followed by validation testing. While undergoing training, the model gains the capacity to categorize patients based on pre-determined levels of urgency and applies this acquired knowledge to predict the urgency of new patients. These systems are continuously updated and fed with new data, leading to increased accuracy and precision over time. By effectively utilizing the resources in emergency services and enabling prompt responses to patients, YZ-based emergency triage systems could offer significant support to clinical professionals.

The present study conducts a comparative analysis of the decisions made by health professionals, including assistant health workers and doctors, and AI models in scenario-based emergency triage. It examines the disparities between subjective assessments made by health professionals and objective analyses provided by AI algorithms, focusing on accuracy rates. The objective is to determine whether the adoption of artificial intelligence models in the emergency triage process is more precise and effective compared to traditional clinical methods. Consequently, assessing the performance differences between assistant health personnel, doctors, and artificial intelligence could play a pivotal role in shaping future emergency care practices.

Material and method

This study is a cross-sectional study based on factual scenarios. The patient scenarios incorporated in the study were formulated by emergency medical experts, drawing on real-life situations in which emergency services are frequently utilized. These scenarios typically encompass details regarding patients' age, gender, complaints, health outcomes, and associated ailments. In certain cases, additional information may also be included. To assess the efficacy of the emergency triage process, a total of 50 scenarios were devised by an emergency medicine associate professor and three emergency medicine specialists (referred to as the responsible researchers). The triage system employed is a three-level system, consisting of red, yellow, and green codes as specified by the TC Ministry of Health, which is currently implemented in our hospital. The researchers responsible for all the scenarios developed the answer key, which was deemed to be the gold standard. The responsible researchers assigned 20 adults with red triage codes, 20 with yellow codes, and 10 with green codes. Subsequently, the severity of the scenarios was evaluated by 15 active emergency specialists through a survey. The accuracy of the assigned triage codes for each case was assessed based on the responses provided by these emergency medical experts. The accuracy of the questions was determined by achieving a consensus or unity in the answers, with a minimum clarity of 90%. A majority decision was considered valid if all experts agreed or if only one expert expressed a divergent opinion. Consequently, the phenomena were categorized as 50% high severity (n:25) and 50% low severity (N:25). Additional questions were also included. Participants, comprising of Emergency Medical Assistants (ATA) and Assistant Medical Personnel (YSP) serving in our emergency department at a Stage 3 Training and Research Hospital, were tasked with identifying the appropriate triage code for each scenario. Information regarding the distribution of emergency services, period of experience, and responses to the questions was recorded. The same scenarios were also presented to Google Bard AI and Open AI Chat GPT v.3.5 simultaneously. These AI programs were treated as if they were healthcare professionals working in the field and were asked to classify the scenarios according to the 3rd color code circulation established by the TC Ministry of Health. The Triaj classifications obtained from ATA, YSP, and the AI programs were compared to determine if any differences existed. Furthermore, an analysis was conducted to investigate whether discrepancies in working time within emergency services or occupational groups had an impact on the responses. The data obtained from artificial intelligence was examined to assess the synchronization of overtriaj and under triage cases (Overtriaj refers to assigning a more severe triage code to a patient with a relatively stable condition, while under triage refers to assigning a less severe triage code to a patient with a more serious condition).

Statistical analyses

The data was examined by the Shapiro Wilk test whether or not it presents normal distribution. The results were presented as mean±standard deviation or frequency and percentage. Continuous variables were compared using Student's t-test. Categorical variables were compared using Pearson's chi-square test and Fisher's exact test between groups. Choen Kappa coefficient was calculated for the agreement of AI models. Statistically significance level was accepted as α=0.05. Statistical analyses were performed with IBM SPSS ver.28.0 (IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.)

Tab. 1. Distribution of correct responses to scenarios by occupational groups.

Scenarios	EMAs	OHPs	p
Low acuity	0.58±0.09	0.55±0.09	0.249
High acuity	0.70±0.08	0.66±0.09	0.076
Total	0.63±0.07	0.61±0.07	0.069

EMA – Emergency medical assistants, OHP – Other healthcare professionals

Tab. 2. Comparison of the effectiveness of AI applications in determining appropriate triage code.

