American Journal of Emergency Medicine 75 (2024) 72–78

Contents lists available at ScienceDirect

American Journal of Emergency Medicine

journal homepage: www.elsevier.com/locate/ajem

Performance of Google bard and ChatGPT in mass casualty incidents triage

Rick Kye Gan [a,*], Jude Chukwuebuka Ogbodo [a,b], Yong Zheng Wee [c], Ann Zee Gan [d,e], Pedro Arcos González [a]

[a] Unit for Research in Emergency and Disaster, Faculty of Medicine and Health Sciences, University of Oviedo, Oviedo 33006, Spain
[b] Department of Primary Care and Population Health, Medical School, University of Nicosia, Nicosia 2408, Cyprus
[c] Faculty of Computing & Informatics, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia
[d] Tenghilan Health Clinic, Tuaran 89208, Sabah, Malaysia
[e] Hospital Universiti Sains Malaysia, 16150 Kota Bharu, Malaysia

ARTICLE INFO

Article history:
Received 20 August 2023
Received in revised form 3 October 2023
Accepted 24 October 2023

Keywords:
Mass casualty incident
Triage
Disaster medicine
Artificial intelligence

ABSTRACT

Aim: The objective of our research is to evaluate and compare the performance of ChatGPT, Google Bard, and medical students in performing START triage during mass casualty situations.
Method: We conducted a cross-sectional analysis to compare ChatGPT, Google Bard, and medical students in mass casualty incident (MCI) triage using the Simple Triage And Rapid Treatment (START) method. A validated questionnaire with 15 diverse MCI scenarios was used to assess triage accuracy and content analysis in four categories: "Walking wounded," "Respiration," "Perfusion," and "Mental Status." Statistical analysis compared the results.
Result: Google Bard demonstrated a notably higher accuracy of 60%, while ChatGPT achieved an accuracy of 26.67% ($p = 0.002$). Comparatively, medical students performed at an accuracy rate of 64.3% in a previous study. However, there was no significant difference observed between Google Bard and medical students ($p = 0.211$). Qualitative content analysis of 'walking-wounded', 'respiration', 'perfusion', and 'mental status' indicated that Google Bard outperformed ChatGPT.
Conclusion: Google Bard was found to be superior to ChatGPT in correctly performing mass casualty incident triage. Google Bard achieved an accuracy of 60%, while chatGPT only achieved an accuracy of 26.67%. This difference was statistically significant ($p = 0.002$).
© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

The field of AI natural language processing has witnessed a significant transformation with the advent of advanced language models, leading to remarkable advancements in various tasks. Notably, prominent examples of these models are ChatGPT and Google Bard [1,2].

ChatGPT, an advanced AI language model initially released in November 2022, shows great promise in various medical applications. One of its significant contributions is in the field of diagnosis, where it has consistently outperformed traditional tools like Google search and symptom checkers in online diagnosis [3-5].

Additionally, ChatGPT serves as an educational tool for emergency physicians and paramedics. A proof-of-concept study has confirmed its effectiveness in providing engaging and enjoyable teaching experiences for medical professionals [6]. Moreover, ChatGPT plays a substantial role in public health by supporting disease surveillance, outbreak management, and resource allocation [7,8].

While the Google Bard is an AI chatbot released by Google on March 21, 2023. It mimics human-like conversations using natural language processing and machine learning. The Bard can be used across digital platforms, giving genuine responses and helping in areas like emergency medicine, public health, and disaster management [9].

In a study comparing ChatGPT, Google Bard, and the paid version of GPT-4 for preparing for neurosurgery oral boards with advanced management cases, the paid GPT-4 scored remarkably well at 82.6%. It did better than both the free ChatGPT3.5 and Google Bard [10].

Both ChatGPT and Google Bard have limitations and challenges in how they are used. Their quality and reliability can be affected by inconsistent information, infrequent updates, and a lack of validation by experts. Moreover, because natural language is complex and often ambiguous, there can be errors and misunderstandings in how users input their queries [11,12].

