

Viewpoint

Evidence-Based Learning Strategies in Medicine Using AI

Juan Pablo Arango-Ibanez^{1*}, MD; Jose Alejandro Posso-Nuñez^{1*}, MD; Juan Pablo Díaz-Solórzano^{1*}, MD; Gustavo Cruz-Suárez^{2,3}, MD

¹Centro de Investigaciones Clínicas, Fundación Valle del Lili, Cali, Colombia

²Departamento de Anestesiología, Fundación Valle del Lili, Cali, Colombia

³Unidad de Inteligencia Artificial, Fundación Valle del Lili, Cali, Colombia

*these authors contributed equally

Corresponding Author:

Juan Pablo Arango-Ibanez, MD

Centro de Investigaciones Clínicas

Fundación Valle del Lili

Carrera 98 #18-49

Cali, 760031

Colombia

Phone: 57 23319090 ext 4022

Email: juan.arango.ib@fvl.org.co

Abstract

Large language models (LLMs), like ChatGPT, are transforming the landscape of medical education. They offer a vast range of applications, such as tutoring (personalized learning), patient simulation, generation of examination questions, and streamlined access to information. The rapid advancement of medical knowledge and the need for personalized learning underscore the relevance and timeliness of exploring innovative strategies for integrating artificial intelligence (AI) into medical education. In this paper, we propose coupling evidence-based learning strategies, such as active recall and memory cues, with AI to optimize learning. These strategies include the generation of tests, mnemonics, and visual cues.

JMIR Med Educ 2024;10:e54507; doi: [10.2196/54507](https://doi.org/10.2196/54507)

Keywords: artificial intelligence; large language models; ChatGPT; active recall; memory cues; LLMs; evidence-based; learning strategy; medicine; AI; medical education; knowledge; relevance

Introduction

e-Learning has revolutionized the way medicine is taught and learned through the use of different internet-based technologies that enhance education [1]. Among these technologies, artificial intelligence (AI) tools, especially large language models (LLMs), have notably garnered significant attention in recent years, given their promising implications for medical education. LLMs are algorithmic models that are trained by extensive data sets, and they have the capability to comprehend text and generate natural-language text in response to a given prompt (input). This allows for interactive engagement with these technologies in a conversational format akin to a “chat” [2,3]. One of the most known LLMs is ChatGPT (owned by OpenAI), and its latest version, ChatGPT-4, was recently released to the public.

Recent studies have demonstrated the great achievements of LLMs in relation to medical knowledge and reasoning, such as ChatGPT-4 scoring 90% when answering USMLE (United States Medical Licensing Examination)–

type questions [4], ChatGPT-4 passing a neurosurgery written board examination [5], and ChatGPT outperforming physicians in terms of providing empathic responses [6]. The educational potential of this technology is immense, encompassing a wide variety of applications. These include but are not limited to tutoring (personalized learning), patient simulation, generation of examination questions, and streamlined access to information [7-9]. The revolutionary potential of LLMs has resulted in researchers and medical students exploring the integration of AI into medical school curricula [10,11]. The rapid advancement of medical knowledge and the need for personalized learning underscore the relevance and timeliness of exploring innovative strategies for integrating AI into medical education [12].

Although numerous publications have examined the implications of LLMs for medicine and medical education, few have explored, in detail, specific strategies whereby LLMs can be used to optimize learning. In this paper, we propose strategies based on active recall, mnemonics, and the use of ChatGPT-4 [13] and DALL·E 3 (through ChatGPT-4)

for enhancing learning outcomes regarding factual knowledge and thus help fill this gap of information. These strategies include the creation of pretests and posttest quizzes, the development of mnemonics, and the use of visual cues and mnemonics. Pretests and posttests serve as effective tools for active recall—a proven method for improving memory retention. Mnemonics simplify complex information into more digestible and memorable formats. Visual cues provide a graphical representation of information, aiding in better understanding and recall.

Active Recall–Based Strategies

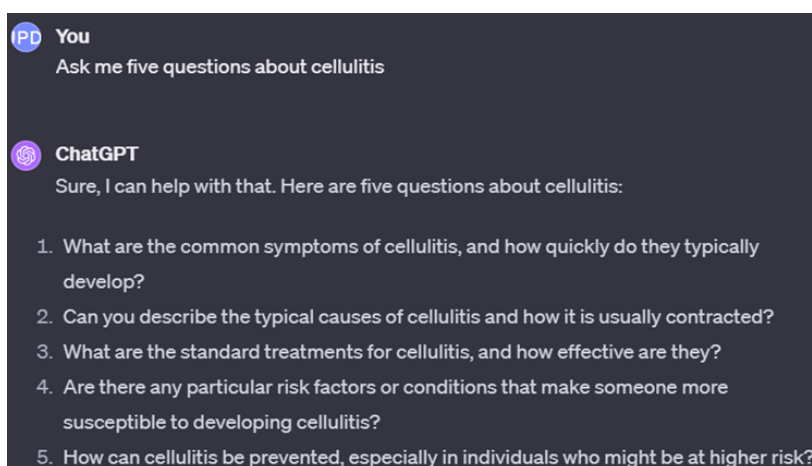
Medical school requires a significant amount of time spent on reading. Research indicates that medical students typically dedicate an average of 1.5 to 6 hours per day to reading [14,15]. Moreover, teacher-centered lectures, which predominantly focus on passive learning, persist as one of the most used strategies despite challenges in the medical education community with regard to encouraging integration with active learning methods that enhance the retention and application of knowledge [16–18]. This may be because medical school students might prefer classic didactic lectures over demonstrations, small group discussions, feedback activities, group work (generating test questions and coming up with solutions to a problem), and other active learning methods that have been reported to better enhance memory and retention [19].

It is essential to adopt evidence-based strategies to enhance learning efficiency, especially considering the substantial academic workload and the ongoing reliance on passive

learning methods. One such strategy is active recall, which involves actively retrieving information that was initially acquired passively through lectures, articles, or videos. This strategy is known to enhance learning significantly in comparison with passive learning strategies [20,21]. In this context, it is beneficial for readers to approach their reading proactively. This can be achieved by prefacing their reading with self-directed inquiries, understanding the main topics of the material, consistently formulating questions, and recognizing key concepts of high significance. Some related techniques include elaborative interrogation, which consists of answering “why” questions about a given concept to enhance medium- to long-term associative memory; self-explanation to relate new information to known information or explain steps taken for solving a problem to improve memory, comprehension, and transfer; and practice testing, which can improve memory, is not as time consuming, and can be applied at different times of the learning process [20].

