



# ChatGPT Responses to Commonly Asked Colon and Rectal Cancer Questions from Patient Perspective



Alyssa Habermann MD<sup>1,3</sup> Tania Gupta BS<sup>2</sup> Stephen Cohen MD<sup>2,3</sup>, Syed Shah MD<sup>2,3</sup>  
Adam Khader MD<sup>1,2,3</sup> Kathryn Fong MD<sup>2,3</sup> Michael F. Amendola MD FSVS FACS<sup>2,3</sup>

1. Virginia Commonwealth University Health System, Richmond, Virginia 2. Virginia Commonwealth University School of Medicine, Richmond, Virginia 3. Central Virginia VA Health System, Richmond, Virginia

## Background

• Patients are turning to large language models (LLMs) such as ChatGPT (Open AI, San Francisco, CA) for medical guidance, prompting concerns. ChatGPT has shown promise in answering clinical questions on topics such as cirrhosis, obstetrics and gynecology, hepatocellular carcinoma, and head/neck cancer<sup>1-3</sup>

## Aim



To assess the proficiency of ChatGPT in answering patient questions regarding the surgical management of colon and rectal cancer

## Methods

- ChatGPT 4.0 (accessed 12/16/2024) was prompted with 10 patient-based questions regarding the surgical management of colon and rectal cancer.
- A five-surgeon panel (one boarded general surgeon, two boarded colorectal surgeons, and two boarded surgical oncologists) evaluated each response for completeness, accuracy, and effective communication (5-point scale)
- ANOVA was used to compare question responses between categories and unpaired t-test was used to compare colon and rectal cancer question scores.

## LLM Answer Assessment

| Domain                         | Question   | Anchors  |
|--------------------------------|--|--|
| <b>Completeness</b>            | Does the answer above include the information that you would discuss with a patient when counseling them about this surgery? | 1. Does not include any or includes very little information<br>3. Neutral<br>5. Includes all/almost all of the information                               |
| <b>Accuracy</b>                | Does the answer above contain any inaccuracies or irrelevant content?  | 1. All or almost all of this information is inaccurate or irrelevant<br>3. Neutral<br>5. Absolutely none of this information is inaccurate or irrelevant |
| <b>Effective Communication</b> | Do you feel that your average patient could be reasonably informed about the procedure by reading the answer above?          | 1. Minimal understanding of the procedure<br>3. Neutral<br>5. As well as or better than if I described it to them myself                                 |

## Results

- The average scores (range: 3-15) of all colon and rectal questions were  $13.0 \pm 0.6$  and  $13.0 \pm 0.5$ , respectively.
- The highest scoring colon question related to postoperative length of stay ( $13.8 \pm 1.3$ ) and the lowest to operative time ( $12.2 \pm 1.9$ ).
- The highest scoring rectal question discussed bowel continence and return to normal activities ( $13.6 \pm 1.3$  and  $13.6 \pm 1.5$ ) and the lowest discussed surgical procedures for removing rectal cancer ( $11.8 \pm 2.3$ ).
- Average scores for completeness, accuracy, and communication were  $4.6 \pm 0.7$ ,  $4.3 \pm 0.7$ , and  $4.2 \pm 0.7$ , respectively ( $p < 0.001$ ).
- There was no significant difference between evaluator scores for colon and rectal questions ( $p > 0.05$ ).

The information contained in this poster is not intended to provide interpretation of Veterans Affairs policy nor specific details about how individual Veterans Affairs Medical Centers operate services within their jurisdiction. The contents do not represent the views of the U.S. Department of Veterans Affairs nor the United States Government.

| Patient-Based Questions  | Colon (mean $\pm$ SD) | Rectal (mean $\pm$ SD) |
|--|-----------------------|------------------------|
| 1. My doctor says I need surgery to remove my cancer. Can you explain the procedure to me? | $12.4 \pm 1.7$        | $11.8 \pm 2.3$         |
| 2. Are there different methods of doing this surgery?                                      | $12.4 \pm 2.2$        | $12.4 \pm 2.2$         |
| 3. What are the risks of this surgery?   | $13.4 \pm 1.1$        | $13.0 \pm 1.2$         |
| 4. Will I need an ostomy?  | $12.8 \pm 1.9$        | $13.0 \pm 1.2$         |
| 5. How long will the surgery take?   | $12.2 \pm 1.9$        | $12.8 \pm 0.8$         |
| 6. How long will I be in the hospital after surgery?                                       | $13.8 \pm 1.3$        | $13.4 \pm 1.1$         |
| 7. Will I have normal bowel movements after surgery?                                       | $12.8 \pm 1.5$        | $13.0 \pm 1.4$         |
| 8. Will I have issues with bowel continence after surgery?                                 | $13.6 \pm 1.3$        | $13.6 \pm 1.3$         |
| 9. How much pain will I be in after surgery?   | $13.4 \pm 2.6$        | $13.4 \pm 2.6$         |
| 10. When can I return to normal activities after surgery?                                  | $13.4 \pm 1.5$        | $13.6 \pm 1.5$         |

## Conclusions

- ChatGPT 4.0 is capable of answering commonly asked questions regarding colon/rectal cancer with reasonable completeness, accuracy, and effective communication with no statistically significant differences in these categories.
- As AI model architecture continues to develop, further advances in patient-based automated responses are likely to occur, leading to more effective communication with patients.

1. Yeo YH, Samaan JS, Ng WH, Ting PS, Trivedi H, Vipani A, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin Mol Hepatol.* 2023;29(3):721-732. doi:10.3350/cmh.2023.0089

2. Grünebaum A, Chervenak J, Pollet SL, Katz A, Chervenak FA. The exciting potential for ChatGPT in obstetrics and gynecology. *Am J Obstet Gynecol.* 2023;228(6):696-705. doi:10.1016/j.ajog.2023.03.009

3. Kuşcu O, Pamuk AE, Sütay Süslü N, Hosal S. Is ChatGPT accurate and reliable in answering questions regarding head and neck cancer? *Front Oncol.* 2023;13:1256459. doi:10.3389/fonc.2023.1256459