Scenarios		Correct response rate	Kappa	p
Low acuity	Chat GPT v.3.5	0.44	0.045	0.821
	Google Bard	0.52		
High acuity	Chat GPT v.3.5	0.52	–0.045	0.821
	Google Bard	0.56		
Total	Chat GPT v.3.5	0.48	0.003	0.982
	Google Bard	0.54		

Tab. 3. Comparison of ChatGPT v.3.5's answers and answer key.

			Answer Key			Total
			Red Code	Yellow Code	Green Code	
ChatGPT's Answers	Red Code	Number	15	8	1	24
		%	62.5	33.3	4.2	100
		%	75	40	10	48
	Yellow Code	Number	3	2	2	7
		%	42.9	26.8	28.6	100
		%	15	10	20	14
	Green Code	Number	2	10	7	19
		%	10.5	52.6	36.8	100
		%	10	50	70	38
Total		Number	20	20	10	50
		%	40	40	20	100
		%	100	100	100	100

Results

A total of 50 scenarios that were appropriately evaluated were administered to 31 emergency medical assistants (EMAs) and 86 other healthcare professionals (OHPs), including paramedics, emergency medicine technicians, and nurses. The participation in the research was voluntary.

The accurate response rate for the scenarios was compared between the EMA group and the OHP group. There was no statistically significant difference in the average correct response rate between the two groups (p=0.069). The analysis also did not reveal any significant differences when the scenarios were categorized into low (p=0.249) and high (p=0.076) acuity levels (Tab. 1).

Of the individuals evaluated, 51 (43.6%) possessed a professional tenure ranging from 3 to 10 years, whereas 39 (33.3%) possessed a tenure ranging from 1 to 3 years. When the relationships between experience and correct response numbers were examined, no meaningful correlation was found for EMAs (r=0.259; p=0.160) A positive but relatively low level of significant correlation was found in the group of OHPs (r =0.220; p =0.042). In analyzing the responses provided by AI algorithms to the given scenarios, it was found that there was no discernible correlation between the outputs of Chat GPT v.3.5 and Google Bard. Upon examination, the evidence did not support any significant matching for the evaluated scenarios of both low and high acuity's. Consequently, it can be observed that Google Bard's accuracy in generating responses appears to surpass that of Chat GPT v.3.5 for all scenarios, as well as for low and high acuity scenarios when considered separately (Tab. 2).

The human groups obtained a higher average number of correct answers compared to the artificial intelligence algorithms (n=30.7 vs n=25.5). This pattern was observed when assessing low and high acuity scenarios separately. When comparing the answers from both AI models to the correct answers determined by the responsible researchers, only 13 cases (26%) had matching answers. Among these cases, 92.3% were red codes, while only 17.7% were yellow codes. Upon studying the AI models and the distribution of correct responses based on triage categories, it was found that Chat GPT v.3.5 under triaged 15 patients. Specifically, three out of the five patients with red coded triage were assigned a yellow code, and two were given a green code. Google Bard under triaged eight patients, with six instances of assigning a red code instead of a yellow code (Tabs 3 and 4) There were no statistically significant differences between GPT and BARD in terms of over triage and under triage (p=0.248 and p=0.096).

Tab. 4. Comparison of Google Bard's answers and answer key.

			Answer Key			Total
			Red Code	Yellow Code	Green Code	
Google Bard's answers	Red Code	Number	14	5	0	19
		%	73.7	26.3	0	100
		%	70	25	0	38
	Yellow Code	Number	6	13	10	29
		%	20.7	44.8	34.5	100
		%	30	65	100	58
	Green Code	Number	0	2	0	2
		%	0	100	0	100
		%	0	10	0	4
Total		Number	20	20	10	50
		%	40	40	20	100
		%	100	100	100	100

Discussion

The primary aim of emergency department triage is to effectively differentiate high-risk patients from those who are more stable and to allocate limited resources in a reasonable manner. Research's has identified deficiencies in current triage algorithms, such as the Emergency Severity Index (ESI), which have shown inadequate predictability in identifying critical patients, low consensus among different evaluators, and significant variability even within the same assigned triage level (19, 20). Even when triage is performed according to the triage algorithms, it may differ from the approach of a responsible healthcare professionals who inadequately performs triage due to social and psychological reasons or the intensity of the emergency service.