One application of an active recall–based strategy involving the use of AI is illustrated in Figure 1, which shows ChatGPT being instructed to generate questions about cellulitis, as an example. Students are encouraged to attempt such questions before starting lectures or reading text. Answering questions before attending a lecture or reading a text (pretesting) is a strategy that enhances the learning process [22]. Interestingly, making mistakes during the study process can enhance learning by improving later memory; generating correct feedback; facilitating active learning; and stimulating the learner to redirect attention appropriately, especially when a mistake is followed by corrective feedback [23,24].

Figure 1. Example of the use of ChatGPT for pretesting.



Taking tests has been proven to enhance learning in various studies [25–28]. Repeated test-taking increases the transfer of learning [25] and improves long-term recall [28], and it even outperformed concept mapping for long-term retention in a previous study [26]. This strategy can be integrated with AI, as shown in Figure 2, which depicts our attempt to extract information from a *StatPearls* article on cellulitis [29] and request ChatGPT to generate relevant questions. The AI system can produce various question formats, such as multiple-choice, true-false, and fill-in-the-blank questions,

when given the appropriate prompts. These questions may be stored and reviewed days or weeks after the initial review to successfully apply spaced repetition, which has been demonstrated to improve learning and the consolidation of knowledge [21].

By using this method, one can input answers to questions and prompt ChatGPT to evaluate the answers' accuracy against the provided text. For instance, using questions from Figure 2, we tested ChatGPT's response by answering

a query about common causative bacteria of cellulitis. We intentionally incorporated broad, correct concepts (gram-positive bacteria) and specific yet erroneous details (emphasizing staphylococci, particularly *Staphylococcus aureus*, as primary

causatives instead of the correct streptococci) (Figure 3). ChatGPT feedback was tested again to contrast it with the feedback on a completely wrong answer (Figure 4).

Figure 2. Example of the use of ChatGPT for creating a posttest.

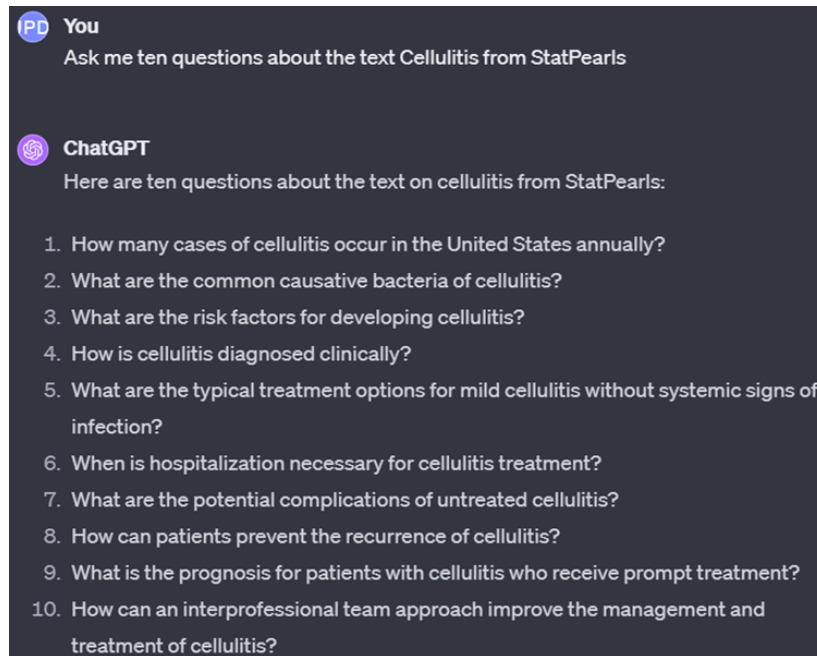


Figure 3. Feedback from ChatGPT on a partially correct answer to a question provided by ChatGPT.

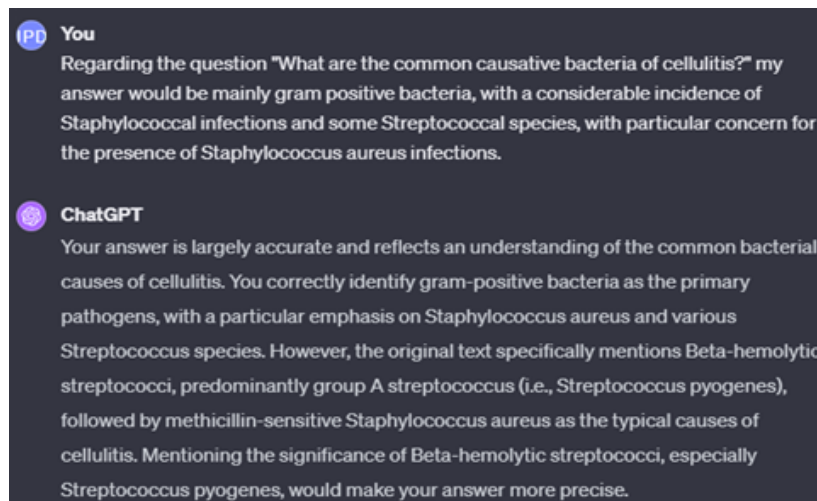
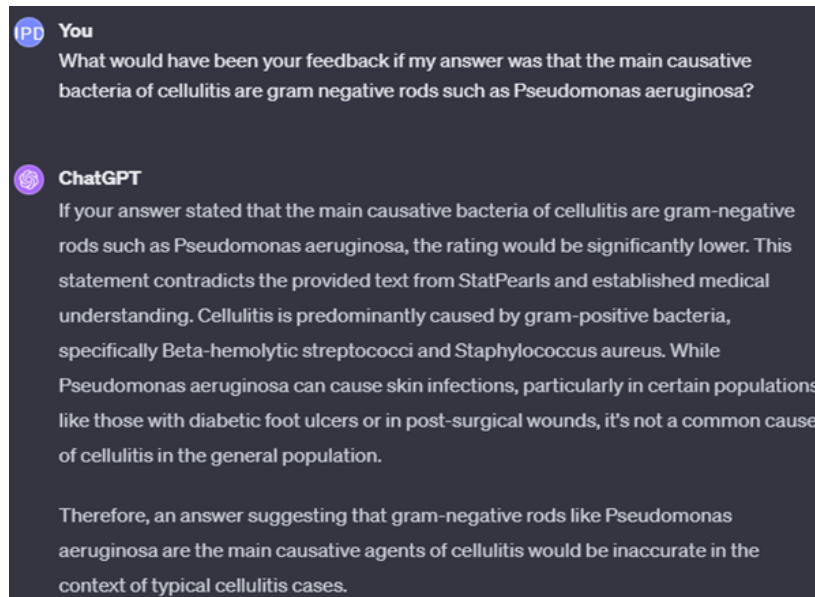


Figure 4. Feedback from ChatGPT on a wrong answer to a question provided by ChatGPT.

