

# ChatGPT: promise and challenges for deployment in low- and middle-income countries

Xiaofei Wang,<sup>a,r</sup> Hayley M. Sanders,<sup>b,r</sup> Yuchen Liu,<sup>a</sup> Kennarey Seang,<sup>c</sup> Bach Xuan Tran,<sup>d,e</sup> Atanas G. Atanasov,<sup>f,g</sup> Yue Qiu,<sup>h</sup> Shenglan Tang,<sup>i</sup> Josip Car,<sup>j,k</sup> Ya Xing Wang,<sup>l,s</sup> Tien Yin Wong,<sup>m,n,o,s</sup> Yih-Chung Tham,<sup>m,p,q,s,\*</sup> and Kevin C. Chung<sup>b,s,\*\*</sup>



<sup>a</sup>Key Laboratory for Biomechanics and Mechanobiology of Ministry of Education, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, China

<sup>b</sup>Section of Plastic Surgery, Department of Surgery, University of Michigan Medical School, Ann Arbor, MI, USA

<sup>c</sup>Grant Management Office, University of Health Sciences, Phnom Penh, Cambodia

<sup>d</sup>Department of Health Economics, Institute for Preventive Medicine and Public Health, Hanoi Medical University, Hanoi, Vietnam

<sup>e</sup>Institute of Health Economics and Technology, Hanoi, Vietnam

<sup>f</sup>Ludwig Boltzmann Institute Digital Health and Patient Safety, Medical University of Vienna, Spitalgasse 23, 1090, Vienna, Austria

<sup>g</sup>Institute of Genetics and Animal Biotechnology of the Polish Academy of Sciences, Jastrzebiec, 05-552, Magdalenka, Poland

<sup>h</sup>Institute for Hospital Management, Tsinghua University, Beijing, China

<sup>i</sup>Duke Global Health Institute, Duke University, Durham, NC, USA

<sup>j</sup>Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University Singapore, Singapore

<sup>k</sup>Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

<sup>l</sup>Beijing Institute of Ophthalmology, Beijing Ophthalmology and Visual Science Key Lab, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China

<sup>m</sup>Singapore Eye Research Institute, Singapore National Eye Centre, Singapore

<sup>n</sup>Tsinghua Medicine, Tsinghua University, Beijing, China

<sup>o</sup>School of Clinical Medicine, Beijing Tsinghua Changgung Hospital, Beijing, China

<sup>p</sup>Centre for Innovation and Precision Eye Health, Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

<sup>q</sup>Ophthalmology and Visual Science Academic Clinical Program, Duke-NUS Medical School, Singapore

## Summary

In low- and middle-income countries (LMICs), the fields of medicine and public health grapple with numerous challenges that continue to hinder patients' access to healthcare services. ChatGPT, a publicly accessible chatbot, has emerged as a potential tool in aiding public health efforts in LMICs. This viewpoint details the potential benefits of employing ChatGPT in LMICs to improve medicine and public health encompassing a broad spectrum of domains ranging from health literacy, screening, triaging, remote healthcare support, mental health support, multilingual capabilities, healthcare communication and documentation, medical training and education, and support for healthcare professionals. Additionally, we also share potential concerns and limitations associated with the use of ChatGPT and provide a balanced discussion on the opportunities and challenges of using ChatGPT in LMICs.

**Copyright** © 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Large language model; ChatGPT; Public health; Global health; Low to middle income countries; Equity

## Introduction

“Can you do this for me?”

Most would wish for a personal assistant, ready at beck and call to tackle everything from mundane tasks to complex decisions. The evolution of artificial

intelligence (AI) chatbots has potentially transformed this fantasy into tangible reality. Worldwide, individuals ranging from students to business executives are increasingly leveraging AI chatbots for diverse purposes: from assisting with homework crafting professional emails, to seeking medical guidance.

Although many use AI chatbots to complete menial tasks, they have considerable potential for improving access to information, facilitating timely communication, enhancing self-care management, and providing contextually relevant support to patients. The World Health Organization (WHO) reports that at least half of the world's 7.3 billion people lack access to essential health services they need.<sup>1</sup> As a result, the life expectancy gap between countries

\*Corresponding author. Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

\*\*Corresponding author. Section of Plastic Surgery, The University of Michigan Health System, 1500 E. Medical Center Drive, 2130 Taubman Center, SPC 5340, Ann Arbor, MI, 48109-5340, USA.

E-mail addresses: [thamyc@nus.edu.sg](mailto:thamyc@nus.edu.sg) (Y.-C. Tham), [kechung@med.umich.edu](mailto:kechung@med.umich.edu) (K.C. Chung).

<sup>r</sup>Joint first authors.

<sup>s</sup>Joint last authors.

The Lancet Regional Health - Western Pacific 2023;41: 100905

Published Online xxx  
<https://doi.org/10.1016/j.lanwpc.2023.100905>

with the most and least expansive health coverage is 21 years.<sup>1</sup> These low- and middle-income countries (LMICs) often face complex public health challenges related to limited resources and practice limitations.<sup>1</sup> An example of this is the COVID-19 pandemic, with LMICs and their marginalized populations disproportionately affected as a result of existing health disparities.<sup>2</sup>

ChatGPT, a large language model (LLM) developed by OpenAI, is a publicly accessible AI chatbot. It is built on the Generative Pre-trained Transformer (GPT) architecture, with GPT-3.5 and GPT-4 as its backbones.<sup>3</sup> These models have been trained on a vast dataset, enabling them to engage in dynamic conversations across many languages. Despite being trained on data dated to September 2021, ChatGPT can integrate real-time or knowledge-based information through plugins, and GPT-4 can now process both image and text input.<sup>3</sup> ChatGPT possesses the potential to impact a substantial portion of the global population, including those in LMICs. This potential serves to elevate opportunities for advancements in public health within these regions.<sup>4</sup> After the release of ChatGPT, several similar LLMs emerged. These include Google's Bard, Microsoft's Bing, Baidu's Ernie, Anthropic's Claude, Alibaba's Tongyi Qianwen and many others. The potential of AI in medical care and public health programs has spurred a wealth of discussions, with a particular focus on the possible benefits for LMICs.<sup>5</sup> Nevertheless, despite AI tools demonstrating impressive performance in research settings, their translation into practical real-world applications, especially in LMICs, has been restricted.<sup>6,7</sup> This is primarily due to the challenges associated with integrating such tools into existing infrastructures, and the need for specific hardware and trained personnel to effectively leverage these technologies. On the contrary, ChatGPT is a publicly available service that requires minimal bandwidth, thus opening up the potential to reach a vast proportion of the global population, potentially effecting transformative change in medicine and public health. As the user base of ChatGPT expands and competitive alternatives emerge on the global stage, the currently high computational cost of ChatGPT might decrease, enhancing its affordability and accessibility, particularly in LMICs.<sup>8</sup> This unique potential positions ChatGPT distinctively from its predecessors in medical AI services.

In this viewpoint, we aim to discuss the potential advantages and drawbacks of employing ChatGPT to advance healthcare in low- and middle-income countries (LMICs) by examining its potential impact on three distinct levels: patients, healthcare professionals, and healthcare systems (Fig. 1). The insights shared herein reflect the authors' perspectives and judgements based on current evidence.

## Patient level

### Improving health literacy

Health literacy is crucial for making informed health decisions and reducing risks. However, in LMICs,

obstacles such as inaccessible or inaccurate information, language differences, and cultural customs often contribute to low health literacy. As a result, health education programs must be tailored to meet the specific needs of each population because approaches used in developed countries may not be directly applicable to LMICs.

ChatGPT has significant potential to increase health literacy, particularly in LMICs where healthcare access is often limited and qualified physicians are scarce.<sup>9</sup> First, ChatGPT can provide reasonably reliable disease information, including potential diagnosis and practical management suggestions relevant to patients' everyday lives. Table 1 contains results from a collection of studies on ChatGPT diagnosing different medical conditions.<sup>10-12</sup> These examples suggest ChatGPT's effectiveness in providing accurate advice on treating a wide range of conditions from more common ailments such as the common cold and cough<sup>13</sup> to more serious conditions like hypertension and diabetes mellitus.<sup>14</sup> ChatGPT even displayed sufficient, quick advice for treating emergencies like acute organophosphate poisoning, and chronic conditions like cirrhosis and hepatocellular carcinoma.<sup>15</sup> Fig. 2 shows a discussion between ChatGPT and one of the authors (HMS) about providing a potential diagnosis.