A mass casualty incident (MCI) involves a significant number of individuals requiring medical attention [13]. These incidents can result from

* Corresponding author.
E-mail address: ganrick@uniovi.es (R.K. Gan).

https://doi.org/10.1016/j.ajem.2023.10.034
0735-6757/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

various causes such as natural disasters, accidents, or terrorist attacks. While the specifics can differ between countries, MCIs generally involve situations that overwhelm local medical resources [14]. The defining feature is when there are more patients than available healthcare resources can handle, typically exceeding ten patients [15].

The START triage method was established in 1983 by the Newport Beach Fire Department and Hoag Hospital in California. It plays a crucial role in quickly and efficiently categorizing MCI victims by injury severity. [16] START, which stands for Simple Triage and Rapid Treatment, employs four categories to prioritize victims: deceased/expectant (black), immediate (red), delayed (yellow), and walking wounded/minor (green). These categories consider factors like the victim's ability to walk, respiratory rate, pulse or capillary refill, and mental status, with the sole intervention being opening the airway for non-breathing victims.

The START triage helps first responders in determining treatment priorities and evacuation, ensuring efficient resource allocation [16]. It involves critical decisions about providing on-site treatment or immediate transportation to the nearest hospital [17]. Accurate and consistent categorization is vital during an MCI, as errors (over-triage or under-triage) can significantly impact disaster response, potentially leading to loss of lives and resource constraints [18]. While AI tools in emergency medicine have been explored in previous research, there is limited comparison of these chatbots' accuracy in START triage.

The objective of our research is to evaluate and compare the performance of ChatGPT, Google Bard, and medical students in performing START triage during mass casualty incidents. We used a validated questionnaire [19] and compared the results of MCI triage.

2. Methods

Our study is a cross-sectional analysis to assess how well ChatGPT and Google Bard perform in mass casualty incident (MCI) triage using Simple Triage And Rapid Treatment (START) triage. We use mixed methods of quantitative descriptive analysis to evaluate their overall MCI triage performance and content analysis to assess their performance in four headings 1. Walking wounded, 2. Respiration, 3. Perfusion, and 4. Mental state. Lastly, we compare their accuracy to that of medical students using the same triage questionnaire performed by Sapp et al. [19].

2.1. Materials

For this research, we used ChatGPT-3.5, developed by OpenAI in San Francisco, CA, and freely available for public use [20]. We also used the Google Bard model, which ran on PaLM 2 and was updated on June 7, 2023. The data was collected in Malaysia and analyzed on July 5, 2023 [21]. No ethical considerations are needed since all data used are from open source and secondary data.

In the study, with written permission, we employed a validated mass casualty incident triage questionnaire from Sapp et al. [19]. The questionnaire's 15 scenarios were expertly crafted by Emergency Medical Services (EMS) Medical Directors and Emergency Faculty affiliated with the University of North Carolina School of Medicine. They have extensive training and experience in crisis management and emergency aid. The scenarios were carefully selected to ensure diverse triage levels and START criteria adherence.

Each scenario provided detailed patient information, including age, symptoms, vital signs (such as breathing rate, heart rate, and capillary refill), and the method of transportation to the hospital. The triage questionnaire considered various medical and traumatic conditions, excluding sarin gas exposure [19]. The triage questionnaire included four cases classified as “Red” (Immediate), four as “Yellow” (Delayed), four as “Green” (Minor), and three as “Black” (Deceased) status. The complete triage questionnaire is available in Appendix 1.

The mean accuracy scores of medical students were obtained from a previously published study conducted with two consecutive classes of first-year students in 2008 and 2009, a total of 315 students, at the University of North Carolina School of Medicine in Chapel Hill. These students had participated in START triage training and had completed a paper-based triage exercise during their orientation. The study's findings were reported by Sapp et al. [19].

2.2. Data collection

Our study tested ChatGPT and Google Bard's ability to perform START triage using the prompts ‘Do you know START triage?’ and ‘Can you perform START triage?’. After confirming their ability, we individually presented questions from the mass casualty triage questionnaire (see Appendix 1). We recorded all responses from both AI chatbots in an Excel spreadsheet for detailed analysis.