Memory Cues

Memory cues are learning strategies in which a process of metacognition transforms information in a way that makes the information easier to recall or understand. Cues can be self-generated or generated by external agents, other people, or AI. Evidence has long suggested that self-generated cues are superior to cues generated by other people [30,31]. Nonetheless, there is available evidence that indicates that memory cues generated by others can still enhance recall [32].

Memory cues are effective because they make difficult-to-remember information into something simpler or meaningful, which facilitates recall [32,33]. For example, a classic memory cue in medical school for remembering descriptors of pain is the use of the mnemonic “SOCRATES” (site, onset, character, radiation, associations, time course, exacerbating factors, and severity). In this context, the name of the great philosopher is repurposed to recall how to properly assess pain in a patient, with the name becoming an acronym. In a recent meta-analysis, a statistically significant effect was found for cueing decreasing the learners’ perceived cognitive load and promoting learning outcomes, namely retention, and the transfer of knowledge [34].

Other modalities of memory cues that are commonly used include pictures, short stories, songs, and rhymes [32]. Evidence indicates conflicting conclusions regarding the superiority of a specific modality of cues over another. For instance, in a study conducted by Pearson and Wilbiks [35], the authors attempted to evaluate the effect of the number of self-generated memory cues and aimed to test the findings of previous research that showed that the use of multisensory memory cues (ie, audiovisual cues) had a greater effect on recall than the use of one modality (ie, either visual cues [written words] or auditory cues [spoken words]). Their findings were that a greater number of cues led to higher recall, with statistical significance, but the modality of the cues did not have an effect on recall.

As previously indicated, one way to enhance the creation of an effective mnemonic is by using a common word as a cue to recall information [33]. This is one of the various ways that learners attempt to encode new vocabulary, abstract concepts, and master knowledge.

Figure 5 is an example of ChatGPT generating a mnemonic, using the word “brains” to recall the absolute contraindications of thrombolysis. Other examples are shown in Figure 6, in which ChatGPT creates a short story, and in Figure 7, in which ChatGPT creates a poem.

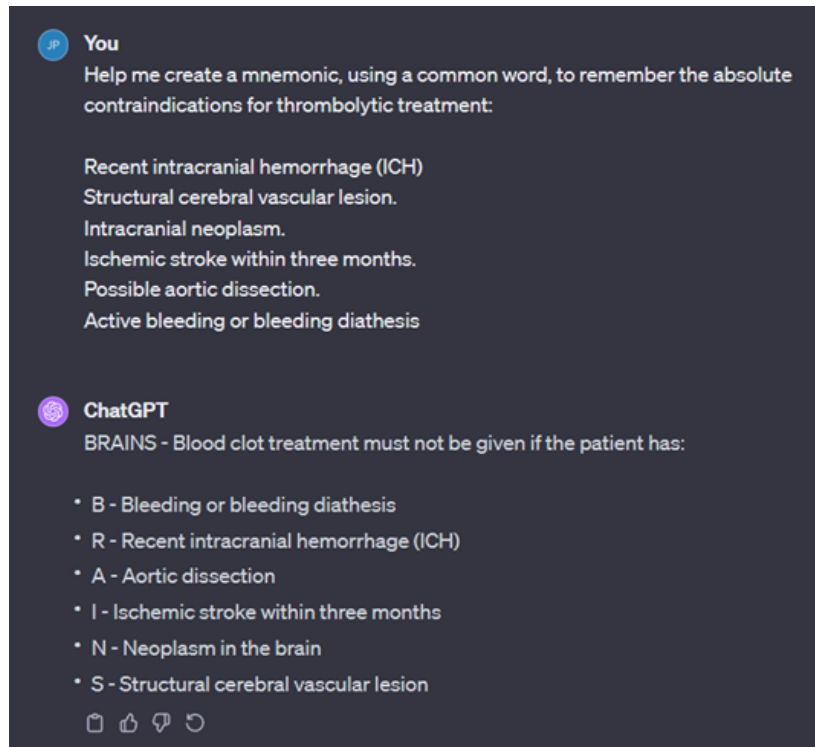
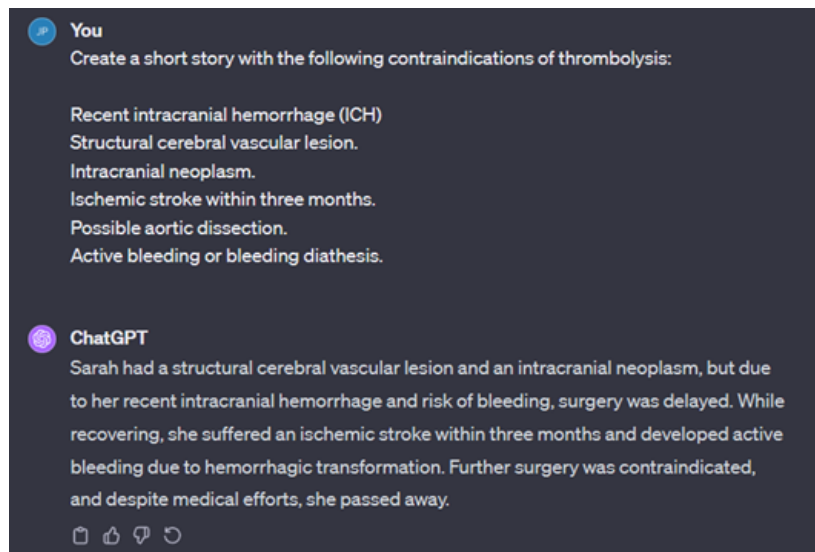
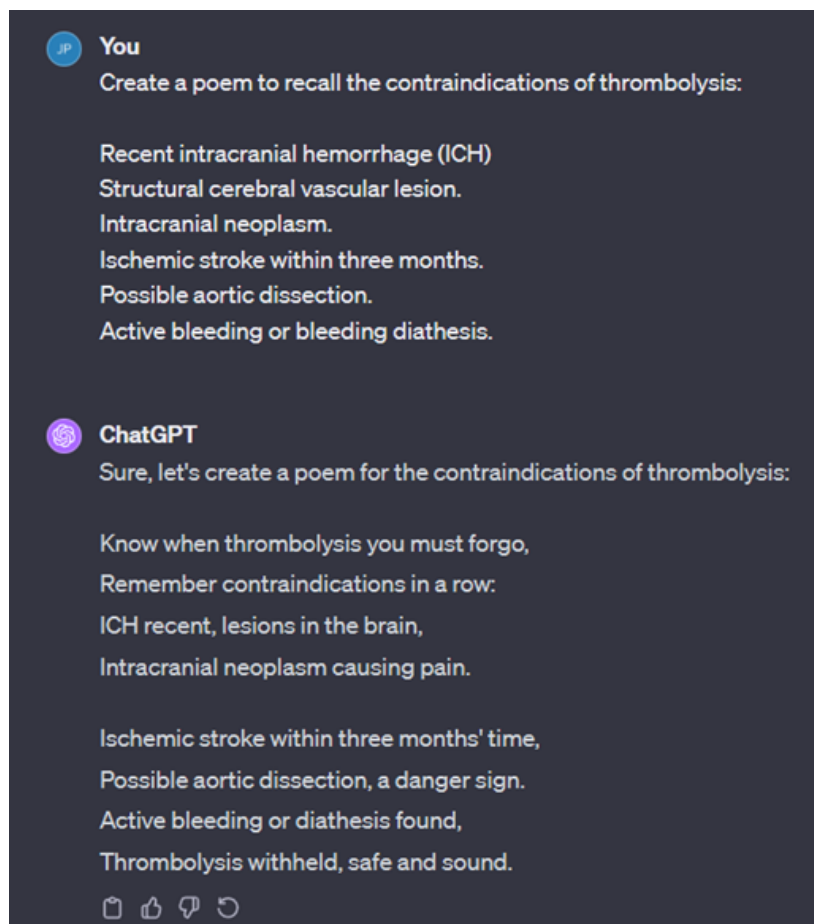
Figure 5. Acronym created by ChatGPT.**Figure 6.** Short story created by ChatGPT.

