



# Comparison of Triage Advice Provided by Artificial Intelligence-based Healthcare Application and Healthcare Providers

Triage Accuracy Study

# Introduction

Timely and accurate triage is essential to improve the efficiency of our national healthcare delivery systems. According to a recent data analysis of private insurance patients by UnitedHealth group, “two-thirds of emergency department visits are not an actual emergency and are avoidable. This results in a staggering amount of United States healthcare expenditure; out of 27 million annual ER visits 18 million ( $\frac{2}{3}$ ) are avoidable which results in costing \$32 billion”.<sup>6,7</sup>

Artificial Intelligence (AI) will be a transformational force in healthcare<sup>1</sup>. Popular digital health options like search engines, chatbots and artificial intelligence based personal health assistant apps offer patients a convenient and timely way to access relevant healthcare information. Since such a large percentage of patients turn to these options, their reliability & effectiveness becomes important.

Artificial Intelligence is increasingly being studied in the fields of Dermatology<sup>2</sup>, Ophthalmology<sup>3</sup>, and medical imaging<sup>4</sup>. However, there is a paucity of data on the role of AI triage in primary and pre-hospital care. A recent study revealed that online symptom checkers might provide unsuitable or incomplete diagnostic or triage advice for users in Australia, resulting in inappropriate care advice<sup>5</sup>. Improper triage may burden emergency departments, leading to their inappropriate use.

Misdiagnosis by physicians occurs in approximately 5% of outpatients<sup>8</sup>. Due to the prevalence of diagnosis error, decision support tools, if reliable, offer a way to help healthcare providers minimize this risk by standardizing the care. However, if symptom checkers or clinical decision support tools are poorly designed or lack rigorous clinical evaluation, it can put patients at risk and likely increase the overall health care costs<sup>9</sup>.

MayaMD is an artificial intelligence- based, advanced clinical decision support tool available for use by attending physicians, resident physicians, nurse practitioners, medical students, and the general public. In this study we seek to understand how MayaMD's (AI) patient triage and primary diagnosis compared to the Physicians and Physician Assistants performance.

## Methodology

The study was conducted in three phases.

Phase one of this triage study was conducted by the University of Utah Bioinformatics Department. They asked an Emergency Room (ER) physician, an ER physician assistant (with over ten years of combined experience), and an AI triaging tool, MayaMD, to make a triage recommendation to 50 clinical vignettes. Also, the primary diagnosis for each clinical vignette is recorded.

Four different triage options follow each clinical vignette.

A: Go to ER or call 911 – Life-threatening injuries or symptoms that need treatment immediately.

B: Go to Urgent Care w/in 24 hours – Non-life threatening, but need treatment the same day.

C: Go to Primary Care (PCP) within 3 days – Not immediately life-threatening that can wait three days before being seen by a primary care physician or specialist.

D: Self Care, remain at home, and only report to Primary Care or Urgent Care if condition worsens.

Phase two of this study was conducted with the core faculty from the University of Michigan / St. Joseph Mercy Hospital Emergency Medicine Residency Program. Six of their Emergency Room Physicians were included in the study. They were provided with 26 clinical vignettes from the initial phase one study which showed variability among the participants. Each physician provided their individual triage first, and then as a group they discussed their responses, and developed an agreed upon consensus answer for all of the vignettes.

Phase three of this study was conducted with the Internal Medicine faculty from the University of California Los Angeles and Case Western University. They were provided with the 50 clinical vignettes from phase one, and asked to provide their triage. Similar to phase two, each doctor on their own completed all 50 cases, and then afterwards; they discussed it as a group and created a consensus answer for each case.

The consensus triage from phase two and three were compared to the results from the phase one study.

## Results

In the phase one study, the overall percentages of triages were nearly matched between MayaMD and the ER Physician, although the individual vignettes were slightly different. Both triaged the patients to ER visits at 40% versus 42%, PCP at 42% versus 40%, Urgent care at 12%, and self-care at 6%. The ER Physician Assistant (ER PA) triaged patients to ER visits at 12%, Urgent care at 5%, PCP at 54% and self-care at 8%. MayaMD, the ER Physician and the ER PA, all agreed on the triage location in 24 cases (48%).