2.3. Data analysis

After the mass casualty incident START triage, we analyzed ChatGPT and Google Bard's performance on the mass casualty triage questionnaire. We categorized their performance into three types: 1) Correct-triage, 2) Over-triage, and 3) Under-triage [22]. We checked ChatGPT and Google Bard's triage performance by collecting data and calculating their proportions in all three triage areas.

For content analysis, we thoroughly analyzed the responses of ChatGPT and Google Bard. We focused on four themes: walking wounded, respiration, perfusion, and mental status. Using START adult triage guidelines as a reference, we categorized their performance as either correct or incorrect to assess their accuracy in these areas [23]. According to the questionnaire, correct responses were those that appropriately reflected the patient's circumstance and triage decision using the START triage algorithm. Responses that didn't match were considered wrong. We recorded all questions and replies from ChatGPT and Google Bard performances in a Microsoft Excel spreadsheet.

We analyzed the data using percentages and two non-parametric statistical tests. First, we used the Kruskal-Wallis test to compare the average accuracies of three groups (ChatGPT, Bard, and medical students) due to our small sample size and non-normally distributed data. Second, we used the Mann-Whitney *U* test to compare any two groups, and we performed these tests using IBM SPSS Statistics 23.

3. Results

We categorized the results into three groups: Correct-triage, Over-triage, and Under-triage, based on response accuracy. ChatGPT's performance showed a notable over-triage rate in 10 out of 15 cases (66.67%), assigning higher care than necessary. Only 4 out of 15 cases (26.67%) were accurately identified for the appropriate level of care. There was a very small number of under-triage cases, 1 out of 15 (6.67%), indicating missed opportunities for higher care when needed.

Google Bard correctly identified and assigned 9 out of 15 patients to the right level of care, making accurate decisions for 60% of them. However, for 6 patients (40%), it assigned a higher level of care than necessary, possibly leading to excess attention and resources. On the positive side, none of the patients were under-triaged, meaning all who needed higher care were correctly identified and treated accordingly. Medical students performance from Sapp et al. shows overall accuracy of (64.3%), overall under-triage (12.6%), and overall over-triage(17.82%) [19]. Overall performance is shown in Fig. 1.

This study compared the overall performance of three methods: ChatGPT, Google Bard, and medical students. The Kruskal-Wallis test showed a significant difference among the three performances ($p = 0.002$). Additionally, the Mann-Whitney *U* test revealed a significant difference in performance between ChatGPT and Google Bard ($p = 0.002$),

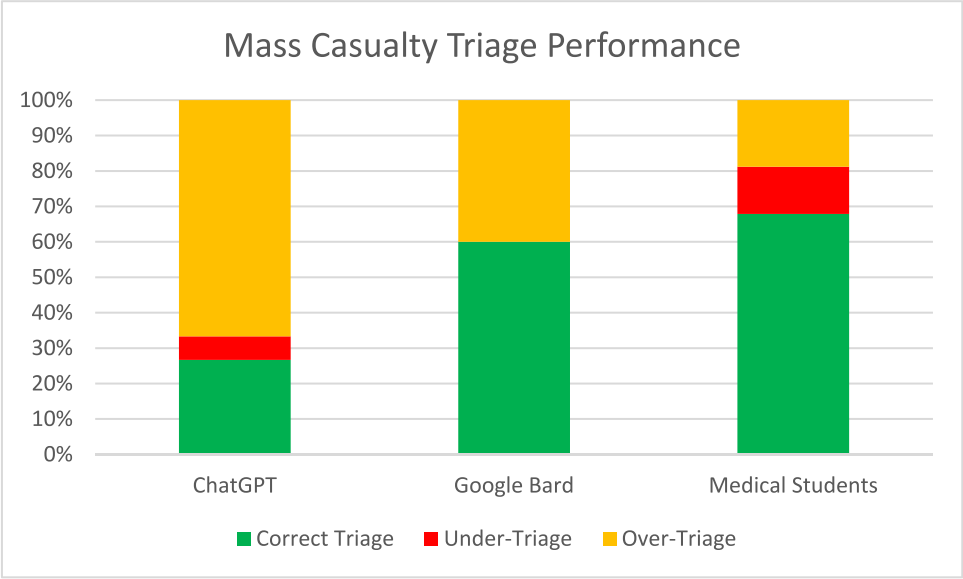

Fig. 1. Performance of START Triage Accuracy of ChatGPT, Google Bard, and medical students.

with Google Bard demonstrating higher mean accuracy compared to ChatGPT.