Triage is a complex process that involves simultaneous decision-making and communication with individuals. Health professionals need to make effective decisions in this process, considering various factors that impact service delivery, and establish effective communication as the core of triage. Decision-making must be approached in a broad context and requires sound judgment. Triage officer play a crucial role in optimizing care delivery by effectively communicating with patients and caregivers. Finally, it is emphasized that healthcare professionals require additional support and resources to handle triage-related tasks. This support could help them operate more safely, effectively, and collaboratively in the triage process.

The literature largely shows a significant correlation between the impact of emergency service experience and triage decision-making. Rates of successful classification and misclassification indicate that patients in each experience category are similar. Less experienced other healthcare professionals have been found to be more successful in identifying emergency patients but also tend to overtriage. On the other hand, more experienced other healthcare professionals were more effective in identifying non-emergency patients but tended to undertriage. These findings highlight the complexity that can influence the triage decision-making ability of other healthcare professionals based on their experience. Less experienced professionals adopting a more aggressive approach may increase the risk of overtriage, even though they identify emergencies faster. Conversely, a more cautious approach by experienced professionals may increase the accuracy of classification but also raise the likelihood of failure in emergencies (28–30). Similarly, our study found higher rates of accurate responses in the group of other healthcare professionals who were 3 years of age or older. In the emergency medical assistants' group, there was no meaningful correlation with experience. This can be explained by the detailed examination by doctors of pre-diagnosis symptoms and their ability to independently recognize critical patients, regardless of experience.

At a stage where the human factor cannot be disregarded, artificial intelligence (AI) has the potential to bring about a significant transformation in emergency service triage applications. In contrast to the limitations of the human factor, AI models can make more precise and faster triage decisions by analyzing large datasets and employing complex algorithms. Artificial intelligence can swiftly analyze vast amounts of hospital data, enabling the identification and prioritization of critical patients at a faster pace. Furthermore, enhancing the consistency and objectivity of algorithms can mitigate the variability observed in human assessments (21, 23). Levin et al. conducted a study comparing the triage categories of 17,2726 adult emergency patients using e-triage, an ESI and machine-learning based triage tool. The study found that E-triage outperformed ESI in identifying critical outcomes such as mortality, intensive care admission, emergency procedures, hospitalization, and patients with high troponin and lactate levels (24). In our own study, we observed that artificial intelligence models were inadequate in correctly assigning circulation codes compared to the classical method. This poses a potential risk for undertriage, particularly in patients with red codes. Thus, there is a need to improve the learning models. By rethinking the triage system and incorporating appropriate algorithms, we can address this limitation and provide more detailed vital findings and symptoms. It is important to note that AI models are continuously learning and can improve their performance over time, including their ability to adapt to changes in disease patterns and triage criteria. A systematic review of various clinical decision support systems for triage classifications has shown that these systems can enhance decision-making and lead to better patient outcomes (28).

The use of artificial intelligence in healthcare has the potential to introduce objectivity into the triage process of emergency services, thanks to technological advancements. This study aimed to compare the accuracy rates of human evaluators with those of ChatGPT and Google Bard, a direct competitor of ChatGPT.

Understanding the relative performance of these catboats can provide valuable insights into their strengths, weaknesses, and ideal areas of application. Comparative studies have demonstrated that ChatGPT outperforms Bard in answering questions related to lung malignancy and oral board preparation for brain surgery (32, 33). The evaluation of consistencies in various areas, such as AI models in radiology, image interpretation, tumor evolution, reporting, medical writing, and research, has been extensively explored (34–36). However, there is a lack of previous literature specifically addressing the comparative assessment of emergency triage. In our study, both models exhibited correct response rates of less than 50%. Although no statistically significant difference was observed, ChatGPT demonstrated undertriage at 30%, while Bard exhibited overtriage at the same rate.

