Exploring the role of an artificial intelligence chatbot on 38 appendicitis management: An experimental study on ChatGPT

Authors: Dylan Gracias, Adrian Siu, Ishith Seth, Dil Dooreemeah, Angus Lee

Introduction

- Appendicitis is a common surgical condition that urgent medical attention
- ChatGPT is a deep learning language model that widespread recognition for its versatility and pote applications in clinical medicine
- As AI and machine learning continue to progress. systems could significantly enhance the clinician' patient's access to reliable, high quality, prompt r information

Aims

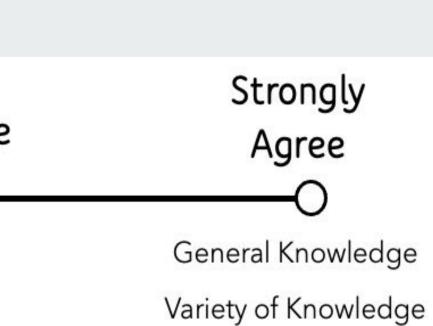
- To assess the precision and comprehensiveness generated by ChatGPT regarding appendicitis ma compared to clinical guidelines and evidence-bas literature
- Assess the quality of references provided by Cha

Methods

- ChatGPT was presented with six questions related appendicitis
- Responses and references were analysed for according comprehensiveness, coherence, appropriateness quality against evidence-based literature
- Further evaluation was performed using a five point Likert scale

Strongly Disagree	Disagree	Neither Agree or Disagree O	Agree O
	Reliability	Accuracy	
	Reference Quality	Comprehensiveness	
		Wholistic	
		In Depth	

Question	Key Positive Points	Key Negative Points	References
1. What is the cause of appendicitis?	 Prefaced that the exact mechanism is not fully understood Identified obstruction as a theory, including fecaliths, lymphoid hyperplasia and tumours contributing to this 	 No elaboration on obstruction leading to inflammation Did not identify any other pathophysiological theories Did not identify environmental, infectious, or ethnic factors which may contribute Repetitive rather than diverse and specific 	 2 existent, relevant publications 3 existent, less relevant publications which deviated from the question
2. What are the preventative measures for appendicitis?	 Included hygiene precautions 	 Failed to acknowledge the limited evidence on the prevention of appendicitis Hygiene precautions was non-descriptive Mentioned techniques found throughout the internet but not validated by evidence-based research (eg: exercise, stress reduction) 	 4 non-existent publications 1 existent, but irrelevant publication
3. What is the most appropriate management for appendicitis?	 Included appendicectomy and antibiotic treatment with a description of each 	 Provided limited information on the specific indications or considerations for each treatment option No data on success rates, complication rates, recurrence rates, negative appendicectomy rates, length of hospital stay and economic implications to compare these two treatment options Lacked depth 	 2 existent, high quality, relevant studies 3 existent, but irrelevant studies Missed key randomised trial, systematic review, and meta-analysis studies
4. Compare the efficacy between surgical versus antibiotics for appendicitis management and what is superior?	 More comprehensive Specified the severity of the disease contributing to the decision, amongst other factors Discussed recurrence rate associated with antibiotic therapy Downsides of antibiotic therapy were explored 	 No quantitative data on success or recurrence rates Downsides of surgical management not explored 	 5 existent, high quality, relevant publications
5. Provide future recommendations for advancement in management of appendicitis and what further research is needed?	 Suggested further evaluation of medical imaging in the diagnostic process 	 Largely inconsistent with current future directions in this field (eg: robotic assisted appendicectomy, intraoperative imaging, immunomodulation) 	 2 existent, valid publications 2 existent, but irrelevant publications 1 non-existent publication
6. What is the incidence of	 Accurate in identifying that pregnancy is a protective factor 	 Scattered statistics No comparison of incidence to the general population 	 5 non-existent publications



Discussion

Breadth of knowledge, fast processing speed, and ability to present the information succinctly Inconsistent in its accuracy, with answers frequently lacking in depth and specificity Commonly referenced fabricated or irrelevant publications, despite the existence of pertinent and real sources Integration of such technology could promote improved healthcare in resource limited settings Ongoing legal and ethical issues of the AI, including patient privacy, plagiarism of resources and spread of misinformation

Conclusion

Demonstrates great potential but none the less numerous pitfalls which question its use in clinical practice Necessitates further refinement before integration into the clinical settings Clinicians and all users must exercise vigilance when utilising ChatGPT for medical information purposes