Figure 7. Poem created by ChatGPT.

Visual Mnemonics

A visual mnemonic or cue is a tool that uses visual imagery to improve the recall of information. This differs from verbal mnemonics, which use words, phrases, or songs, as visual mnemonics use pictorial cues to forge memorable links. Their effectiveness stems from the incorporation of visual representations, analogies, or symbolism, which fortifies the associations and makes them more distinct. Visual mnemonics aid in recalling abstract or intricate information and facilitate both the sequential and the immediate retrieval of memorized material [36,37]. The use of mnemonics can be highly useful for learning difficult or abstract information [30], which is often found in the field of medicine [38]. Multiple studies have demonstrated that using visual or pictorial mnemonics can enhance learning outcomes [35,39].

DALL·E 3 is an AI system created by OpenAI that generates images based on prompts provided by the user

and can be used for the creation of visual mnemonics. An example is given in Figure 8; a prompt was given to DALL·E 3 to create an image. For this example, which we created via DALL·E 3, the prompt “Fat purple man with long hair falling into a trap in a dry desert” was used to help recall some important features of hairy cell leukemia. “Fat” was used to recall the massive splenomegaly seen in patients with this condition; “purple” was used to make an association with lymphocytes, which are commonly seen as purple cells via hematoxylin and eosin staining and are involved in the pathogenesis of this neoplasm; “long hair” helps with recalling the filamentous projections of cells in hairy cell leukemia; “trap” was used to remember that this disease stains positively in tartrate-resistant acid phosphatase staining; and “dry desert” was used to recall that bone marrow fibrosis leads to dry tap on aspiration.

Figure 8. DALL-E 3 creation with the prompt “Fat purple man with long hair falling into a trap in a dry desert.”



Discussion

Active Recall

Active recall is a highly effective learning strategy and significantly outperforms passive restudying when it comes to certain learning outcomes, such as conceptualization and long-term retention [21]. It has yielded better evaluation testing performance than traditional studying or rereading. Kornell et al [22] reviewed recall when participants were presented with fictional and nonfictional information, modifying the time for pretesting and read-only strategies. The testing strategy yielded a greater amount of correct answers than the read-only strategy, with statistical significance when equal or more time was allocated to the testing condition when the final test was performed more than 24 hours after the learning exercise, as well as in the fictional topic scenarios ($P < .01$). Other studies supporting the use of pretesting have been reported [22,40,41]. This highlights the role of pretesting in learning new information.

The benefits of active learning through testing have also been supported by other authors. Butler [25] tested students' recall ability when they were either passively restudying or studying via repeated testing. Butler [25] found that repeated testing resulted in better performance on a recall test than passive learning strategies and concluded that repeated test-taking increases the transfer of learning. In another study performed by Karpicke and Blunt [26], when retrieval practice (testing) was evaluated against passive learning strategies and even concept mapping, it proved to be better for verbatim and inference question answering, resulting in

an improvement of about 50% in long-term retention scores ($d=1.50$; $F_{1,38}=21.63$; $\eta_p^2=0.36$). Additionally, the superiority of retesting over passive restudying for long-term retention has even been proven in a randomized controlled trial, wherein pediatric and emergency medicine residents were randomized to study the same text passages either via testing or repeated studying (ie, rereading). They were then tested on day 1, week 2, week 4, and month 6. The test results showed that the scores of participants who studied via testing were, on average, 13% higher than the scores of participants who performed repeated studying ($P < .001$), with an effect size of 0.91 [28]. Spaced testing (taking tests on different days between study sessions) has an even better effect on retention, long-term memory, and evaluation performance than repeated test-taking [27]. On the other hand, research indicates that spaced repetition (regardless of whether studying is done actively or passively) promotes more efficient and effective learning [42].

The previously mentioned studies highlight the importance of leveraging evidence-based techniques for studying rather than passive learning strategies. As we exemplified, these strategies can be coupled with AI. This approach addresses the limitation of relying solely on teacher-provided tests or textbook tests [20]. Moreover, ChatGPT is available on different platforms (web application and mobile app), and it can save chats (interactions) across these platforms. Therefore, students can easily access ChatGPT wherever it is needed and space their study sessions. Further, self-testing with AI reduces the pressure of graded assessments and leverages errors as learning opportunities [23,24], which research has shown to be particularly effective when the

learner is confident in their incorrect answers [23]. This could be related to the effect of feedback.

