ChatGPT's response regarding ankylosing spondylitis-related questions

Sirs,

With the rapid development of the times, AI technology, represented by ChatGPT, is increasingly being applied in the field of healthcare. One of its roles is engaging in language communication with patients, answering health-related questions, and providing medical advice and guidance.

Ankylosing spondylitis (AS) is a commonly encountered chronic inflammatory disease that primarily affects the sacroiliac joints, spine, adjacent soft tissues, and peripheral joints. It may also manifest with extra-articular features such as anterior uveitis, psoriasis, inflammatory bowel disease, and in severe cases, spinal deformities or even ankyloses, leading to a decline in the patient's quality of life (1). AS commonly occurs in young and middle-aged individuals, and due to the desire to save time or healthcare costs, many of these patients prefer seeking help online when facing relevant issues.

To assess the accuracy of ChatGPT's responses and measure the potential benefit to patients, we conducted the following study. We have selected 14 questions about AS that are most commonly asked by patients or their family members in clinical practice to ask ChatGPT (ChatGPT-3.5) (online Supplementary file). The selection of these questions is based on our clinical experience and not from direct patient interactions. Each question was asked to ChatGPT six times. While the answers provided by ChatGPT varied, their underlying meaning remained consistent. Variations primarily existed in wording or sentence structure. Therefore, we selected the first response as the final answer. Afterwards, we invited five rheumatology experts with ten to twenty-five years of experience in the field, holding positions as associate chief physicians or chief physicians, and possessing extensive clinical expertise, to assess ChatGPT's responses. The scoring range is from 1 to 10 points, where 1 indicates completely incorrect and 10 indicates entirely correct. The average score obtained was 9.38±0.66, indicating that ChatGPT's responses received approval from the experts. However, it was observed that ChatGPT's answers were not comprehensive or accurate in some instances.

For instance, in Question 6: "What tests are needed to diagnose ankylosing spondylitis?" ChatGPT's response did not mention CT scans. CT, especially high-resolution CT, can provide clearer visualisation of structural changes in the sacroiliac joints, such as erosion, sclerosis, and ankylosis, compared to x-rays (1). In clinical practice, when x-ray results are inconclusive and structural changes are suspected, or when a patient cannot undergo an MRI, a complementary CT evaluation may be necessary.

In Question 8: "What medications are used to treat ankylosing spondylitis?" ChatGPT's response was not sufficiently accurate. In clinical practice, for AS patients whose symptoms persist after NSAIDs treatment, the consideration would first shift to the use of biologic DMARDs, such as tumour necrosis factor (TNF)α inhibitors or interleukin (IL) 17 inhibitors (2). While the efficacy of traditional synthetic DMARDs for axial involvement in AS has not been confirmed, they may play a role in treating peripheral manifestations when coexisting with axial arthritis (3).

Through this study, we found that ChatGPT demonstrated a high level of accuracy in responding to questions related to AS, providing valuable assistance to patients seeking medical advice. Additionally, it helped disseminate disease-related knowledge, facilitating shared decision-making between healthcare providers and patients and improving patient compliance. However, ChatGPT also exhibited some limitations, such as occasional inaccuracies or incomplete answers to certain questions, leading to potential misinformation for patients. Further efforts are needed to enhance its accuracy and reliability.

In addition, we have noticed that apart from providing answers to the questions, at the end of each response, ChatGPT would advise patients to consult healthcare professionals or rheumatologists if they were troubled by related issues. This indicates that ChatGPT only offers answers to the questions posed by patients, and the ultimate decision-making lies with healthcare professionals, showing that doctors cannot be replaced by it at present.

In summary, as of now, ChatGPT cannot provide patients with highly accurate and comprehensive answers. However, with continuous updates and optimisations, this issue is expected to be resolved. As for whether it can offer personalised diagnosis and treatment services for patients, the answer is uncertain. Perhaps in the future, this could be achieved by adding some form of programmatic settings.

S. Li
C. Huang
X. Chen

Department of General Practice of The Second Affiliated Hospital of Fujian Medical University, Quanzhou City, Fujian Province, China.

Please address correspondence to:
Xiaoqing Chen
Department of General Practice, The Second Affiliated Hospital of Fujian Medical University, no. 950 Donghai Road, Feng ze District, Quanzhou 362000, Fujian, China.
E-mail: xiaoqingchen202301@outlook.com

Competing interests: none declared.

© Copyright Clinical and Experimental Rheumatology 2025.

References

1. Sieper J, Poddubnyy D: Axial spondyloarthritis. *Lancet* 2017; 390(10089): 73-84. https://doi.org/10.1016/S0140-6736(16)31591-4
2. van der Heijde D, Ramiro S, Landewé R *et al.*: 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* 2017; 76(6): 978-91. https://doi.org/10.1136/annrheumdis-2016-210770
3. Ward MM, Deodhar A, Gensler LS *et al.*: 2019 Update of the American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network Recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. *Arthritis Rheumatol* 2019; 71(10): 1599-613. https://doi.org/10.1002/art.41042