ChatGPT can improve the quality of life for people living with disabilities or chronic illnesses. Be My Eyes, a mobile app aimed for people with blindness or visual impairment, has integrated OpenAI's GPT-4 in its Virtual Volunteer tool.<sup>23</sup> ChatGPT can potentially offer a more cost-effective care service for users with visual impairment in LMICs, including India, Pakistan, Philippines, Iran and Ghana.<sup>24</sup> Through its image-to-text interpretation capabilities, it can further empower these individuals with greater independence.

Medical records laden with intricate terminology can be perplexing for patients, potentially leading to misunderstandings. ChatGPT offers a promising solution by making healthcare knowledge more digestible.<sup>17</sup> Fig. 3 illustrates an example of conversation between ChatGPT and one of the authors (HMS), where ChatGPT explains a diagnosis of ankylosing spondylitis in simple terms. A study conducted in Delhi, India, highlighted the efficacy of chatbots in delivering culturally attuned breastfeeding education, with chatbots addressing 88% of participants' inquiries.<sup>25</sup> By accessing the chatbot, patients are able to learn and ask questions about sensitive information in private environment. For instance, given its advanced capabilities, ChatGPT can potentially offer precise and user-friendly information on maternal and child health to expectant mothers and new parents. This exemplifies how ChatGPT can enhance health literacy in a secure, anonymous manner. Nonetheless, it is imperative for patients to consult with their physicians for definitive advice and before making any major health decisions.

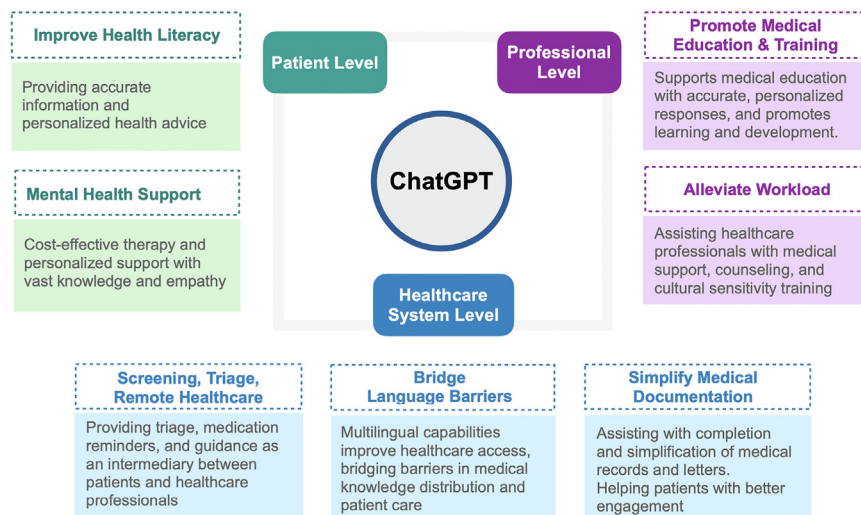


Fig. 1: The impact of ChatGPT on medicine and public health in low- and middle-income countries.

Physicians, as licensed experts, possess the requisite training to provide medical counsel, whereas ChatGPT disseminates information based on its database, which may not always be current or accurate. Therefore, patients should prioritize physician consultations before taking any major medical actions.

### Mental health support

Inequities in mental health care emerge as a prevalent issue in LMICs.<sup>26</sup> According to a WHO report, around 76%–85% patients with serious mental disorders did not receive any treatment in these countries.<sup>27</sup> The National Mental Health Survey of India reported a treatment gap of 83% for any mental disorder.<sup>28</sup> In Cambodia, a LMIC in Southeast Asia, the WHO reports the fewest mental health service providers relative to other nations in the Western Pacific Region.<sup>29</sup> This shortfall stems from inadequate training of mental health professionals to manage and treat severe mental illness, further exacerbated by resource constraints.

ChatGPT and similar applications are emerging as viable alternatives for reducing the gap in mental health services. With the ability to produce human-like responses, ChatGPT potentially offers companionship, support, and even therapy to individuals grappling with psychological challenges. Notably, ChatGPT is not only well-versed in psychiatric conditions but also exhibits a semblance of “empathy” in its responses. Research has demonstrated that ChatGPT provided high-caliber answers to patient queries in an online forum. Impressively, ChatGPT’s responses were perceived as markedly more empathetic than those of physicians, with the frequency of empathetic or highly empathetic responses being 9.8 times greater for ChatGPT than for medical professionals.<sup>30</sup> Additionally, certain

iterations of ChatGPT are freely accessible from virtually any location, enabling patients to confide their personal experiences and feelings without the potential unease of an in-person consultation. Consequently, ChatGPT, when viewed as a mental health resource, holds promise in mitigating or even obliterating financial and logistical barriers to care.

A few chatbots have already been developed for mental health purposes. For example, Shim is a smartphone app that engages users in dialogues on positive psychology, providing tailored responses based on pre-written dialogues and user inputs.<sup>31</sup> The Cambodian Ministry of Women’s Affairs created an app with a chatbot to assist women and girls in reporting violence or abuse and offer timely and vital counseling support.<sup>32</sup> However, these chatbots are often considered shallow and confusing by users.<sup>33</sup> ChatGPT, an advanced instance of NLP embodied in a LLM, was trained on an extensive dataset. This training enables it to produce a broader spectrum of responses tailored to user preferences, such as tone and response length. In contrast, older chatbots are often confined to specific contexts or predetermined response templates, limiting their utility.<sup>34</sup> One of ChatGPT’s notable features is its ability to continuously adapt to user interactions.<sup>3</sup> As a result, it has the potential to offer more personalized emotional advice and support. Fig. 4 shows a discussion between ChatGPT3.5 version and one of the authors (HMS), simulating mental health symptoms. ChatGPT was able to respond to user queries and offers strategies to improve mental health.

While recognizing its potential benefits, it’s important to recognise that ChatGPT, like any tool, is not without limitations, especially when deployed in sensitive domains such as mental health support. An

Medical domain	Study details	Performance metrics	Findings	Reference
Diabetes	Differentiate ChatGPT-generated and human answers with 10 questions about diabetes.	59.5% participants could identify ChatGPT-generated answers with a binary assessment.	The strongest correlation is with previous ChatGPT use (p = 0.003) among all factors. Linguistic features might hold a higher predictive value than the content itself.	Hulman et al., 2023 <sup>14</sup>
Microbiology	Answers to 96 first- and second-order questions based on the competency-based medical education (CBME) curriculum.	Mean score was 4.07 (±0.32)/5, which had an accuracy rate of 80%.	ChatGPT has the potential to be an effective tool for automated question-answering in microbiology.	Das et al., 2023 <sup>10</sup>
Medical responses	Answers to 284 medical questions across 17 specialties.	Largely correct responses with mean accuracy score of 4.8/6 and mean completeness score of 2.5/3.	Substantial improvements were observed in 34 of 36 questions scored 1-2. ChatGPT generated largely accurate information although with important limitations.	Johnson et al., 2023 <sup>11</sup>
Cancer	Comparison with NCI's answers about 13 questions on a cancer-related web page.	96.9% overall agreement on the accuracy of cancer information.	Few noticeable differences were in the number of words or the readability of the answers from NCI or ChatGPT.	Johnson et al., 2023 <sup>16</sup>
Cirrhosis and hepatocellular carcinoma (HCC)	Performance comparison between ChatGPT and physicians or trainees in 164 questions.	An overall accuracy of 76.9% in quality measures.	ChatGPT demonstrated strong knowledge of cirrhosis and HCC, but lacked comprehensiveness in diagnosis and preventive medicine.	Yeo et al., 2023 <sup>15</sup>
Prostate cancer	Comparison between five state-of-the-art large language models in providing information on 22 common prostate questions.	ChatGPT had the highest accuracy rate and comprehensiveness score among all five Large Language Model, and had satisfactory patient readability.	Large Language Models with internet-connected dataset was not superior to ChatGPT. Paid version of ChatGPT did not show superiority over the free version.	Zhu et al., 2023 <sup>17</sup>
Toxicology	Response to a case of acute organophosphate poisoning.	Subjective assessment.	ChatGPT fared well in answering and offered good explanations of the underlying reasoning.	Abdel-Messih et al., 2023 <sup>18</sup>
Shoulder Impingement Syndrome (SIS)	Analyzed the ability of providing information for SIS.	Subjective assessment.	ChatGPT could provide useful medical information and treatment options for patients with SIS, including symptoms, similar diseases, orthopedic tests, and exercise recommendations. But potential biases and inappropriate information must be taken into account.	Kim J hee, 2022 <sup>19</sup>
Infection	Antimicrobial advice in eight hypothetical infection scenario based questions.	Subjective assessment of appropriateness, consistency, safety, and antimicrobial stewardship implications.	ChatGPT exhibited deficiencies in situational awareness, inference, and consistency in clinical practice.	Howard et al., 2023 <sup>20</sup>
Ophthalmology	Tested the accuracy on two question banks used for the high-stakes Ophthalmic Knowledge Assessment Program (OKAP) exam.	55.8% and 42.7% accuracy in the two 260-question simulated exams.	Performance of ChatGPT varied across ophthalmic subspecialties. Domain-specific and pre-training might improve its performance.	Antaki et al., 2023 <sup>21</sup>
Neuropathic pain	Tested the performance on 50 pairs of causal relationships in neuropathic pain diagnosis.	ChatGPT tended to make false negative mistakes, and showed high precision and low recall from the binary assessments of true/false.	ChatGPT lacked in consistency and stability in the context of neuropathic pain diagnosis. Using Large Language Models' causal claims as causal discovery results requires caution because of fundamental differences in the tasks and biases in causal benchmarks.	Tu et al., 2023 <sup>22</sup>