Several articles on active recall and learning emphasize the role of feedback in enhancing learning processes, which is a characteristic that passive studying lacks. Roediger and Butler [27] compared traditional studying with test-taking studying without feedback and test-taking studying with differently timed feedback to determine whether test-taking studying results in better retention and whether retention is enhanced by feedback. They tested participants at different times and highlighted the efficiency of test-taking as a studying strategy, which was superior to that of traditional reading and restudying (22%, 32%, and up to a 43% difference between traditional studying and test-taking studying without feedback, test-taking studying with immediate feedback, and test-taking studying with delayed feedback, respectively). Since feedback has a clear impact on learning, especially when coupled with active recall strategies, non-AI-mediated active learning strategies could be limited by the lack of opportunities to offer feedback. Feedback generally comes from a reliable source, such as a teacher or an expert, or is obtained through an appropriate literature search, which can be time consuming. Sometimes, it is not possible to have the timely intervention of a teacher or an expert if there is a lot to study, and it may not be possible for a student to conduct a proper literature search if the student is new to a given topic. Furthermore, there are different types of feedback; some feedback is self-directed (ie, obtained through an introspective process). Feedback in the learning process can be used to enhance or develop skills for setting goals, monitoring one's own learning process, and assimilating input (feedback) toward enhancing performance [43]. All of these are important skills to have when one attempts to obtain feedback on their own, such as when using LLMs for feedback.

AI can enhance medical education by offering feedback and explanations to clarify incorrect responses, thereby increasing study efficiency. By using AI tools like ChatGPT, students can receive detailed feedback on their answers, including the identification of errors and the provision of correct information, as we have shown. ChatGPT presents promising implications in providing technically accurate medical feedback, given the exceptional knowledge it has exhibited, as we previously described. This process, however, should not replace thorough literature research or foundational knowledge acquisition. AI models can also serve as tutors to facilitate discussions on specific knowledge areas, similar to existing models in other fields, such as Khan Academy's Khanmigo, which serves as a fully personalized tutor [44]. One limitation of AI systems like ChatGPT is their character limit for inputs, which can be managed by breaking text into sections or using multiple prompts. Additionally, web searching is only available with ChatGPT's paid subscription; for the free version, one should provide ChatGPT with the reference text by copying and pasting it. Further research is needed to explore the potential of AI-assisted tutors in medical education, especially in education on basic subjects.

Memory Cues

Memory cues can be used as effective learning tools that ease the studying experience for a student attempting not only memorization but also mastery of complex concepts and new vocabulary. Evidence has described the superiority of self-generated memory cues over cues created by others [30,31]. The explanations behind the superiority of self-generated memory cues are (1) the generation effect of creating such cues, that is, the act of generation requires significant cognitive effort, which boosts memory, and (2) the cue selection process itself and the consequential metamnemonic effect. This means that students who identify the learning formats that work best for them are able to create their own memory cues by using a modality that is tailored to their needs [31]. For example, if a student prefers to have a visual representation of the ideas that they are attempting to memorize, they might lean toward the creation of mnemonics that create a mental picture to integrate information. Tullis and Fraundorf [31] proposed evidence that the benefits of self-generated cues come in great part from the correct selection of a cue from a list of candidates. If students can create multiple cues, they can, with greater effectiveness, select the cue that best benefits retrieval. Tullis and Fraundorf [31] further suggested that allowing a learner to select from multiple options of cues requires less cognitive work, takes less time, and may not hinder memorization.

As we described, the act of generation is effortful and may be time consuming. AI tools like ChatGPT can help students as a result of their seemingly tireless and effortless generative capacity. In addition, these tools can create multiple cues with different modalities (textual-based cues and pictorial cues) when prompted to do so, thereby allowing learners to focus on understanding the material and selecting the most appropriate cue that fits their educational needs. The downside to the use of this method is that multiple attempts may be required for ChatGPT to produce a mnemonic that is subjectively good or fitting enough for a particular student. In addition, it is our opinion that these tools are best used when the user has an idea of what they should learn or memorize, and the user should prompt the AI tool to create a mnemonic device that facilitates the recall of the information they wish to encode. This is because there is abundant evidence of ChatGPT not only making errors but also blatantly providing false information [45], which is known as "artificial hallucination."

Visual Mnemonics

Previous research has explored the effectiveness of visual mnemonics in improving learning outcomes. An experimental study that compared pictorial mnemonic use to traditional study methods found that pictorial mnemonics aid in learning from text passages by improving the recall of factual knowledge and long-term memory retention in college students [46]. Additionally, a randomized trial compared audiovisual mnemonics against traditional text-based learning for retaining medical knowledge; participants who used mnemonics demonstrated significant improvements in free-recall tests, with scores improving by

65%, 161%, and 208% immediately, after 1 week, and after 1 month, respectively, when compared to those who used text materials ($P<.001$). Moreover, the group that used mnemonics outperformed the group that used text materials by 55% in a 1-week–delayed multiple-choice test that focused on higher-order thinking ($P<.001$) [47]. In a comparative study of visual mnemonics versus traditional lectures for learning the porphyrin pathway, there was no significant difference in quiz scores immediately or 1 week after the intervention; however, the mnemonic group exhibited a 20% higher score 3 weeks later ($P=.02$) [48]. In another randomized trial that compared story-based audiovisual mnemonics with traditional text reading for memory retention among medical students, the audiovisual mnemonics group demonstrated significantly better performance in multiple-choice tests immediately after the intervention ($P=.04$), as well as at 1 week, 2 weeks, and 4 weeks after the intervention [49]. These results underscore story-based mnemonics' superior effectiveness in enhancing immediate and long-term memory retention in medical education. Although there is some variation in the visual mnemonic techniques across studies (eg, the studies by Yang et al [47] and Abdalla et al [49] used some audiovisual mnemonics), they consistently demonstrated that factual knowledge can be represented visually and that the use of this type of mnemonic enhances both the recall and long-term retention of knowledge, with large effect sizes.

The visual mnemonic proposed in our study highly resembles the strategy used in the experiment by Rummel et al [46], in which visual mnemonics were created from texts about psychologists, incorporating elements for recalling both the psychologists' names and the key aspects of their theories. In our mnemonic, "long hair" aids in recalling the name of the disease (hairy cell leukemia), and the other elements in the image are used to help recall the disease's main features. The Picmonic System, which uses mnemonics from a web-based educational platform [50] that was used in the studies by Yang et al [47] and Abdalla et al [49], also adopts the visual mnemonic approach by combining visual elements and storytelling to enhance the recall of information; this is also highly similar to our approach. Thus, using DALL-E 3 for mnemonic generation shows promise for improving different learning outcomes,

such as test performance, long-term retention, and free recall. Future studies should experimentally investigate the effectiveness of visual mnemonics generated by text-to-image models in learning processes. A significant limitation of using DALL-E 3 for medical mnemonic generation is its restriction on explicit content, prohibiting prompts with terms like "blood." By recognizing this limitation, knowledge area-specific text-to-image models can be developed to more accurately describe the information needed and enable the use of words that are commonly used in a knowledge area but are censored in current models. Another limitation is that creating stories that accurately reflect the intended factual knowledge for mnemonic cues can be complex, particularly for certain subjects. Effective prompt engineering techniques could help in creating more relevant and coherent visual mnemonics.