Moreover, the test indicated a significant difference between ChatGPT and medical students' performance ($p = 0.000$), with medical students performing more accurately. However, there was no statistically significant difference between Google Bard and medical students ($p = 0.211$).

3.1. Content analysis

The content analysis comprises four categories: "Walking wounded," "Respiration," "Perfusion," and "Mental Status." ChatGPT's performance in identifying scenarios is notable.

In the "Walking wounded" category, ChatGPT accurately identified 12 out of 15 scenarios, achieving 80% accuracy (Fig. 2). Some of the incorrect responses are as follows:

Q1*: "Walking Ability: No information is provided about the patient's walking ability…".*

However, it has been clearly stated that the patient is capable of assisting someone else at the scene.

Conversely, in the "Respiration" analysis, ChatGPT scored 4 out of 15 (26.67%). Some of the incorrect responses are as follows:

Q2*: "Respiratory Status: There is no specific information provided regarding the patient's respiratory status…".*

Where Q2 clearly stated that the respiratory rate of the patient was $R = 12$. Hinting that ChatGPT is unable to correctly identify medical abbreviations.

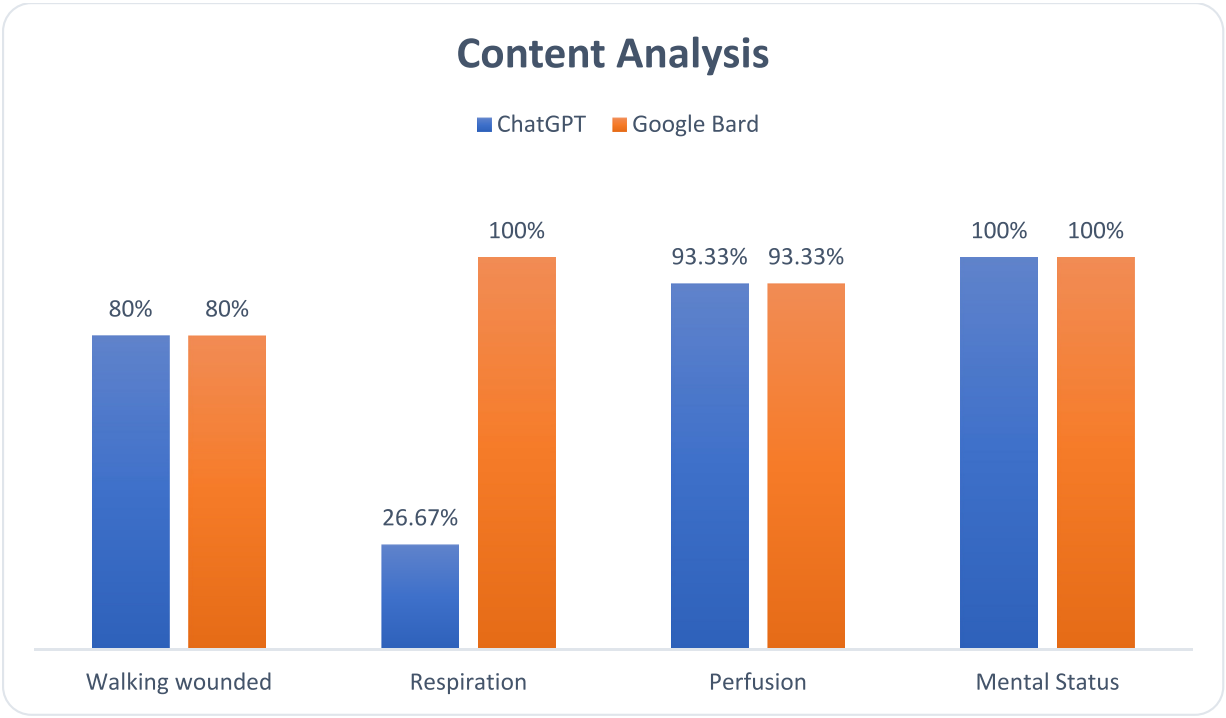

Fig. 2. Content analysis result of ChatGPT and Google Bard Performance in Mass Casualty Incident Triage.