**Table 1: Summary of studies evaluating ChatGPT's performance in various medical domains.**


unfortunate incident in Belgium, involving an interaction between an adult male and an AI chatbot on the Chai app, underscores the potential implications if AI chatbots do not adhere strictly to necessary safety protocols.<sup>35</sup> This incident serves as a stark reminder of the utmost importance of these safeguards. ChatGPT, as part of its safeguards, employs safety mechanisms to prohibit the generation of content that promotes self-harm or suicide.<sup>36</sup> If such content is detected, ChatGPT promptly alerts the user of the content's violation against its policy and encourages the individual to seek support from trusted sources. Nevertheless, as the aforementioned case demonstrates, the possibility of manipulating prompts to circumvent these safety protocols remains a pressing concern.<sup>37</sup> Therefore, it is crucial that ChatGPT and similar chatbots are




continually monitored and updated based on the latest research and findings in mental health and AI ethics. This will require a multidisciplinary approach, bringing together experts in AI, mental health, ethics, regulators, and other relevant fields to ensure that these systems are as safe and effective as possible.

## Healthcare professional level

### Promoting medical education and training

Medical training in LMICs has continuously encountered barriers and challenges in the effort to produce well-trained physicians. In a study in Pakistan,<sup>38</sup> only 56% of physicians met an acceptable diagnostic standard for viral diarrhea, and only 35% met the acceptable standard for treatment. A study conducted to evaluate


 Hi ChatGPT, my doctor is unavailable to see me right now. My symptoms are stiffness, fever, and swollen joints, especially in my fingers. Can you provide me a list of potential diagnoses?




 I'm sorry to hear that you're not feeling well. Based on the symptoms you've described, the following are some potential diagnoses that you may want to consider:  

1. Rheumatoid Arthritis: Rheumatoid arthritis is a chronic autoimmune disease that causes inflammation in the joints, leading to stiffness, swelling, and pain. It commonly affects the fingers and can cause fever.
2. Systemic Lupus Erythematosus (SLE): SLE is an autoimmune disease that can cause joint pain and swelling, fatigue, and fever. It can also affect other parts of the body, such as the skin, kidneys, and lungs.
3. Gout: Gout is a type of arthritis that occurs when uric acid builds up in the joints, leading to sudden and severe pain, redness, and swelling. It can also cause fever.
4. Lyme Disease: Lyme disease is an infectious disease caused by the bacterium *Borrelia burgdorferi*, which is transmitted to humans through the bite of infected ticks. It can cause joint pain and swelling, fever, and other flu-like symptoms.
5. Viral Arthritis: Viral infections, such as hepatitis B, hepatitis C, and parvovirus B19, can cause joint pain and swelling, fever, and other symptoms.
6. Reactive Arthritis: Reactive arthritis is a type of arthritis that occurs as a result of an infection in another part of the body, such as the urinary or gastrointestinal tract. It can cause joint pain and swelling, fever, and other symptoms.

Please note that this list is not exhaustive and is provided for informational purposes only. It's important to consult with a medical professional for a proper diagnosis and treatment plan. If your symptoms are severe or worsening, please seek immediate medical attention.

Fig. 2: ChatGPT provides potential diagnoses when presented with a list of symptoms.

 My doctor diagnosed me with ankylosing spondylitis, but I'm not sure what that means. Can you explain it in simple terms?

 Ankylosing spondylitis is a type of arthritis that primarily affects the spine, although other joints may also be involved. It is a chronic inflammatory condition that can cause pain, stiffness, and limited mobility, especially in the lower back and hips.  


The exact cause of ankylosing spondylitis is unknown, but it is believed to be related to a combination of genetic and environmental factors. Symptoms usually develop in early adulthood and may include:




- Pain and stiffness in the lower back and hips, which may be worse in the morning or after periods of inactivity
- Pain and stiffness in other joints, such as the shoulders, knees, and ankles
- Fatigue and general feeling of being unwell
- Reduced flexibility in the spine and other joints
- Difficulty breathing if the chest wall becomes affected

Ankylosing spondylitis is a chronic condition, meaning it can last for a long time or even for the rest of your life. However, with proper treatment and management, it is possible to control symptoms and improve quality of life. Treatment may include medications to reduce inflammation and manage pain, physical therapy to improve mobility and flexibility, and lifestyle modifications such as exercise and maintaining good posture.

It's important to work closely with your doctor to develop a treatment plan that is tailored to your individual needs and symptoms. They can also provide you with additional information and support as you navigate your diagnosis and manage your condition.

Fig. 3: ChatGPT simplifies complex medical terminology.

 I'm feeling lonely and sad. Is there something wrong with me? What should I do?

 Feeling lonely and sad is a common experience that many people go through at various times in their lives. It's important to remember that you are not alone in this feeling, and there are things you can do to help yourself feel better.  

Firstly, it's important to understand that feeling sad or lonely does not necessarily mean that there is something wrong with you. It's a normal part of the human experience to feel sad or lonely from time to time. However, if these feelings persist for an extended period of time or interfere with your daily life, it may be helpful to seek professional support from a therapist or counselor.

In the meantime, here are some things you can do to help improve your mood:

1. Reach out to others: Connect with friends or family members who can provide support and comfort. Consider joining a social group or club that aligns with your interests.
2. Practice self-care: Engage in activities that make you feel good, such as exercising, reading, or listening to music. Take care of your physical health by getting enough sleep and eating a balanced diet.
3. Seek professional help: Consider talking to a therapist or counselor who can provide guidance and support. They can help you develop coping strategies and work through any underlying issues that may be contributing to your feelings of loneliness and sadness.

Remember that it's okay to ask for help when you need it, and there is no shame in seeking support. Take small steps each day to improve your mood and well-being, and remember that things can and will get better.

**Fig. 4:** ChatGPT provides empathetic support when presented with mental health symptoms.


pediatric care in Papua New Guinea found that the majority of physicians and other team members (69%) only checked for two out of four examination criteria for pneumonia cases.<sup>39</sup> Challenges for medical education in LMICs can be attributed to insufficient financial and technological resources, undeveloped education models (i.e., teacher-centered approach instead of student-centered approach), and a lack of updated medical information.<sup>40</sup> Below, we outline the capability of ChatGPT to complement medical teaching, particularly in LMICs.




ChatGPT, particularly with the GPT-4 engine, can serve as a valuable supplementary tool for medical education in resource-limited settings. ChatGPT performed well on medical examinations such as United States Medical Licensing Examination (USMLE)<sup>41,42</sup> and Korean general surgery board exam.<sup>43</sup> A recent study evaluated ChatGPT's potential in medical question answering, finding it provided logical, informative responses, equivalent to a third-year medical student's passing score.<sup>44</sup> The performance of ChatGPT on the Medical College Admission Test (MCAT) demonstrated its ability to accurately answer standardized test questions while offering sufficient, logical explanations.<sup>45</sup>

Kung et al. conducted a study that demonstrated ChatGPT's capability to produce comprehensive reasoning and valid clinical insights during the USMLE, suggesting its potential to assist medical students in

education.<sup>42</sup> ChatGPT demonstrated the capability to generate answers and explanations with a notable concordance of 94.6% across all USMLE questions.<sup>42</sup> Such advancements potentially enable human learners to more effectively understand the logical sequence and inherent relational directionality present in the explanations offered by ChatGPT. In addition, ChatGPT allows medical students to ask questions about specific medical concepts and receive accurate and personalized responses to help them better structure their knowledge around each concept. [Fig. 5](#) shows an interaction between ChatGPT and one of the authors (HMS), in which ChatGPT furnishes the user with mnemonic devices for remembering the Krebs Cycle. This holds promising potential, especially in LMICs where traditional medical education resources may be scarce. Furthermore, ChatGPT can offer suggestions to medical educators regarding course outlines and structure, further augmenting the medical education process.