Conclusions

LLMs, as a form of AI, are transforming the landscape of medical education. They offer a vast range of applications, and their potential has sparked discussions about integrating them into medical school curricula. Active recall-based learning strategies can be integrated with AI and can promisingly improve the recall and retention of information. This integration can be effectively applied by using AI to generate pretests and posttest quizzes. Memory cues, including self-generated mnemonics and mnemonics created by AI, can effectively simplify and transform complex information, thereby enhancing recall and optimizing learning. ChatGPT can create multiple types of memory cues, such as acronyms, short stories, and even poems. Moreover, AI tools, like DALL-E 3, can create images based on text and thus can be used to create visual mnemonics. However, crafting the right prompts can be challenging and time consuming, and results may vary. Thus, we believe that the use of new AI-based technologies, such as ChatGPT and DALL-E 3, is a highly useful strategy for learning, especially when these technologies are used with evidence-based principles. Further research is warranted to elucidate the impact of these strategies within the context of medical education.

Acknowledgments

JPAI designed and conceptualized the work. JPAI, JAPN, JPDS, and GCS gathered the data and performed the interpretation. JPAI, JAPN, and JPDS drafted the manuscript. JPAI, JAPN, JPDS, and GCS reviewed and corrected the final version of the manuscript. All authors have read and approved the final version of the paper.

Conflicts of Interest

None declared.

References

1. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. *Acad Med*. Mar 2006;81(3):207-212. [doi: [10.1097/00001888-200603000-00002](https://doi.org/10.1097/00001888-200603000-00002)] [Medline: [16501260](https://pubmed.ncbi.nlm.nih.gov/16501260/)]
2. Naveed H, Khan AU, Qiu S, et al. A comprehensive overview of large language models. arXiv. Preprint posted online on Oct 5, 2023. [doi: [10.48550/arXiv.2307.06435](https://doi.org/10.48550/arXiv.2307.06435)]
3. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. Aug 2023;29(8):1930-1940. [doi: [10.1038/s41591-023-02448-8](https://doi.org/10.1038/s41591-023-02448-8)] [Medline: [37460753](https://pubmed.ncbi.nlm.nih.gov/37460753/)]