Q3: *"Respiratory Status: The patient is complaining of shortness of breath, indicating respiratory distress…".*

Although it is mentioned that the patient is having shortness of breath after running, the respiratory rate $R = 24$ which is acceptable according to the START triage algorithm.

In contrast, the "Perfusion" analysis demonstrated impressive capabilities, with an outstanding accuracy rate of 93.33%, by correctly identifying 14 out of 15 scenarios. Some of the incorrect responses are as follows:

Q10: *"Perfusion Status: No specific information is provided about the patient's perfusion status or circulation…".*

But the question clearly stated that the patient has a capillary refill of >2 s and barely palpable radial pulse.

The "Mental Status" analysis showcased remarkable proficiency, with a perfect score of 15 out of 15 (100%), highlighting ChatGPT's competence in this domain. Some of the responses are as follows:

Q11: *"Mental Status: The patient is described as alert but has symptoms of slurred speech and blurry vision..."*

As correctly described in the questionnaire.

In Google Bard, the "Walking wounded" category achieved 80% accuracy (12/15). Some of the incorrect responses are as follows:

Q1: Google Bard did not mention anything regarding "Walking wounded," while it is clearly stated that the patient was able to assist someone.

"Respiration" in Google Bard demonstrated a perfect accuracy of 100% (15/15). Some of the correct responses are as follows:

Q12: "*The patient's respiratory rate of 36 breaths per minute is within the normal range…*".

"Perfusion" in Google Bard achieved 93.33% accuracy (14/15). The incorrect response is as follows:

Q2: "*…The patient also has a capillary refill time of <2 seconds, which indicates that they have poor perfusion…*".

This is incorrect since a capillary refill time of <2 s is considered normal according to START triage.

In Google Bard, "Mental Status" exhibited flawless accuracy of 100% (15/15). Some of the correct responses are as follows:

Q7: "*…The patient is unresponsive after having a seizure…".*

Which correctly shows that the patient is immobile.

4. Discussion

Overall, the findings underscore ChatGPT's and Google Bard's strengths and limitations across different areas. In ChatGPT, the high accuracy in "Perfusion" and "Mental Status" showcases the model's capabilities, while lower accuracy in "Respiration" indicates areas for further development, especially in recognizing medical abbreviations. Whereas in Google Bard, these findings indicate varying levels of accuracy across the categories, suggesting possible areas for improvement in "Walking wounded" and "Perfusion."

By reducing over-triage and enhancing the precision of mass casualty incident triage, we can optimize the allocation of resources for more effective emergency medical care. Investigating further and improving the triage system is essential to achieve better outcomes and enhance patient care during emergencies.

Since the advent of AI-powered chatbots like ChatGPT and Google Bard, they have garnered considerable attention from the media and professionals across diverse industries. These technologies are being hailed as transformative milestones in the ongoing AI revolution. AI chatbots are software applications that leverage natural language processing and machine learning to engage users through text or voice interfaces. Their applications in the field of medicine are particularly noteworthy [24]. AI chatbots could be used in medicine for several purposes, one of which is to share medical information with patients. AI chatbots could have advanced algorithms that are trained on extensive healthcare datasets, including disease symptoms, diagnostic techniques, markers, and available treatment options. They can answer basic patient queries and offer helpful resources, especially when the patient's primary healthcare provider is unavailable [24,25].

Chatbots could potentially provide real-time, evidence-based recommendations to healthcare providers to improve patient outcomes [12]. They can suggest appropriate treatment options, flag potential drug interactions, and provide clinical guidelines for complex medical cases [26]. This is a valuable capability because it can help healthcare providers make better decisions about patient care [26]. Chatbots can access and process large amounts of data quickly, which allows them to identify potential problems that may not be immediately obvious to human providers [12]. They could also provide up-to-date information on the latest treatment options and guidelines [12,26]. Therefore, chatbots can help to improve patient safety and ensure that patients receive the best possible care.