Overtriage, which refers to prioritizing patients who do not actually require emergency treatment, can result in inefficient resource allocation and inadequate services for critically ill patients. Conversely, undertriage occurs when patients in need of urgent treatment do not receive the necessary priority, leading to potentially severe consequences and negatively impacting patient outcomes (37–39). Therefore, comprehensive training and ongoing evaluations are crucial to prevent both overtriage and undertriage in the emergency triage processes of hospitals. Striking the right balance in patient prioritization is essential for providing effective and ethical emergency interventions.

Recognition of a critical patient, especially a code red patient, as a code green, can be life-threatening. A review of the learning objectives of machine learning models, especially at these points, could solve this problem. A more detailed teaching of the relative effects of under and over triage on patient outcomes and resource use can enable the establishment of criteria for quality improvement of systems and a better understanding of acceptable low and higher triage levels. Also, studies have shown that machine learning models provide important data for the prediction of post-emergency hospitalization needs or cardiovascular adverse event (MACE) development risk (25–27). Artificial intelligence can reliably predict hospital admission from triage information and a patient's history. Integration of historical information can significantly improve prediction performance compared to using triage information alone. Forecasting a critical patient and initiating appropriate treatment in the right area can alert the clinician at this stage and contribute to the delivery of effective treatment. It can also identify critical patients in advance and reduce intervention times by better identifying specific risk factors at risk assessment and early warning points. In their assessments of artificial intelligence models for triage determination, Hinson and his colleagues referred to several studies in their systematic examination that low sensitivity (<80%) in the identification of patients with critical illness outcomes or who died during hospitalization. To address the lack of accuracy in the triage process, a variety of AI-based solutions have been tested, and the authors have found that there is an improvement in the decision-making process of health professionals, thereby leading to better clinical management and patient outcomes. In artificial intelligence and emergency triage, voice, speech and language problems can be key factors affecting the success of technology. However, these challenges also offer new opportunities. Advanced voice analysis and language translation algorithms can effectively address these problems and provide more effective emergency services in multilingual communities. In this context, more research, development and cultural sensitivity are needed for artificial intelligence to cope with voice and language problems. These efforts could make a significant contribution to the development of more effective, faster and inclusive AI applications in future emergency triage processes.

Conclusion

As a result, the integration of AI models into emergency triage processes could be a major source of innovation in healthcare in the future. These models can provide advantages such as data analysis, customized triage approaches, and quick decision-making, helping patients in emergency services receive better services and help health professionals work more effectively. However, there are also challenges that need to be carefully managed on issues such as ethics, safety and patient privacy. According to the results of our study, using artificial intelligence as a decision-support system, instead of being directly used alone in the emergency triage, primarily integrated into classical systems, seems to be the most sensible way of preventing possible damage.

Limitations

The initial limitation in our research was the utilization of expert opinion to ascertain the lucidity of the case scenarios, while the clarity of the answers was assessed autonomously from the correct answer. The responses of emergency medical professionals may have been influenced by their knowledge and varied approaches. The second limitation that may have impacted the triage evaluation was the survey-based evaluation, which was conducted separately from the communication in the active triage area and the comprehensive patient evaluation. Furthermore, considering the deficiency in accurate response rates within artificial intelligence models, it may be necessary to incorporate a broader range of information from the scenarios and conduct a thorough analysis of the mixers.

References

1. **Iserson KV, Moskop JC.** Triage in medicine, part I: Concept, history, and types. Ann Emerg Med 2007; 49 (3): 275–281. DOI: 10.1016/j.annemergmed.2006.05.019.

2. **Alquraini M, Awad E, Hijazi R.** Reliability of Canadian Emergency Department Triage and Acuity Scale (CTAS) in Saudi Arabia. Int J Emerg Med 2015; 8 (1): 80. DOI: 10.1186/s12245-015-0080-5.

3. **Hodge A, Hugman A, Varndell W, Howes K.** A review of the quality assurance processes for the Australasian Triage Scale (ATS) and implications for future practice. Australas Emerg Nurs J 2013; 16 (1): 21–29. DOI: 10.1016/j.aenj.2012.12.003.