4. Brin D, Sorin V, Vaid A, et al. Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments. *Sci Rep*. Oct 1, 2023;13(1):16492. [doi: [10.1038/s41598-023-43436-9](https://doi.org/10.1038/s41598-023-43436-9)] [Medline: [37779171](https://pubmed.ncbi.nlm.nih.gov/37779171/)]
5. Ali R, Tang OY, Connolly ID, et al. Performance of ChatGPT and GPT-4 on neurosurgery written board examinations. *Neurosurgery*. Dec 1, 2023;93(6):1353-1365. [doi: [10.1227/neu.0000000000002632](https://doi.org/10.1227/neu.0000000000002632)] [Medline: [37581444](https://pubmed.ncbi.nlm.nih.gov/37581444/)]
6. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. Jun 1, 2023;183(6):589-596. [doi: [10.1001/jamainternmed.2023.1838](https://doi.org/10.1001/jamainternmed.2023.1838)] [Medline: [37115527](https://pubmed.ncbi.nlm.nih.gov/37115527/)]
7. Tsang R. Practical applications of ChatGPT in undergraduate medical education. *J Med Educ Curric Dev*. May 24, 2023;10:23821205231178449. [doi: [10.1177/23821205231178449](https://doi.org/10.1177/23821205231178449)] [Medline: [37255525](https://pubmed.ncbi.nlm.nih.gov/37255525/)]
8. 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*. Mar 6, 2023;9:e46885. [doi: [10.2196/46885](https://doi.org/10.2196/46885)] [Medline: [36863937](https://pubmed.ncbi.nlm.nih.gov/36863937/)]
9. Mohammad B, Supti T, Alzubaidi M, et al. The pros and cons of using ChatGPT in medical education: a scoping review. *Stud Health Technol Inform*. Jun 29, 2023;305:644-647. [doi: [10.3233/SHTI230580](https://doi.org/10.3233/SHTI230580)] [Medline: [37387114](https://pubmed.ncbi.nlm.nih.gov/37387114/)]
10. Cooper A, Rodman A. AI and medical education - a 21st-century Pandora's box. *N Engl J Med*. Aug 3, 2023;389(5):385-387. [doi: [10.1056/NEJMp2304993](https://doi.org/10.1056/NEJMp2304993)] [Medline: [37522417](https://pubmed.ncbi.nlm.nih.gov/37522417/)]
11. Civaner MM, Uncu Y, Bulut F, Chalil EG, Tatli A. Artificial intelligence in medical education: a cross-sectional needs assessment. *BMC Med Educ*. Nov 9, 2022;22(1):772. [doi: [10.1186/s12909-022-03852-3](https://doi.org/10.1186/s12909-022-03852-3)] [Medline: [36352431](https://pubmed.ncbi.nlm.nih.gov/36352431/)]
12. Sapci AH, Sapci HA. Artificial intelligence education and tools for medical and health informatics students: systematic review. *JMIR Med Educ*. Jun 30, 2020;6(1):e19285. [doi: [10.2196/19285](https://doi.org/10.2196/19285)] [Medline: [32602844](https://pubmed.ncbi.nlm.nih.gov/32602844/)]
13. ChatGPT. URL: <https://chat.openai.com> [Accessed 2024-01-14]
14. Leff B, Harper GM. The reading habits of medicine clerks at one medical school: frequency, usefulness, and difficulties. *Acad Med*. May 2006;81(5):489-494. [doi: [10.1097/01.ACM.0000222273.90705.a6](https://doi.org/10.1097/01.ACM.0000222273.90705.a6)] [Medline: [16639211](https://pubmed.ncbi.nlm.nih.gov/16639211/)]
15. Klatt EC, Klatt CA. How much is too much reading for medical students? assigned reading and reading rates at one medical school. *Acad Med*. Sep 2011;86(9):1079-1083. [doi: [10.1097/ACM.0b013e31822579fc](https://doi.org/10.1097/ACM.0b013e31822579fc)] [Medline: [21785317](https://pubmed.ncbi.nlm.nih.gov/21785317/)]
16. Prober CG, Heath C. Lecture halls without lectures--a proposal for medical education. *N Engl J Med*. May 3, 2012;366(18):1657-1659. [doi: [10.1056/NEJMp1202451](https://doi.org/10.1056/NEJMp1202451)] [Medline: [22551125](https://pubmed.ncbi.nlm.nih.gov/22551125/)]
17. Prober CG, Khan S. Medical education reimaged: a call to action. *Acad Med*. Oct 2013;88(10):1407-1410. [doi: [10.1097/ACM.0b013e3182a368bd](https://doi.org/10.1097/ACM.0b013e3182a368bd)] [Medline: [23969367](https://pubmed.ncbi.nlm.nih.gov/23969367/)]
18. Bucklin BA, Asdigian NL, Hawkins JL, Klein U. Making it stick: use of active learning strategies in continuing medical education. *BMC Med Educ*. Jan 11, 2021;21(1):44. [doi: [10.1186/s12909-020-02447-0](https://doi.org/10.1186/s12909-020-02447-0)] [Medline: [33430843](https://pubmed.ncbi.nlm.nih.gov/33430843/)]
19. Roffler M, Sheehy R. Self-reported learning and study strategies in first and second year medical students. *Med Sci Educ*. Mar 18, 2022;32(2):329-335. [doi: [10.1007/s40670-022-01533-w](https://doi.org/10.1007/s40670-022-01533-w)] [Medline: [35528305](https://pubmed.ncbi.nlm.nih.gov/35528305/)]
20. Dunlosky J, Rawson KA, Marsh EJ, Nathan MJ, Willingham DT. Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol Sci Public Interest*. Jan 2013;14(1):4-58. [doi: [10.1177/1529100612453266](https://doi.org/10.1177/1529100612453266)] [Medline: [26173288](https://pubmed.ncbi.nlm.nih.gov/26173288/)]
21. Augustin M. How to learn effectively in medical school: test yourself, learn actively, and repeat in intervals. *Yale J Biol Med*. Jun 6, 2014;87(2):207-212. [Medline: [24910566](https://pubmed.ncbi.nlm.nih.gov/24910566/)]
22. Kornell N, Hays MJ, Bjork RA. Unsuccessful retrieval attempts enhance subsequent learning. *J Exp Psychol Learn Mem Cogn*. Jul 2009;35(4):989-998. [doi: [10.1037/a0015729](https://doi.org/10.1037/a0015729)] [Medline: [19586265](https://pubmed.ncbi.nlm.nih.gov/19586265/)]
23. Metcalfe J. Learning from errors. *Annu Rev Psychol*. Jan 3, 2017;68:465-489. [doi: [10.1146/annurev-psych-010416-044022](https://doi.org/10.1146/annurev-psych-010416-044022)] [Medline: [27648988](https://pubmed.ncbi.nlm.nih.gov/27648988/)]
24. Mera Y, Rodríguez G, Marin-García E. Unraveling the benefits of experiencing errors during learning: definition, modulating factors, and explanatory theories. *Psychon Bull Rev*. Jun 2022;29(3):753-765. [doi: [10.3758/s13423-021-02022-8](https://doi.org/10.3758/s13423-021-02022-8)] [Medline: [34820785](https://pubmed.ncbi.nlm.nih.gov/34820785/)]
25. Butler AC. Repeated testing produces superior transfer of learning relative to repeated studying. *J Exp Psychol Learn Mem Cogn*. Sep 2010;36(5):1118-1133. [doi: [10.1037/a0019902](https://doi.org/10.1037/a0019902)] [Medline: [20804289](https://pubmed.ncbi.nlm.nih.gov/20804289/)]
26. Karpicke JD, Blunt JR. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*. Feb 11, 2011;331(6018):772-775. [doi: [10.1126/science.1199327](https://doi.org/10.1126/science.1199327)] [Medline: [21252317](https://pubmed.ncbi.nlm.nih.gov/21252317/)]
27. Roediger HL 3rd, Butler AC. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci*. Jan 2011;15(1):20-27. [doi: [10.1016/j.tics.2010.09.003](https://doi.org/10.1016/j.tics.2010.09.003)] [Medline: [20951630](https://pubmed.ncbi.nlm.nih.gov/20951630/)]
28. Larsen DP, Butler AC, Roediger HL 3rd. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ*. Dec 2009;43(12):1174-1181. [doi: [10.1111/j.1365-2923.2009.03518.x](https://doi.org/10.1111/j.1365-2923.2009.03518.x)] [Medline: [19930508](https://pubmed.ncbi.nlm.nih.gov/19930508/)]