AI chatbots can be used to automate the medical recordkeeping process by generating summaries of patient interactions and medical histories. This can save time and improve accuracy, as chatbots can access and process large amounts of data quickly. Chatbots can also extract relevant information from patient records, such as lab results or imaging reports. This can help healthcare providers quickly identify important information and make better decisions about patient care [12,26]. In medical research, Chatbots possess the capability to scrutinize vast quantities of medical data, identifying patterns and trends that can serve as valuable guidance in the development of innovative treatments and diagnostic tools [26]. Furthermore, they exhibit proficiency in medical transcription, efficiently converting medical reports and notes into written form, thereby eliminating reliance on human transcriptionists and significantly reducing the likelihood of errors [26].

Chatbots powered by AI could monitor and track health status by gathering information from users, such as temperature, oxygen saturation, and exposure history. This data can be analyzed to keep track of changes in health status and identify potential issues at an early stage. These chatbots can also give alerts and reminders for follow-up actions, like isolation, quarantine, or vaccination. This can help enhance patient outcomes and prevent the spread of diseases [11]. Chatbots can provide emotional support, counselling, and therapy to users with mental health issues. They are trained on a large corpus of text data and can offer personalized support [11,25].

One issue that needs to be addressed is the potential risk to privacy and security when it comes to AI Chatbots. There is a possibility that they could unintentionally reveal personal or health information to unauthorized parties, leaving them vulnerable to harmful attacks. It is crucial to obtain proper consent and provide adequate information to users regarding data handling and storage to avoid violating data protection laws and ethical standards [11,25]. Furthermore, there is the possibility that chatbots might exhibit a deficiency in empathy or cultural sensitivity during their interactions, which could subsequently impact user satisfaction and engagement [11,25].

Human oversight and accountability are crucial aspects to consider. Due to their limitations, AI chatbots may not be able to handle situations requiring human judgment, empathy, or specialized expertise. Additionally, they might lack transparency in generating outputs or recommendations. Therefore, constant supervision and review by human professionals become imperative, as they hold responsibility for the final decisions and actions [12].

Further research is essential to investigate the applicability of AI tools, including ChatGPT and Google Bard, in the context of mass casualty incidents (MCIs) triage and disaster medicine. This research should encompass various MCI modalities, such as SALT triage, which aligns with the Model Uniform Core Criteria for mass casualty incidents. Expanding our understanding of the capabilities and effectiveness of these AI technologies in such critical situations.

5. Conclusion

Google Bard was found to be superior to ChatGPT in correctly performing mass casualty incident triage. Google Bard achieved an accuracy of 60%, while chatGPT only achieved an accuracy of 26.67%. This difference was statistically significant ($p = 0.002$).

Author's contribution

All authors made substantial contributions to the design, data collection, data analysis, and writing; drafted the article and reviewed the final manuscript; gave final approval of the version to be published and agreed to be accountable for all aspects of the work. All figures in this manuscript were produced by the first author.

CRediT authorship contribution statement

Rick Kye Gan: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jude Chukwuebuka Ogbodo:** Writing – review & editing, Writing – original draft, Software, Formal analysis, Data curation. **Yong Zheng Wee:** Visualization, Investigation, Data curation, Conceptualization. **Ann Zee Gan:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Pedro Arcos González:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

All authors declared no conflict of interest.

Acknowledgement

We would like to thank the editor-in-chief Prof. Dr. J. Douglas White and two anonymous reviewer who guided us and provided us with so much insight to improve the quality of our article.