ChatGPT can revolutionize medical education by simulating case reports, offering a dynamic, case-based learning approach. This is invaluable when access to real patient data is limited due to consent or confidentiality concerns. ChatGPT's ability to generate realistic yet fictionalized patient interactions not only ensures diverse clinical exposure but also upholds ethical standards. By integrating this tool into medical curricula, students can engage with simulated cases in real-time,

 I'm a first-year medical student learning about the Krebs cycle and I can't seem to fully remember the sequence of events. Can you teach me another way to remember this?

 Sure, there are several mnemonic devices that can help you remember the sequence of events in the Krebs cycle. Here are a few examples:  

1. "Citrate Is Krebs' Starting Substrate For Making Oxaloacetate"  
This mnemonic uses the first letter of each word to represent each step in the cycle: citrate, isocitrate, alpha-ketoglutarate, succinyl-CoA, succinate, fumarate, malate, and oxaloacetate.
2. "Can I Keep Selling Seashells For Money Officer?"  
This mnemonic uses the first letter of each word to represent each step in the cycle in reverse order: citrate, isocitrate, alpha-ketoglutarate, succinyl-CoA, succinate, fumarate, malate, and oxaloacetate.
3. "Our City Is Kept Safe And Secure From Malicious Outside Forces"  
This mnemonic uses the first letter of each word to represent each step in the cycle, starting with oxaloacetate: oxaloacetate, citrate, isocitrate, alpha-ketoglutarate, succinyl-CoA, succinate, fumarate, and malate.

By using these or other mnemonics, you may find it easier to remember the steps of the Krebs cycle. Additionally, it may be helpful to practice drawing out the cycle and to work through practice problems to reinforce your understanding. Good luck with your studies!

Fig. 5: ChatGPT provides learning strategies.

while fostering interactive learning and critical thinking. Leveraging on ChatGPT as a pedagogy tool may represent a new paradigm shift in blended learning. Finally, medical training and education are lifelong processes for healthcare professionals, and keeping up with the latest research, techniques, and guidelines can be challenging, especially in LMICs. Through plugins, ChatGPT can provide instant access to relevant up-to-date medical information and resources for healthcare professionals, making it easier to support their ongoing learning and development, improve their skills and knowledge, and therefore provide better care to patients.<sup>46</sup>

### Alleviating workload of healthcare professionals

Healthcare professionals in LMICs are often confronted with overwhelming work pressure that stems from a complex interplay of factors. The patient load in these regions is typically high, a situation exacerbated by a chronic shortage of healthcare providers. Resources are often limited, which makes patient care even more challenging. Additionally, the lack of sufficient institutional and professional support in terms of professional development opportunities—compounds these challenges. These challenges potentially compromise the quality of care and the well-being of healthcare providers themselves. In this context, tools like ChatGPT may present opportunities to alleviate some of these pressures.

For instance, ChatGPT can assist in answering primary care level medical queries, offering comprehensive treatment strategies and recommendations for healthcare professionals.<sup>47–49</sup> It can potentially assist health professionals in differential diagnosis, evaluating risks, formulating personalized treatment plans, and diagnosing complex conditions by providing relevant information. In a recent study by Balas et al.,<sup>50</sup> ChatGPT was found to have potential value in diagnosing ophthalmic conditions. For provisional diagnosis, the correct diagnosis was identified in 9/10 cases. For differential diagnosis, the correct answer was obtained in all 10/10 cases from all listed possibilities of differentials.

Despite the promising potentials, deployment of ChatGPT and LLMs alike must be done judiciously and with careful consideration of the unique challenges inherent to LMICs. A salient issue in many LMICs is the pronounced shortage of physicians.<sup>51</sup> This scarcity means that the responsibility of patient care often rests heavily on a limited number of healthcare professionals. Introducing ChatGPT into this healthcare landscape presents both potential challenges and benefits. On one hand, there's a risk that the already limited physicians might face an increased workload. This could arise if patients, after consulting with ChatGPT, require medical intervention due to following potentially incorrect advice. Conversely, ChatGPT could serve as a valuable tool in alleviating some of the pressure on these

physicians. By offering prompt solutions to routine medical inquiries, ChatGPT might reduce the number of standard consultations, enabling physicians to dedicate more time to complex cases.

## Healthcare systems level

### Screening, triage and remote healthcare support

The lack of access to healthcare is a major issue for LMICs, with several causes contributing to significant challenges. These causes include physician shortages, inadequate healthcare infrastructure, and extensive travel distances to healthcare facilities, particularly in regions lacking specialists. In areas with limited healthcare access, like in remote regions, it is crucial to provide basic information that can help individuals identify symptoms and triage health concerns. In such settings, technology emerges as a useful assistive tool. Telemedicine services and AI-assisted triage and screening tasks have shown great promise in addressing these healthcare gaps in LMICs.<sup>52–54</sup>

ChatGPT can serve as a valuable bridge between patients and healthcare professionals, aiding in primary health assessments, appointment scheduling, triage recommendations, medication reminders, and follow-up care. This integration can be realized by incorporating ChatGPT services into existing systems via its APIs or plugins.<sup>55</sup> Through interactive conversations and targeted questioning, ChatGPT can gather patient medical histories and offer insights into the urgency and type of medical attention needed. By utilizing plugins, ChatGPT can actively monitor patient conditions, send medication reminders, schedule check-ups, and assess if urgent care is necessary. Furthermore, the insights provided by ChatGPT can be invaluable for physicians, enhancing their diagnostic and treatment strategies.

Furthermore, ChatGPT presents capabilities that could enhance telemedicine, potentially broadening the scope of remote medical services. Utilizing its API, ChatGPT can potentially be integrated into telemedicine platforms, offering timely, evidence-based insights to aid clinical decision-making. For example, when users engage with a telemedicine platform, ChatGPT can streamline the initial patient surveys, collecting health histories and capturing essential symptom data. This information can be systematically organized and relayed to physicians for subsequent consultation, whether virtual or in-person. Moreover, during these consultations, ChatGPT can support practitioners by accessing pertinent medical literature, guidelines, and drug information. The reliability and effectiveness of ChatGPT as a virtual medical assistant have been tested in various fields (Table 1),<sup>12</sup> including prostate cancer,<sup>17</sup> cardiovascular disease,<sup>56</sup> infection,<sup>20</sup> ophthalmology,<sup>21</sup> and neuropathic pain.<sup>22</sup> In addition, ChatGPT can provide drug information such as side effects, interactions,

contraindications, dosage, administration guidelines, storage requirements, and potential alternatives. However, it is important to acknowledge that, in its current version, ChatGPT's responses cannot replace medical advice from a doctor or other healthcare professional. Hence, patients should not blindly follow ChatGPT's advice without the proper direction from their doctor. In addition, by acting as virtual assistants, ChatGPT can promote remote healthcare management through appointment booking, treatment facilitation, and health information management.

### Simplifying medical documentation

In LMICs, medical professionals often encounter challenging circumstances such as inadequate resources, manpower, and deficient infrastructure. The documentation of medical records is particularly time-consuming and resource-intensive. ChatGPT has the potential to assist with administrative tasks such as clinical letter writing, discharge summaries, clinical vignettes, and simplified reports. ChatGPT has been shown to create clinical letters with naturalness and readability comparable to human-generated letters.<sup>57</sup> ChatGPT can also help fine-tune language and style by suggesting appropriate vocabulary, grammar, and sentence structures. Discharge summaries, frequently sidelined due to time limitations and pressing clinical duties, can be crafted by ChatGPT. This not only eases the burden on junior doctors but also bolsters patient safety during transitions in care.<sup>58</sup> Nevertheless, users should remain vigilant, refraining from entering sensitive or private data into ChatGPT and always verifying its outputs for accuracy. A recent study suggested that most radiologists agree that ChatGPT-generated simplified reports are complete and safe for patients, allowing for better understanding and preparation for healthcare interactions.<sup>59</sup>

### Bridging language barriers in medical knowledge distribution and patient care

English is the predominant language used in medicine,<sup>60,61</sup> on the internet<sup>62</sup> and in academic publishing.<sup>63</sup> Among 80 LMICs, 27 countries' official languages include English, while 25 countries' official languages include other high-resource languages (>1% of data in the CommonCrawl), 12 countries' official languages include medium-resource languages (>0.1%), and 16 countries' official languages are low-resourced (<0.1%).<sup>64,65</sup> English is an academic subject, yet not commonly spoken in LMICs. Conversely, some LMICs might not include English in their curricula at all. This linguistic gap can hinder physicians in accessing current and accurate medical data, predominantly available in English. Patients, too, might struggle to find pertinent health details due to language barriers. The linguistic tapestry of many LMICs is rich; for instance, India recognizes 22 distinct official languages. In such



multilingual contexts, physicians often rely on interpreters or translation services to bridge communication gaps with patients.<sup>66</sup> ChatGPT's exceptional multilingual capabilities, enabled by the GPT-4 backbone,<sup>67</sup> addressing the intricacies of language diversity and enhancing communication. These capabilities offer potential benefits in two significant ways.