29. Brown BD, Watson KLH. Cellulitis. In: StatPearls. StatPearls Publishing; 2023. URL: <https://www.ncbi.nlm.nih.gov/books/NBK549770/> [Accessed 2024-01-15]
30. Tullis JG, Qiu J. Generating mnemonics boosts recall of chemistry information. *J Exp Psychol Appl*. Mar 2022;28(1):71-84. [doi: [10.1037/xap0000350](https://doi.org/10.1037/xap0000350)] [Medline: [33939460](https://pubmed.ncbi.nlm.nih.gov/33939460/)]
31. Tullis JG, Fraundorf SH. Selecting effectively contributes to the mnemonic benefits of self-generated cues. *Mem Cognit*. May 2022;50(4):765-781. [doi: [10.3758/s13421-021-01245-3](https://doi.org/10.3758/s13421-021-01245-3)] [Medline: [34731430](https://pubmed.ncbi.nlm.nih.gov/34731430/)]
32. Tullis JG, Finley JR. Self-generated memory cues: effective tools for learning, training, and remembering. *Policy Insights Behav Brain Sci*. Oct 2018;5(2):179-186. [doi: [10.1177/2372732218788092](https://doi.org/10.1177/2372732218788092)]
33. Tullis JG, Finley JR. What characteristics make self-generated memory cues effective over time? *Memory*. Nov 2021;29(10):1308-1319. [doi: [10.1080/09658211.2021.1979585](https://doi.org/10.1080/09658211.2021.1979585)] [Medline: [34546833](https://pubmed.ncbi.nlm.nih.gov/34546833/)]
34. Xie H, Wang F, Hao Y, et al. The more total cognitive load is reduced by cues, the better retention and transfer of multimedia learning: a meta-analysis and two meta-regression analyses. *PLoS One*. Aug 30, 2017;12(8):e0183884. [doi: [10.1371/journal.pone.0183884](https://doi.org/10.1371/journal.pone.0183884)] [Medline: [28854205](https://pubmed.ncbi.nlm.nih.gov/28854205/)]
35. Pearson HC, Wilbiks JMP. Effects of audiovisual memory cues on working memory recall. *Vision (Basel)*. Mar 19, 2021;5(1):14. [doi: [10.3390/vision5010014](https://doi.org/10.3390/vision5010014)] [Medline: [33808715](https://pubmed.ncbi.nlm.nih.gov/33808715/)]
36. Higbee KL. Mnemonics, psychology of. In: *International Encyclopedia of the Social & Behavioral Sciences*. Elsevier; 2001:9915-9918. [doi: [10.1016/B0-08-043076-7/01517-5](https://doi.org/10.1016/B0-08-043076-7/01517-5)]
37. O'Hanlon R, Laynor G. Responding to a new generation of proprietary study resources in medical education. *J Med Libr Assoc*. Apr 2019;107(2):251-257. [doi: [10.5195/jmla.2019.619](https://doi.org/10.5195/jmla.2019.619)] [Medline: [31019395](https://pubmed.ncbi.nlm.nih.gov/31019395/)]
38. Qiao YQ, Shen J, Liang X, et al. Using cognitive theory to facilitate medical education. *BMC Med Educ*. Apr 14, 2014;14:79. [doi: [10.1186/1472-6920-14-79](https://doi.org/10.1186/1472-6920-14-79)] [Medline: [24731433](https://pubmed.ncbi.nlm.nih.gov/24731433/)]
39. Reed LA, Hoffman LG. Pictorial cues and enhancement of patient recall of instructions or information. *J Am Optom Assoc*. Apr 1986;57(4):312-315. [Medline: [3700955](https://pubmed.ncbi.nlm.nih.gov/3700955/)]
40. Richland LE, Kornell N, Kao LS. The pretesting effect: do unsuccessful retrieval attempts enhance learning? *J Exp Psychol Appl*. Sep 2009;15(3):243-257. [doi: [10.1037/a0016496](https://doi.org/10.1037/a0016496)] [Medline: [19751074](https://pubmed.ncbi.nlm.nih.gov/19751074/)]
41. Latimier A, Riegert A, Peyre H, Ly ST, Casati R, Ramus F. Does pre-testing promote better retention than post-testing? *NPJ Sci Learn*. Sep 24, 2019;4:15. [doi: [10.1038/s41539-019-0053-1](https://doi.org/10.1038/s41539-019-0053-1)] [Medline: [31583117](https://pubmed.ncbi.nlm.nih.gov/31583117/)]
42. Kang SHK. Spaced repetition promotes efficient and effective learning: policy implications for instruction. *Policy Insights Behav Brain Sci*. Mar 2016;3(1):12-19. [doi: [10.1177/2372732215624708](https://doi.org/10.1177/2372732215624708)]
43. Lüdeke ABEK, Olaya JFG. Effective feedback, an essential component of all stages in medical education. *Universitas Medica*. 2020;61(3). [doi: [10.11144/Javeriana.umed61-3.feed](https://doi.org/10.11144/Javeriana.umed61-3.feed)]
44. Meet Khanmigo: Khan Academy's AI-powered teaching assistant & tutor. Khan Academy. URL: <https://www.khanacademy.org/khan-labs> [Accessed 2024-01-19]
45. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*. Feb 19, 2023;15(2):e35179. [doi: [10.7759/cureus.35179](https://doi.org/10.7759/cureus.35179)] [Medline: [36811129](https://pubmed.ncbi.nlm.nih.gov/36811129/)]
46. Rummel N, Levin JR, Woodward MM. Do pictorial mnemonic text-learning aids give students something worth writing about? *J Educ Psychol*. 2003;95(2):327-334. [doi: [10.1037/0022-0663.95.2.327](https://doi.org/10.1037/0022-0663.95.2.327)]
47. Yang A, Goel H, Bryan M, et al. The Picmonic(®) Learning System: enhancing memory retention of medical sciences, using an audiovisual mnemonic web-based learning platform. *Adv Med Educ Pract*. May 8, 2014;5:125-132. [doi: [10.2147/AMEP.S61875](https://doi.org/10.2147/AMEP.S61875)] [Medline: [24868180](https://pubmed.ncbi.nlm.nih.gov/24868180/)]
48. De Moll EH, Routt E, Heinecke G, Tsui C, Levitt J. The use of an imagery mnemonic to teach the porphyrin biochemical pathway. *Dermatol Online J*. Apr 16, 2015;21(4):13030/qt4j51j8wt. [Medline: [25933070](https://pubmed.ncbi.nlm.nih.gov/25933070/)]
49. Abdalla MMI, Azzani M, Rajendren R, et al. Effect of story-based audiovisual mnemonics in comparison with text-reading method on memory consolidation among medical students: a randomized controlled trial. *Am J Med Sci*. Dec 2021;362(6):612-618. [doi: [10.1016/j.amjms.2021.07.015](https://doi.org/10.1016/j.amjms.2021.07.015)] [Medline: [34606752](https://pubmed.ncbi.nlm.nih.gov/34606752/)]
50. Picmonic® picture mnemonics - medical school, nursing school and more! Picmonic. URL: <https://www.picmonic.com/> [Accessed 2024-01-03]

Abbreviations

AI: artificial intelligence

LLM: large language model

SOCRATES: site, onset, character, radiation, associations, time course, exacerbating factors, and severity

USMLE: United States Medical Licensing Examination

Edited by Gunther Eysenbach; peer-reviewed by Lorentz Jantschi, Shankar Ganesh; submitted 12.11.2023; final revised version received 20.01.2024; accepted 23.03.2024; published 24.05.2024

Please cite as:

Arango-Ibanez JP, Posso-Nuñez JA, Díaz-Solórzano JP, Cruz-Suárez G

Evidence-Based Learning Strategies in Medicine Using AI

JMIR Med Educ 2024;10:e54507

URL: <https://mededu.jmir.org/2024/1/e54507>

doi: [10.2196/54507](https://doi.org/10.2196/54507)

© Juan Pablo Arango-Ibanez, Jose Alejandro Posso-Nuñez, Juan Pablo Díaz-Solórzano, Gustavo Cruz-Suárez. Originally published in JMIR Medical Education (<https://mededu.jmir.org>), 24.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Education, is properly cited. The complete bibliographic information, a link to the original publication on <https://mededu.jmir.org/>, as well as this copyright and license information must be included.