Appendix 1

Triage Scenarios	
R Y G B	25 y/o male appears alert c/o headache and cough after assisting someone from scene. Denies trauma. R 18
R Y G B	32 y/o c/o is unable to walk due to severe weakness, sweating and uncontrolled diarrhea. No visible trauma. Cap refill <2 sec, R 12
R Y G B	33y/o male asthmatic alert c/o shortness of breath after running from scene. Vital signs R 24
R Y G B	34 y/o female alert c/o severe R leg pain with a fractured lower leg with bone sticking out of the skin after being pushed down approx 6 stairs. Bleeding appears to be minor. Denies other complaint. Cap refill <2 sec, radial pulse palpable. R 26.
R Y G B	38y/o confused male who initially c/o severe shortness of breath and crushing chest pain but became unresponsive in the ambulance. CPR in progress by EMS on arrival at hospital.
R Y G B	40y/o female is confused, falling repeatedly and unable to stand. Is unable to follow simple commands. Cap refill <2sec, good radial pulse. R 14
R Y G B	44 y/o male having seizure that stops on arrival at the hospital. He is unresponsive,There is a bruise on his forehead. Cap refill <2 sec, good radial pulse. R 14
R Y G B	45y/o male is walking alert c/o right wrist pain after fall. Evident deformity of R wrist. No visible bleeding.Vital signs R 20
R Y G B	48y/o alert male with large cuts on thigh after putting his leg through a glass door. He is unable to stand. There is visible muscle and tendon with controllable bleeding. Cap refill <2 sec, good radial pulse. R 18
R Y G B	51y/o male unresponsive in drivers seat of car at hospital entrance. He has no visible trauma. His face is blue and he is not breathing. Cap refill >2 sec, barely palpable radial pulse. R O after opening his airway.
R Y G B	55 y/o alert male c/o headache, slurred speech and blurry vision. Falls repeatedly on standing but has no visible trauma. Cap refill < 2 sec, good radial pulse. R 14.
R Y G B	56 y/o female alert, c/o shortness of breath, chest tightness, blurry vision, drooling and weakness. Cap refill <2 sec, good radial pulse. R 36
R Y G B	56 y/o female unresponsive according to EMS she, " Just stopped breathing!" No visible trauma. Cap refill <2 sec, weak radial pulse. R 0 after opening her airway.
R Y G B	60 y/o female police officer presents after driving herself to hospital. She was the first arriving officer and aided the wounded. She is alert c/o generalized weakness, wet with sweat, and drooling. No visible trauma R 14
R Y G B	63 y/o female face down unresponsive. Her face is blue and she does not appear to be breathing.There is no visible trauma. Cap refill <2 sec, good radial pulse. R 0, Begins breathing spontaneously after opening her airway.

Sapp © 2010 Prehospital and Disaster Medicine

Fig. A1. Shows mass casualty incident triage questionnaire by Sapp et al. © Sapp 2010 All rights reserved. Used with permission from Prehospital and Disaster Medicine, WADEM.

Appendix 2. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajem.2023.10.034.

References

[1] Rahsepar AA, Tavakoli N, Kim GHJ, Hassani C, Abtin F, Bedayat A. How AI responds to common lung cancer questions: ChatGPT versus Google bard. Radiology. 2023 Jun 1.;307(5).

[2] Eysenbach G. The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. JMIR Med Educ. 2023 Mar 6(9):e46885.

[3] Bilal M, Jamil Y, Rana D, Shah HH. Enhancing awareness and self-diagnosis of obstructive sleep apnea using AI-powered Chatbots: the role of ChatGPT in revolutionizing healthcare. Ann Biomed Eng. 2023 Jun 30..

[4] Irwin P, Jones D, Fealy S. What is ChatGPT and what do we do with it? Implications of the age of AI for nursing and midwifery practice and education: an editorial. Nurse Educ Today. 2023 Aug;127:105835.

[5] Ramadan I. The Main and basic differences between the Google search and ChatGPT application. Int J Sci Res Pub. 2023 Jan 24.;13(1).

[6] Graefen B, Fazal N. Gpteacher: examining the efficacy of ChatGPT as a tool for public health education. Eur J Edu Stud. 2023 Jul 5.;10(8).

[7] Morita PP, Abhari S, Kaur J, Lotto M, Miranda PADSES, Oetomo A. Applying ChatGPT in public health: a SWOT and PESTLE analysis. Front Public Health. 2023 Jul 3.;11.

[8] Parray AA, Inam ZM, Ramonfaur D, Haider SS, Mistry SK, Pandya AK. ChatGPT and global public health: applications, challenges, ethical considerations and mitigation strategies. Glob Transit. 2023;5:50–4.