In the phase two study, the individual physician's triage scores varied from 19 to 24 out of 26 compared to their group's consensus answers. MayaMD's answers matched at 23 out of 26 (88.5%), ER Physicians on an average was 21.3 out of 26 (81.9%). Only one out of six physicians had a better match than MayaMD with their group's consensus responses. MayaMD had zero down triages i.e., patients were not inappropriately sent to lower levels of care. Using the phase two study's consensus in phase one, MayaMD matched at 94%, ER Physician at 80% and the ER PA was at 62% ( Figure 1)

In the phase three study, like the previous results, there was significant variability between individual providers. Compared to the consensus answers, accuracy level ranged from 66% to 94% among physicians and MayaMD was at 90%. After some corrections to answers where both answers could be appropriate, MayaMD was at 96% (Figure 3). MayaMD had 4% of down triages. (Two cases. In the first patient with nasal congestion and mild hearing loss, MayaMD suggested outpatient care, whereas the physician consensus was urgent care. In the second case, a patient with bacterial conjunctivitis, MayaMD, suggested urgent care, and the consensus answer was emergency room). Further review from the same physician team suggested that those answers are acceptable and can be considered as reasonable recommendations. This suggests that 100% of the triage recommendations from MayaMD can be considered safe in the phase three study.

In phase two, MayaMD's triage matched with 88.5% with the consensus response and better than average ER physician response of 81.9%. In phase three, MayaMD's triage matched 96% with the consensus responses compared to the average Internal Medicine Physician match of 82 %. Average physician accuracy rate compared to the consensus in all three studies were 80%, 81.9% and 82% respectively and MayaMD was at 92.3% (Figure 2).

# Discussion

Overall results showed that MayaMD had performed at par or better than individual clinicians in the three- phased triage study. Based on the phase 1 study, the ER PA had underestimated ER visit's need in 70% of cases (14/20 cases). However, it was just a single provider and cannot be generalized.

Interestingly, all three phases of the study showed significant variability between clinicians and this might be reflective of the cognitive errors that lead to medical mistakes. In phase two and phase three, the average physician score was around 80% and MayaMD was around 90%. However, in both the studies, there were physicians who had outperformed the AI tool. Percentage of unsafe advice (downtriage) provided by Maya was 0% in both phase two and phase three studies.

Research studies show that 43% of hospital visits originate from search engines. Searching for health related information online is actually the third most common online activity. Avoidable ER visits are a source of high cost on our healthcare system which with appropriate decision support can be minimized. The average cost of treating common primary care treatable conditions at a hospital ED is \$2,032; which is 12 times higher than visiting a physician office (\$167) and 10 times higher than traveling to an urgent care center (\$193) for help with those same issues. So, the results of this study are very significant.

These three studies were conducted at various teaching hospitals across the United States of America in the Emergency and Internal Medicine disciplines. Although the sample size is small, these initial studies hold a significant promise of the utility of MayaMD and the artificial intelligence-based tools in appropriately triaging the patients. Further large-scale studies are needed to replicate the findings in this study.

# Conclusions

The pilot studies conducted have shown that MayaMD (AI health assistant) can be a safe triage tool and clinical decision support tools can potentially reduce the unnecessary variability among clinicians.