Firstly, ChatGPT's multilingual capabilities enable it to deliver informed responses across all supported languages, even if not directly trained in a specific language. In certain NLP tasks, ChatGPT demonstrated comparable or superior performance for some low-resource languages relative to high- and medium-resource ones. This suggests that ChatGPT's efficacy is not necessarily diminished for low-resource languages prevalent in LMICs.<sup>68</sup> Such capability may be invaluable for physicians and patients seeking medical and public health insights in their native tongues that is not English. Access to dependable medical information in native languages can empower both healthcare providers and recipients, fostering improved health outcomes.

Secondly, ChatGPT offers high-quality translations,<sup>67</sup> enhancing the readability and accuracy of translated information. If harnessed responsibly, physicians can gain a deeper understanding of intricate medical scenarios and leverage this tool for enhanced communication with linguistically diverse patients. Patients will also be able to develop a clearer understanding of their health conditions and treatment options, and become more engaged in their care, ultimately resulting in better health outcomes. The continued evolution and deployment of ChatGPT and other LLMs are poised to significantly bridge the global health chasm, championing universal access to health information and services.

### Potential concerns, limitations, proposed approaches, and roadmap for implementation

While the potential of ChatGPT to enhance healthcare in LMICs is considerable, it is also crucial to identify and recognise its limitations and potential concerns. In particular, there is a concern that application of ChatGPT in LMICs with lower literacy levels and limited digital literacy may not directly benefit the intended population, despite the well intention. On this note, ensuring data accuracy, addressing misinformation and biases, comprehending the relevance of context, adhering to ethical guidelines, mitigating overreliance, and promoting equal access to technology should be considered holistically to effectively integrate ChatGPT into medicine and public health settings. Furthermore, validation of ChatGPT-based interventions in community and clinical settings should be further explored and the legal and infrastructure requirements for implementing ChatGPT and similar services in LMICs should be assessed.

### ChatGPT's reproducibility and reliability

As a LLM, ChatGPT statistically predicts words based on their precedence in previously seen texts without truly comprehending the knowledge generated or the prompts provided.<sup>69</sup> Its reliability depends on the accuracy and quality of its training data, which remains undisclosed and likely has a limited focus on medicine and public health. This undisclosed data may contain inaccuracies, leading to the absorption of misinformation, imbalanced content, and biases during the training process.<sup>69</sup> Due to these factors, ChatGPT can occasionally yield inconsistent answers and may not fully capture the intricacies of medical sciences.<sup>70</sup> Moreover, when tasked with recognizing unanswerable questions, ChatGPT's accuracy stood at a mere 27.8%. This suggests its inclination to offer speculative responses rather than declining such queries, a critical aspect in the medical field.<sup>71</sup> Its output is also influenced by its assigned role or perspective (such as "you are a helpful assistant") in conversations. In more severe instances, it may deliver potentially harmful response if the prompt is malicious.<sup>48</sup> It is also vulnerable to adversarial attacks, where even an alternation of a single character can significantly impact its response.<sup>71</sup> Thus, responses generated by ChatGPT in the medical context cannot be entirely trusted and ought to be cross checked by a trained physician. However, in LMICs, the scarcity of specialized medical training might amplify these flaws and biases, as physicians may struggle to identify inaccurate information. Further research and development are needed to enhance the reliability of ChatGPT in the medical domain, ensuring its effective application in LMICs.

In addition to ChatGPT, there are domain-specific LLMs like Google's Med-PaLM and medAlpaca, which are fine-tuned specifically for medical question answering. Research suggests that while GPT-4 hasn't been tailored specifically for medical knowledge, it surpasses Med-PaLM in the medical domain.<sup>41</sup> Yet, with the introduction of the Med-PaLM 2 model, its performance either matches or even exceeds that of the general-purpose ChatGPT-4 in certain medical question-answering scenarios.<sup>72</sup> The continuous advancement of domain-specific LLMs bodes well for the future, promising to enhance the precision and dependability of LLM-generated responses, which holds significant potential for patient-focused applications.

Traditionally, individuals turned to internet searches for health information, requiring discernment to assess the accuracy of online content. ChatGPT, while facing similar challenges, has been shown to provide answers that are either superior to or on par with Google search results in cardiovascular areas.<sup>73</sup> While it can offer valuable information and guidance for common diseases and public health related topics, the extent to which it can function

autonomously, i.e., whether its use should be limited to generic medical queries or extend to more advanced, patient-centric contexts, remains to be determined. Currently, in response to standard medical questions in English, ChatGPT using the GPT-4 engine often prefaces with “I am not a doctor, but...”. While this disclaimer emphasizes ChatGPT’s non-medical nature, it might not deter all from exclusively depending on its guidance. Rigorous evaluation of ChatGPT’s medical accuracy and safety is crucial to guide its use and help users discern its advice. Enhancing its accuracy requires not only diverse data, particularly from LMICs, but also a more structured, context-aware approach to health queries to prevent potentially misleading or harmful responses.

### **Disproportionate proficiency in language and social context**

Although multilingual, ChatGPT’s training data are primarily English-based and fine-tuned with prompts mostly in English,<sup>74,75</sup> leading to a disparity in language proficiency and potentially inferior performance in non-English languages<sup>68</sup> which are often spoken in LMICs. This potential limitation ought to be highlighted. Furthermore, recent report indicated that ChatGPT’s responses tend to align closer with Western culture and may adapt less effectively to other cultural contexts.<sup>76</sup> This could indicate that ChatGPT may lack cultural and religious appropriateness when implemented in LMICs. Thus, ChatGPT’s limited understanding of complex social and cultural aspects relevant to medicine and public health could impede its applications in LMICs. One example where cultural consideration is especially important is in the realm of medical education. Students must be able to care for patients while considering cultural context and customs. ChatGPT’s Western-typical responses may impede this important education if ChatGPT is used to create curricula in a LMIC. For instance, cultural variations are relevant especially in the discussion of mental health. Presentation of mental health symptoms have been known to vary across cultures, and therefore may require cultural specific-sensitivity when devising management and treatment plans for patients.<sup>77</sup> The vast cultural diversity across LMIC regions may exacerbate this challenge. While evaluation tasks exist to gauge cultural differences in LLMs, there is a notable scarcity of studies that delve into cultural and religious appropriateness within healthcare. This gap underscores the need to infuse patient perspectives and localized contexts into the next phase of LLM development and refinement. Embracing a patient- and population-centric methodology could address this disparity, fostering tailored care in LMIC environments.<sup>78</sup> Furthermore, adherence to ethical norms and best practices in digital safety and privacy is paramount. This would further ensure safety and contextual relevance in LMICs.

### **Infrastructure and accessibility challenges**

The computational power demanded by ChatGPT and similar LLMs currently necessitates dedicated servers and internet connections, which poses challenges for adoption in LMICs due to infrastructure limitations and insufficient internet access or electronic devices. According to a report from the Alliance for Affordable Internet (A4AI), basic internet access is only available to half of the population in LMICs, and of those, only 10% have what is referred to as “meaningful connectivity” or a decent internet connection.<sup>79</sup> However, the ongoing development of LLMs and innovative internet solutions hold promise for mitigating these challenges in the near future.

Recent advancements, such as Stanford’s Alpaca model fine-tuned from Meta’s LLaMA, demonstrate the feasibility of operating LLMs on mobile devices, potentially enabling offline use in rural areas with limited connectivity. Additionally, creative internet solutions like SpaceX’s satellite-based Starlink system and the continued exploration of balloon technology for internet provision, such as Project Loon pioneered by Google, can help bridge the connectivity gap in remote regions.<sup>80</sup>

By partnering with NGOs and alternative internet providers, it is plausible that advanced LLMs like ChatGPT will become more universally accessible and available, regardless of geographical locations.