[9] Siad SM. The promise and perils of Google ' s bard for scientific research. HASTAC. 2023:0–3.

[10] Ali R, Tang OY, Connolly ID, Fridley JS, Shin JH, Zadnik Sullivan PL, et al. Performance of ChatGPT, GPT-4, and Google bard on a neurosurgery oral boards preparation question bank. Neurosurgery 93(5):p 1090-1098. November 2023..

[11] Fisher S, Rosella LC. Priorities for successful use of artificial intelligence by public health organizations: a literature review. BMC Public Health. 2022 Nov 22;22(1):2146.

[12] Pryss R, Kraft R, Baumeister H, Winkler J, Probst T, Reichert M, et al. Using Chatbots to support medical and psychological treatment procedures: challenges, opportunities, technologies, reference architecture. In. 2019:249–60.

[13] ReliefWeb. Guide: Mass casualty preparedness and response in emergency units - World | ReliefWeb [Internet]; 2022. [cited 2023 Jul 26]. Available from: https://reliefweb.int/report/world/guide-mass-casualty-preparedness-and-response-emergency-units.

[14] Jenkins JL, McCarthy ML, Sauer LM, Green GB, Stuart S, Thomas TL, et al. Mass-casualty triage: time for an evidence-based approach. Prehosp Disaster Med. 2008 Feb 28;23(1):3–8.
[15] Shin K, Lee T. Emergency medical service resource allocation in a mass casualty incident by integrating patient prioritization and hospital selection problems. IISE Trans. 2020 Oct 2;52(10):1141–55.
[16] Lin YK, Chen KC, Wang JH, Lai PF. Simple triage and rapid treatment protocol for emergency department mass casualty incident victim triage. Am J Emerg Med. 2022 Mar;53:99–103.
[17] Giannou C, Turnbull J. Mass casualty triage. In: Kravitz AS, van Tulleken A, editors. Oxford handbook of humanitarian medicine. Oxford University Press; 2019. p. 275–88.
[18] Hinson JS, Martinez DA, Schmitz PSK, Toerper M, Radu D, Scheulen J, et al. Accuracy of emergency department triage using the emergency severity index and independent predictors of under-triage and over-triage in Brazil: a retrospective cohort analysis. Int J Emerg Med. 2018 Dec 15;11(1):3.
[19] Sapp RF, Brice JH, Myers JB, Hinchey P. Triage performance of first-year medical students using a multiple-casualty scenario. Paper Exercise Prehosp Disaster Med. 2010 Jun 28;25(3):239–45.
[20] Natalie. ChatGPT — Release Notes | OpenAI Help Center [Internet]; 2023. [cited 2023 Jul 24]. Available from: https://help.openai.com/en/articles/6825453-chatgpt-release-notes.
[21] Krawczyk J. July Bard update: New features, languages, countries [Internet]; 2023. [cited 2023 Jul 24]. Available from: https://blog.google/products/bard/google-bard-new-features-update-july-2023/.
[22] Peng J, Xiang H. Trauma undertriage and overtriage rates: are we using the wrong formulas? Am J Emerg Med. 2016 Nov;34(11):2191–2.
[23] Purwadi H, Breaden K, McCloud C, Pranata S. The SALT and START Triage Systems for Classifying Patient Acuity Level: A Systematic Review. , Vol. 11Nurse Media Journal of Nursing. Diponegoro University- Department of Nursing, Faculty of Medicine; 2021; 413–27.
[24] Dave T, Athaluri SA, Singh S. ChatGPT in medicine: An overview of its applications, advantages, limitations, future prospects, and ethical considerations. Frontiers in Artificial Intelligence. Frontiers Media S.A; 2023.
[25] Prasad A, Jennifer BSSS, Ghsoh D, Busshetty H, J.T DrT. Chatbot in healthcare. Int J Eng Res Comp Sci Eng. 2022 Nov 1;9(11):26–8.
[26] Zagabathuni Yeshwanth. Applications, scope, and challenges for AI in healthcare. Int J Emerg Trends Eng Res. 2022 Apr 7;10(4):195–9.