# References

1. Ten Ways Artificial Intelligence will transform Primary Care. Steven Y Lin, Megan R Mahoney, Christine A Sinsky. J Gen Intern Med . 2019 Aug;34(8):1626-1630. doi: 10.1007/s11606-019-05035-1. Epub 2019 May 14.
2. Artificial Intelligence in Dermatology: A Primer. Albert Young, Mulin Xiong, Jacob Pfau, Micheal Keiser, Maria Wei. J Invest Dermatol 2020 Aug;140(8):1504-1512. doi: 10.1016/j.jid.2020.02.026. Epub 2020 Mar 27.
3. Artificial Intelligence in Ophthalmology during COVID-19 and in the post COVID-19 era. Joelle A Hallak, Angelica C Scanzera, Dimitri T Azar, R V Paul Chan. Curr Opin Ophthalmol 2020 Sep;31(5):447-453. doi: 10.1097/ICU.0000000000000685.
4. Artificial intelligence to diagnose ischemic stroke and identify large vessel occlusions: a systematic review. Nick M Murray, Mathias Unberath, Gregory Hager, Ferdinand K Hui. J Neurointerv Surg. 2020 Feb;12(2):156-164. doi: 10.1136/neurintsurg-2019-015135. Epub 2019 Oct 8.
5. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. Michella G Hill, Moira Sim, Brennen Mills. Medical journal of Australia, 2020-06, Vol.212 (11), p.514-519

6. [The High Cost of Avoidable Hospital Emergency Department Visits - UnitedHealth Group](#)

7. ['Avoidable' ER Visits Fuel U.S. Health Care Costs | Health News | US News](#)

8. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. Singh H, Meyer AN, Thomas EJ. *BMJ Qual Saf* 2014; 23: 727–31

9. Correspondence: Safety of patient-facing digital symptom checkers. Hamish Fraser, Enrico Coiera, David Wong. [Volume 392, Issue 10161](#), 24–30 November 2018, Pages 2263–2264. [https://doi.org/10.1016/S0140-6736\(18\)32819-8](https://doi.org/10.1016/S0140-6736(18)32819-8)

## Figure 1

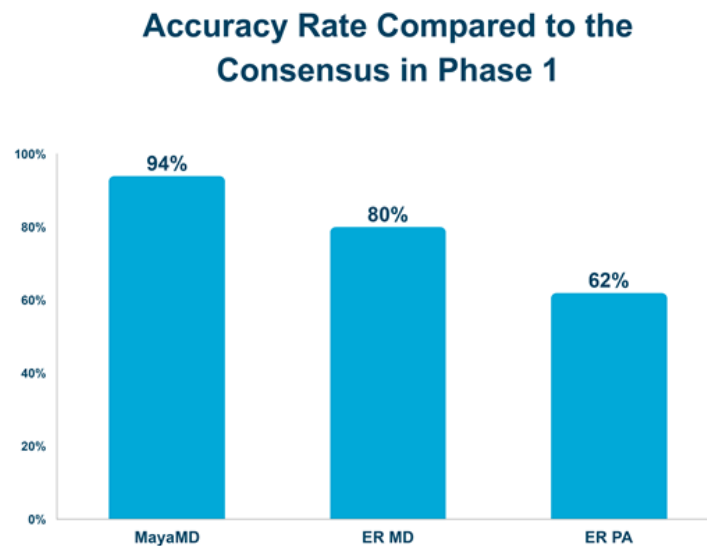




Figure 2

**Average Physician Accuracy Rate Compared to the Consensus in all Three Studies**

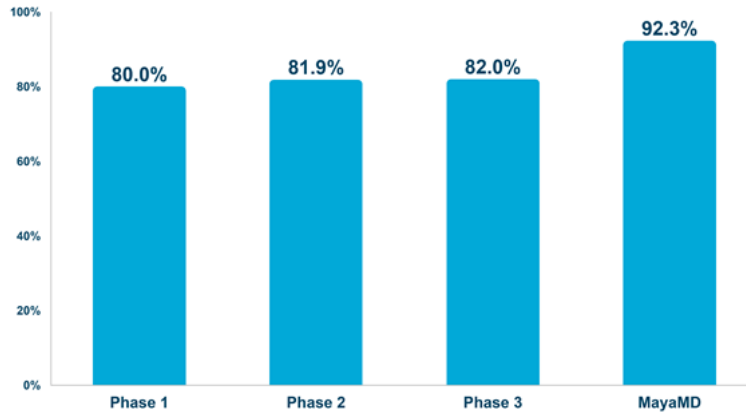


Figure 3

**Phase 3: Accuracy Rate Compared to the Group's Consensus**