### **Implementation issues of ChatGPT in LMICs**

The potential integration of ChatGPT into the medical landscape of LMICs necessitates a thorough examination of liability and ethics. Key questions arise: Who bears responsibility for adverse outcomes stemming from ChatGPT’s advice? How would medical malpractice involving ChatGPT be addressed? How is user data safeguarded? A recent viewpoint By Harrer offers a potential regulatory blueprint, emphasizing the formation of AI ethics committees with decisive authority and advocating for collaboration among developers, providers, patients, and other stakeholders.<sup>69</sup>

User privacy and data protection represent a critical issue that requires careful attention. It is essential to adhere to strict data protection guidelines and ethical standards, as well as maintaining transparency in communication to users. According to OpenAI, user’s conversations, including any personal data the user input, were shared with the company by default and may be used to improve their models.<sup>81</sup> Research has demonstrated that, despite ChatGPT’s dialog safety measures against revealing personal information, certain prompting engineering methods, such as a combination of jailbreak and Chain-of-Thoughts prompts, are able to compromise its safety defence mechanism.<sup>82</sup> When analyzing health-related data with ChatGPT, there is a risk of inadvertently violating patient privacy, particularly when dealing with sensitive personal health information. To mitigate this risk, it is

crucial to ensure that such applications follow strict data protection guidelines and adhere to privacy regulations. In addition, users must be aware of what information is appropriate to input into ChatGPT. Also, transparency is vital in building trust and maintaining ethical standards. Many countries, including the European Union, the United States, and China, have started discussions on the need to regulate services like ChatGPT.<sup>83</sup> Recently, Italy took a significant step by temporarily banning ChatGPT, highlighting the need to address ethical concerns related to generative AI at an international level. However, many LMICs, particularly in South-Asia and Southeast Asia, lack the legal provisions and enforcement mechanisms to counter data breaches and cybercrime, potentially complicating the deployment of ChatGPT in LMICs. Consequently, while ChatGPT's potential is evident, its practical implementation in LMICs is shrouded in uncertainties, attributed to sparse legal backing for e-health solutions, health system limitations, and a slower digital transition pace in LMICs. ChatGPT's current limitation in offering evidence-based advice contrasts with the rigorous standards of clinical practice. Physicians typically provide recommendations after thorough patient evaluations, a nuance ChatGPT might miss in its broad response spectrum. Additionally, cultural and social beliefs can impede AI adoption. Research indicates a patient preference for in-person consultations over remote alternatives, even if more costly. While ChatGPT and similar digital health tools aren't intended to supplant face-to-face interactions, they serve as valuable adjuncts. The COVID-19 pandemic exemplified this, with a surge in digital health technology adoption, both in direct pandemic response and broader health service delivery during lockdowns. While the irreplaceability of traditional medical examinations stands, emerging evidence suggests potential benefits of remote health services, especially for underserved groups.<sup>84,85</sup> Nevertheless, rigorous oversight and monitoring are crucial to prevent misinformation from LLM chatbots like ChatGPT.

#### **Overreliance on ChatGPT could stifle critical thinking and decision making**

Excessive reliance on ChatGPT-generated insights and recommendations might result in a decline in critical thinking and decision-making abilities among healthcare professionals. In fact, studies have shown that incorrect AI predictions could cause clinicians to doubt their own initial judgment, regardless of their experience, especially when they lacked confidence in their judgment in the first place.<sup>86</sup> This could be exacerbated in LMICs where healthcare professionals generally have limited access to specialized and high-quality training and resources, making them more susceptible to the influence of AI-generated information and potentially compromising the quality of care delivered to patients. It is essential to strike a balance between leveraging AI

technology and maintaining human expertise, ensuring that professionals continue to exercise their judgment and expertise even when supplemented with AI-generated information.<sup>78</sup>

Furthermore, ChatGPT may not have a complete understanding of the complex intricacies associated with various medicine and public health issues. Overreliance on AI-generated information may also inadvertently give rise to oversimplified solutions for complex medical topics. Such oversimplified solution generated by ChatGPT will not be sufficient and effective in addressing multifaceted public health challenges.

Taken together, user education plays a critical role in ensuring that users understand ChatGPT and the role it can play in their healthcare or health education. It is important to prioritize user education programs that explain the working principles of ChatGPT in simple and accessible language. Moreover, it is essential to emphasize the role of ChatGPT as an augmenting tool rather than a replacement for human decision-making. By emphasizing on the importance of user education, one can better ensure that the use of ChatGPT in healthcare is done responsibly and ethically.

#### **Roadmap and future considerations**

For the effective implementation of ChatGPT and similar LLMs in LMICs, a deep understanding of both barriers and potential solutions is paramount. Given the rapid evolution of LLMs and the potential of ChatGPT in healthcare, we propose the following roadmap: ChatGPT's initial applications in LMICs can be multifaceted. It can generate culturally sensitive health messages, broadening its reach to wider audiences. Additionally, it can potentially serve as an administrative assistant, supporting health professionals in tasks such as drafting clinical letters and creating discharge summaries. Furthermore, as detailed above, ChatGPT can also significantly influence medical training in LMICs. As we progress, fine-tuning, customization, and the subsequent evaluation of LLM become paramount. LLMs should be adapted to resonate with local languages and cultures in LMICs, ensuring cultural appropriateness. Research should be conducted to assess the utility of ChatGPT in diverse settings, especially in aiding clinical decision-making. Concurrently, there is a need to develop robust regulations for its application in complex scenarios. In light of the considerable potential for LLMs like ChatGPT in healthcare, particularly in LMICs, there are several key considerations to ponder for the future. User induction and education will be central to ChatGPT's success. Initiatives should be launched to test ChatGPT's practicality in varied healthcare contexts within LMICs. These efforts should be complemented with capacity-building initiatives and public awareness campaigns, ensuring that both professionals and the public are well-informed. In addition, monitoring and research will be

the backbone of this roadmap. Frameworks should be established to rigorously assess ChatGPT's efficacy, safety, and overall impact, all while considering the unique contexts of LMICs. This evaluation should encompass observational studies, randomized trials, and user feedback. On the regulatory front, there is a pressing need for comprehensive oversight. Regulatory blueprints addressing liability, while promoting innovation and safeguarding patient well-being, should be crafted. These discussions should actively involve healthcare bodies, tech companies, and legal experts.

It is also crucial to recognize that the challenges of healthcare accessibility are not exclusive to LMICs. The strategies we have proposed have the potential to benefit healthcare systems universally, and may be applicable to underserved regions of high-income countries. Lastly, before deploying ChatGPT in LMIC healthcare settings, a comprehensive strategic plan should be mutually agreed upon by governing bodies and healthcare institutions. This plan should ensure data privacy and security, establish clear ethical guidelines, and prepare for potential risks or challenges that may arise from the use of this technology. By adopting these measures, we may be better positioned to improve patient care, support healthcare providers, and work towards a more balanced and efficient healthcare ecosystem.

## Conclusion

Despite the inherent challenges arising from its training data constraints and potential inaccuracies, ChatGPT holds promise as a valuable tool in LMICs where access to quality healthcare is limited. ChatGPT stands to make promising contributions to medicine and public health by facilitating more immediate access to essential medical knowledge and education. This encompasses domains like health literacy, screening, triage, remote assistance, mental health, healthcare communication, documentation, and broader medical training. With the growing availability of function-enhancing plugins, ChatGPT's capabilities could potentially be augmented by incorporating more up-to-date and specialized medical information. Nonetheless, when contemplating its adoption in LMICs, a nuanced approach is essential, factoring in ethical considerations, cultural relevance, existing infrastructural constraints, prevalent low literacy rates, and limited digital proficiency. To optimize ChatGPT's utility in LMICs, it is crucial to institute suitable user onboarding and relevant educational initiatives. These initiatives can help set realistic expectations regarding ChatGPT's capabilities while fostering responsible and ethical use of this technology in LMICs.

## Contributors

Conceptualisation: XW, HS, ST, JC, YXW, TYW, YCT, KC.

Literature review: XW, HS, YL, ST, JC, YXW, YCT, KC.

Project administration: XW, YL.

Supervision: XW, JC, YXW, TYW, YCT, KC.

Writing—original draft: XW, HS, YL, YXW.

Writing—review & editing: XW, HS, YL, BXT, KS, AGA, YQ, ST, JC, YXW, TYW, YCT, KC.

## Declaration of interests

YQ received payment from Asia Pacific Medical Technology Association for presentation.