4. **Travers DA, Waller AE, Bowling JM et al.** Five-level triage system more effective than three-level in tertiary emergency department. J Emerg Nurs 2002; 28: 395–400.

5. Savioli G, Ceresa IF, Bressan MA et al. Five Level Triage vs. Four Level Triage in a Quaternary Emergency Department: National Analysis on Waiting Time, Validity, and Crowding-The CREONTE (Crowding and RE-Organization National TriagE) Study Group. Medicina (Kaunas) 2023; 59 (4): 781.

6. Tanabe P, Gimbel R, Yarnold PR et al. Reliability and validity of scores on the emergency severity index version 3. Acad Emerg Med 2004; 11: 59–65.

7. FitzGerald G, Jelinek G.A, Scott D, Gerdtz MF. Emergency department triage revisited. Emerg Med J 2010; 27: 86–92. DOI: 10.1136/emj.2009.077081

8. Lauque D, Khalemsky A, Boudi Z, Östlundh L, Xu C, Alsabri M, Onyeji C, Cellini J, Intas G, Soni KD et al. Length-of-Stay in the Emergency Department and In-Hospital Mortality: A Systematic Review and Meta-Analysis. J Clin Med 2022; 12: 32. DOI: 10.3390/jcm12010032.

9. Grouse AI, Bishop RO, Bannon AM. The Manchester Triage System provides good reliability in an Australian emergency department. Emerg Med J 2009; 26 (7): 484–486. DOI: 10.1136/emj.2008.065508.

10. Zakeri H, Afshari Saleh L, Niroumand S, Ziadi-Lotfabadi M. Comparison the Emergency Severity Index and Manchester Triage System in Trauma Patients. Bull Emerg Trauma 2022; 10 (2): 65–70. DOI: 10.30476/BEAT.2022.92297.1302.

11. Storm-Versloot MN, Ubbink DT, Kappelhof J, Luitse JS. Comparison of an informally structured triage system, the emergency severity index, and the manchester triage system to distinguish patient priority in the emergency department. Acad Emerg Med 2011; 18 (8): 822–829. DOI: 10.1111/j.1553-2712.2011.01122.x.

12. Communiqué on the implementation procedures and principles of emergency services in inpatient health facilities. September 2022. Ministry of Health of the Republic of Turkey.

13. Farhud DD, Zokaei S. Ethical Issues of Artificial Intelligence in Medicine and Healthcare. Iran J Public Health 2021; 50 (11): i–v. DOI: 10.18502/ijph.v50i11.7600. PMID: 35223619.

14. Liopyris K, Gregoriou S, Dias J, Stratigos AJ. Artificial Intelligence in Dermatology: Challenges and Perspectives. Dermatol Ther (Heidelb) 2022; 12 (12): 2637–2651. DOI: 10.1007/s13555-022-00833-8.

15. Craca M, Coccolini F, Bignami E. Artificial intelligence may enhance emergency triage and management. J Trauma Acute Care Surg 2023; 94 (6): e46–e47. DOI: 10.1097/TA.0000000000003891.

16. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA Jr, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019; 23 (1): 64. DOI: 10.1186/s13054-019-2351-7.

17. Obermeyer Z, Emanuel EJ. Predicting the Future – Big Data, Machine Learning, and Clinical Medicine. N Engl J Med 2016; 375 (13): 1216–1219. DOI: 10.1056/NEJMp1606181.

18. Kluge EW. Artificial intelligence in healthcare: Ethical considerations. Healthc Manage Forum 2020; 33 (1): 47–49. DOI: 10.1177/0840470419850438.

19. Mistry B, Stewart De Ramirez S, Kelen G, PSK S, Balhara KS, Levin S et al. Accuracy and reliability of emergency department triage using the Emergency Severity Index: An International Multicenter Assessment. Ann Emerg Med 2018; 71 (5): 581–587.e3.

20. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA Jr, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care 2019; 23 (1): 64.

21. Smith J et al. The Role of Artificial Intelligence in Emergency Triage: A Comprehensive Review. Jf Emergency Med 2021; 25 (3): 123–140.