TYW received editorial support and medical writing from ApothecCom, study funding and article processing charges from Bayer AG, Leverkusen, Germany; funding of editorial support and medical writing from Bayer Consumer Care AG, Basel, Switzerland; study funding from Regeneron Pharmaceuticals, Inc; consulting fees from Aldropika Therapeutics, Bayer, Boehringer Ingelheim, Genetech, Iveric Bio, Novartis, Oxurion, Plano, Roche, Sanofi and Shanghai Henlius. He is also an inventor, patent-holder and a cofounder of start-up companies EyiRiS and Visre.

KC received funding from the National Institutes of Health, book royalties from Wolters Kluwer and Elsevier, and a research grant from Sonex to study carpal tunnel outcomes.

## Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

- World Health Organization. *Improving health care in developing countries using quality assurance: report of a World Health Organization pre-ISQUA joint review meeting*. Budapest, Hungary: World Health Organization; 1998:5–6.
- World Health Organization. *COVID-19 and the social determinants of health and health equity: evidence brief*. 2021.
- OpenAI. *GPT-4 technical report*; 2023. <http://arxiv.org/abs/2303.08774>. Accessed March 27, 2023.
- Eloundou T, Manning S, Mishkin P, Rock D. *GPTs are GPTs: an early look at the labor market impact potential of large language models*; 2023. <http://arxiv.org/abs/2303.10130>. Accessed March 21, 2023.
- Schalwe N, Wahl B. Artificial intelligence and the future of global health. *Lancet*. 2020;395:1579–1586.
- van de Sande D, Van Genderen ME, Smit JM, et al. Developing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter. *BMJ Health Care Inform*. 2022;29:e100495.
- Ciecierski-Holmes T, Singh R, Axt M, Brenner S, Barteit S. Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review. *NPJ Digit Med*. 2022;5:162.
- Chen L, Zaharia M, Zou J. *FrugalGPT: how to use large language models while reducing cost and improving performance*. 2023. <https://doi.org/10.48550/arXiv.2305.05176>.
- Guo J, Li B. The application of medical artificial intelligence technology in rural areas of developing countries. *Health Equity*. 2018;2:174–181.
- Das D, Kumar N, Longjam LA, et al. Assessing the capability of ChatGPT in answering first- and second-order knowledge questions on microbiology as per competency-based medical education curriculum. *Cureus*. 2023;15(3):e36034. <https://doi.org/10.7759/cureus.36034>.
- Johnson D, Goodman R, Patrinely J, et al. Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the chat-GPT model. *Research Square (preprint)*. 2023. <https://doi.org/10.21203/rs.3.rs-2566942/v1>.
- Ifikhar L, Ifikhar MF, Hanif MI. *DocGPT: impact of ChatGPT-3 on health services as a virtual doctor*. 2023.
- Hossain MM, Krishna Pillai S, Dansy SE, Bilong AA. Mr Dr. Health-assistant chatbot. *Int J Artif Intell*. 2021;8:58–73.
- Hulman A, Dollerup OL, Mortensen JF, et al. ChatGPT- versus human-generated answers to frequently asked questions about diabetes: a Turing test-inspired survey among employees of a danish diabetes center. *PLoS One*. 2023;18(8):e0290773. <https://doi.org/10.1371/journal.pone.0290773>.
- Yeo YH, Samaan JS, Ng WH, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin Mol Hepatol*. 2023;29(3):721–732. <https://doi.org/10.3350/cmh.2023.0089>.

- 16 Johnson SB, King AJ, Warner EL, Aneja S, Kann BH, Bylund CL. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. *JNCI Cancer Spectrum*. 2023;7:kad015.
- 17 Zhu L, Mou W, Chen R. Can the ChatGPT and other large language models with internet-connected database solve the questions and concerns of patient with prostate cancer?. 2023.03.06.23286827. 2023.
- 18 Abdel-Messih MS, Boulos MNK. ChatGPT in clinical toxicology. *JMIR Med Education*. 2023;9:e46876.
- 19 Kim J. Search for medical information and treatment options for musculoskeletal disorders through an artificial intelligence Chatbot: focusing on shoulder impingement syndrome. *medRxiv*. 2022. <https://doi.org/10.1101/2022.12.16.22283512>.
- 20 Howard A, Hope W, Gerada A. ChatGPT and antimicrobial advice: the end of the consulting infection doctor? *Lancet Infect Dis*. 2023;23(4):405–406. [https://doi.org/10.1016/S1473-3099\(23\)00113-5](https://doi.org/10.1016/S1473-3099(23)00113-5).
- 21 Antaki F, Touma S, Milad D, El-Khoury J, Duval R. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. *Ophthalmol Sci*. 2023;3(4):100324. <https://doi.org/10.1101/2023.01.22.23284882>.
- 22 Tu R, Ma C, Zhang C. Causal-discovery performance of ChatGPT in the context of neuropathic pain diagnosis. 2023. <https://doi.org/10.48550/arXiv.2301.13819>.
- 23 Be my eyes. <https://openai.com/customer-stories/be-my-eyes>. Accessed June 22, 2023.
- 24 Community stories. <https://www.bemyeyes.com/community-stories>. Accessed June 22, 2023.
- 25 Yadav D, Malik P, Dabas K, Singh P. Feedpal: understanding opportunities for chatbots in breastfeeding education of women in India. *Proc ACM Hum-Comput Interact*. 2019;3:1–30.
- 26 Parry SJ, Wilkinson E. Mental health services in cambodia: an overview. *BJPsych Int*. 2020;17:29–31.
- 27 Demyttenaere K, Bruffaerts R, Posada-Villa J, et al. Prevalence, severity, and unmet need for treatment of mental disorders in the world health organization world mental health surveys. *JAMA*. 2004;291:2581–2590.
- 28 Singh O. Artificial intelligence in the era of ChatGPT - opportunities and challenges in mental health care. *Indian J Psychiatry*. 2023;65:297.
- 29 World Health Organization. *Mental health atlas 2020*. 2021.
- 30 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*. 2023;183:589.
- 31 A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. *Internet Interv*. 2017;10:39–46.
- 32 Kimmarita L. Ministry to launch app, chatbot to help counter gender violence. <https://phnompenhpost.com/national/ministry-launch-app-chatbot-help-counter-gender-violence>. Accessed June 19, 2023.
- 33 Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and opinions of patients about mental health chatbots: scoping review. *J Med Internet Res*. 2021;23:e17828.
- 34 Oh K-J, Lee D, Ko B, Choi H-J. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In: *2017 18th IEEE international conference on mobile data management (MDM)*. Daejeon, South Korea: IEEE; 2017:371–375.
- 35 'He would still be here': man dies by suicide after talking with AI chatbot, widow says. <https://www.vice.com/en/article/pkadgm/mandies-by-suicide-after-talking-with-ai-chatbot-widow-says>. Accessed May 31, 2023.
- 36 Usage policies. <https://openai.com/policies/usage-policies>. Accessed May 31, 2023.
- 37 Xiang C. The amateurs jailbreaking GPT say they're preventing a closed-source AI dystopia. *Vice*; 2023. <https://www.vice.com/en/article/5d9z55/jailbreak-gpt-openai-closed-source>. Accessed May 31, 2023.
- 38 Thaver IH, Harpham T, McPake B, Garner P. Private practitioners in the slums of Karachi: what quality of care do they offer? *Soc Sci Med*. 1998;46:1441–1449.
- 39 Beracochea E, Dickson R, Freeman P, Thomason J. Case management quality assessment in rural areas of Papua New Guinea. *Trop Doct*. 1995;25:69–74.
- 40 Al-Shamsi M. Addressing the physicians' shortage in developing countries by accelerating and reforming the medical education: is it possible? *J Adv Med Educ Prof*. 2017;5:210–219.
- 41 Nori H, King N, McKinney SM, Carignan D, Horvitz E. *Capabilities of GPT-4 on medical challenge problems*; 2023. <http://arxiv.org/abs/2303.13375>. Accessed March 27, 2023.
- 42 Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS Digital Health*. 2023;2:e0000198.
- 43 Oh N, Choi G-S, Lee WY. ChatGPT goes to operating room: evaluating GPT-4 performance and future direction of surgical education and training in the era of large language models. *Ann Surg Treat Res*. 2023;104(5):269–273. <https://doi.org/10.4174/astr.2023.104.5.269>.
- 44 Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ*. 2023;9:e45312.
- 45 Bommineni VL, Bhagwagar S, Balcarcel D, Davazitkos C, Boyer D. *Performance of ChatGPT on the MCAT: the road to personalized and equitable premedical learning*. 2023.03.05.23286533. 2023.
- 46 Moons P, Van Bulck L. ChatGPT: can artificial intelligence language models be of value for cardiovascular nurses and allied health professionals. *Eur J Cardiovasc Nurs*. 2023. <https://doi.org/10.1093/eurjcn/zvad022>.
- 47 Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare*. 2023;11:887.
- 48 Casella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. *J Med Syst*. 2023;47:33.
- 49 Rao A, Kim J, Kamineni M, Pang M, Lie W, Succ MD. Evaluating ChatGPT as an adjunct for radiologic decision-making. 2023.02.02.23285399 *medRxiv*. 2023.
- 50 Balas M, Ing EB. Conversational AI models for ophthalmic diagnosis: comparison of ChatGPT and the isabel pro differential diagnosis generator. *JFO Open Ophthalmology*. 2023;1:100005.
- 51 Saluja S, Rudolphson N, Massenburg BB, Meara JG, Shrim MG. The impact of physician migration on mortality in low and middle-income countries: an economic modelling study. *BMJ Glob Health*. 2020;5:e001535.
- 52 Ting DS, Gunasekaran DV, Wickham L, Wong TY. Next generation telemedicine platforms to screen and triage. *Br J Ophthalmol*. 2020;104:299–300.
- 53 Garzon-Chavez D, Romero-Alvarez D, Bonifaz M, et al. Adapting for the COVID-19 pandemic in Ecuador, a characterization of hospital strategies and patients. *PLoS One*. 2021;16:e0251295.
- 54 Love SM, Berg WA, Podilchuk C, et al. Palpable breast lump triage by minimally trained operators in Mexico using computer-assisted diagnosis and low-cost ultrasound. *JGO*. 2018;4:1–9.
- 55 Dwivedi YK, Pandey N, Currie W, Micu A. Leveraging ChatGPT and other generative artificial intelligence (AI)-based applications in the hospitality and tourism industry: practices, challenges and research agenda. *IJCHM*. 2023. <https://doi.org/10.1108/IJCHM-05-2023-0686>.
- 56 Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. *JAMA*. 2023;329(10):842–844.
- 57 Ali SR, Dobbs TD, Hutchings HA, Whitaker IS. Using ChatGPT to write patient clinic letters. *Lancet Digital Health*. 2023;5(4):e179–e181. [https://doi.org/10.1016/S2589-7500\(23\)00048-1](https://doi.org/10.1016/S2589-7500(23)00048-1).
- 58 Patel SB, Lam K. ChatGPT: the future of discharge summaries? *Lancet Digital Health*. 2023;5:e107–e108.
- 59 Jeblik K, Schachtner B, Dext J, et al. *ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports*. 2022. <https://doi.org/10.48550/arXiv.2212.14882>.
- 60 Chan SMH, Mamat NH, Nadarajah VD. Mind your language: the importance of English language skills in an International Medical Programme (IMP). *BMC Med Educ*. 2022;22:405.
- 61 Maher J. The development of English as an international language of medicine. *Applied Linguistics*. 1986;7:206–218.
- 62 *Historical yearly trends in the usage statistics of content languages for websites*; 2023. [https://w3techs.com/technologies/history\\_overview/content\\_language/ms/y](https://w3techs.com/technologies/history_overview/content_language/ms/y). Accessed March 29, 2023.
- 63 Di Bitetti MS, Ferreras JA. Publish (in English) or perish: the effect on citation rate of using languages other than English in scientific publications. *Ambio*. 2017;46:121–127.
- 64 World bank country and lending groups – world bank data help desk. <https://datahelpdesk.worldbank.org/knowledgebase/articles/>

- 906519-world-bank-country-and-lending-groups. Accessed August 14, 2023.
- 65 List of official languages by country and territory. Wikipedia; 2023. [https://en.wikipedia.org/w/index.php?title=List\\_of\\_official\\_languages\\_by\\_country\\_and\\_territory&oldid=1169893098](https://en.wikipedia.org/w/index.php?title=List_of_official_languages_by_country_and_territory&oldid=1169893098). Accessed August 14, 2023.
- 66 Al Shamsi H, Almutairi AG, Al Mashrafi S, Al Kalbani T. Implications of language barriers for healthcare: a systematic review. *Oman Med J*. 2020;35:e122.
- 67 Jiao W, Wang W, Huang J, Wang X, Tu Z. Is ChatGPT A good translator? Yes with GPT-4 as the engine. *arXiv*. 2023. <https://doi.org/10.48550/ARXIV.2301.08745>.
- 68 Lai VD, Ngo NT, Veyseh APB, et al. *ChatGPT beyond english: towards a comprehensive evaluation of large language models in multilingual learning*; 2023. <http://arxiv.org/abs/2304.05613>. Accessed May 7, 2023.
- 69 Harrer S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *eBioMedicine*. 2023;90:104512.
- 70 Dahmen J, Kayaalp M, Ollivier M, et al. Artificial intelligence bot ChatGPT in medical research: the potential game changer as a double-edged sword. *Knee Surg Sports Traumatol Arthrosc*. 2023;31:1–3.
- 71 Shen X, Chen Z, Backes M, Zhang Y. In *ChatGPT we trust? Measuring and characterizing the reliability of ChatGPT*; 2023. <http://arxiv.org/abs/2304.08979>. Accessed May 31, 2023.
- 72 Singhal K, Tu T, Gottweis J, et al. *Towards expert-level medical question answering with large language models*; 2023. <http://arxiv.org/abs/2305.09617>. Accessed July 18, 2023.
- 73 Van Bulck L, Moons P. What if your patient switches from Dr. Google to Dr. ChatGPT? A vignette-based survey of the trustworthiness, value, and danger of ChatGPT-generated responses to health questions. *Eur J Cardiovasc Nurs*. 2023. <https://doi.org/10.1093/eurjcn/zvad038>.
- 74 Ouyang L, Wu J, Jiang X, et al. *Training language models to follow instructions with human feedback*; 2022. <http://arxiv.org/abs/2203.02155>. Accessed March 30, 2023.
- 75 Brown TB, Mann B, Ryder N, et al. *Language models are few-shot learners*; 2020. <http://arxiv.org/abs/2005.14165>. Accessed March 30, 2023.
- 76 Cao Y, Zhou L, Lee S, Cabello L, Chen M, Hershcovich D. *Assessing cross-cultural alignment between ChatGPT and human societies: an empirical study*; 2023. <http://arxiv.org/abs/2303.17466>. Accessed May 7, 2023.
- 77 Haroz EE, Bolton P, Gross A, Chan KS, Michalopoulos L, Bass J. Depression symptoms across cultures: an IRT analysis of standard depression symptoms using data from eight countries. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51:981–991.
- 78 McCradden MD, Kirsch RE. Patient wisdom should be incorporated into health AI to avoid algorithmic paternalism. *Nat Med*. 2023;29:765–766.
- 79 Alliance for Affordable Internet. *Meaningful connectivity: a new target to raise the bar for internet access*; 2020. [https://a4ai.org/wp-content/uploads/2021/02/Meaningful-Connectivity\\_Public.pdf](https://a4ai.org/wp-content/uploads/2021/02/Meaningful-Connectivity_Public.pdf).
- 80 *Satellites beat balloons in race for flying internet*. BBC News; 2021. <https://www.bbc.com/news/technology-55770141>. Accessed May 9, 2023.
- 81 How your data is used to improve model performance | OpenAI Help Center. <https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance>. Accessed May 7, 2023.
- 82 Li H, Guo D, Fan W, Xu M, Song Y. *Multi-step jailbreaking privacy attacks on ChatGPT*; 2023. <http://arxiv.org/abs/2304.05197>. Accessed May 7, 2023.
- 83 China mandates security reviews for AI services like ChatGPT - bloomberg. <https://www.bloomberg.com/news/articles/2023-04-11/china-to-mandate-security-reviews-for-new-chatgpt-like-services>. Accessed May 7, 2023.
- 84 Mgbako O, Miller EH, Santoro AF, et al. COVID-19, telemedicine, and patient empowerment in HIV care and research. *AIDS Behav*. 2020;24:1990–1993.
- 85 Cao B, Bao H, Oppong E, et al. Digital health for sexually transmitted infection and HIV services: a global scoping review. *Curr Opin Infect Dis*. 2020;33:44–50.
- 86 Tschandl P, Rinner C, Apalla Z, et al. Human–computer collaboration for skin cancer recognition. *Nat Med*. 2020;26:1229–1234.