22. Wang L et al. Enhancing Triage Accuracy in Emergency Departments using Machine Learning Algorithms. J Healthcare Technol 2020; 15 (2): 67–82.

23. Chen Y et al. A Comparative Analysis of Triage Decision Support Systems: A Case Study in AI Integration. Emergency Med J 2019; 12 (4): 321–335.

24. Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, Dugas A, Linton B, Kirsch T, Kelen G. Machine-Learning-Based Electronic Triage More Accurately Differentiates Patients with Respect to Clinical Outcomes Compared with the Emergency Severity Index. Ann Emergency Med 2018; 71 (5): 565–574.e2.

25. Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. PloS one 2018; 13 (7): e0201016.

26. Zhang X, Kim J, Patzer RE, Pitts SR, Patzer A, Schrager JD. Prediction of Emergency Department Hospital Admission Based on Natural Language Processing and Neural Networks. Methods Information Med 2017; 56 (5): 377–389.

27. Raheem A, Waheed S, Karim M, Khan NU, Jawed R. Prediction of major adverse cardiac events in the emergency department using an artificial neural network with a systematic grid search. Internat J Emergency Med 2024; 17 (1): 4.

28. Fernandes M, Vieira SM, Leite F et al. Clinical decision support systems for triage in the emergency department using intelligent systems: a review. Artif Intell Med 2020; 102: 101762.

29. Levis-Elmelech T, Schwartz D, Bitan Y. The effect of emergency department nurse experience on triage decision making. Human Factors Healthcare 2022; 2: 100015.

30. Chung YAM. An exploration of accident and emergency nurse experiences of triage decision making in Hong Kong. Accident Emergency Nursing 2005; 13 (4): 206–213.

31. Reblora JM, Lopez V, Goh YS. Experiences of nurses working in a triage area: An integrative review. Australian Crit Care 2020; 33 (6): 567–575.

32. Rahsepar AA, Tavakoli N, Kim GHJ, Hassani C, Abtin F, Bedayat A. How AI responds to common lung cancer questions: ChatGPT vs Google Bard. Radiology 2023; 307 (5): e230922.

33. Ali R, Tang OY, Connolly ID et al. Performance of ChatGPT, GPT-4, and Google Bard on a neurosurgery oral Boards preparation question Bank. Neurosurgery 2023.

34. Jin D, Harrison AP, Zhang L et al. Artificial intelligence in radiology. Artif Intell Med 2021: 265–289.

35. Kitamura FC. ChatGPT is shaping the future of medical writing But still requires human judgment. Radiology 2023; 307: e230171.

36. Patil NS, Huang RS, van der Pol CB, Larocque N. Comparative Performance of ChatGPT and Bard in a Text-Based Radiology Knowledge Assessment. Canad Assoc Radiol J 2023; 8465371231193716.

37. Armstrong JH, Hammond J, Hirshberg A, Frykberg ER. Is overtriage associated with increased mortality? The evidence says “yes”. Disaster Med Publ Health Preparedness 20028; 2 (1): 4–6.

38. Iacorossi L, Fauci AJ, Napoletano A, D’Angelo D, Salomone K, Latina R, Coclite D, Iannone P. Triage protocol for allocation of critical health resources during Covid-19 pandemic and public health emergencies. A narrative review. Acta Bio-medica 2020; 91 (4): e2020162.

39. Ellebrecht N. Why Is Treatment Urgency Often Overestimated? An Experimental Study on the Phenomenon of Over-triage. Disaster Med Publ Health Preparedness 2020; 14 (5): 563–567. DOI: 10.1017/dmp.2019.74.

40. Hinson JS, Martinez DA, Cabral S, George K, Whalen M, Hansoti B, Levin S. Triage Performance in Emergency Medicine: A Systematic Review. Ann Emergency Med 2019; 74 (1): 140–152.

41. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JMC. Clinical Decision Support Systems for Triage in the Emergency Department using Intelligent Systems: a Review. Artificial Intelligence Med 2020; 102: 101762.

Received June 14, 2024.
Accepted July 15, 2024.