Reading capsule endoscopy: Why not AI alone?

OPEN ACCESS

(cc) (i) (S) (=)

Authors
Maren Haslach-Häfner[1], Klaus Mönkemüller[2]

Institutions
1   HELIOS Frankenwaldklinik Kronach, Kronach, Germany
2   Gastroenterology, Virginia Tech Carilion Research Institute, Roanoke, United States

Corresponding author
Prof. Klaus Mönkemüller, Virginia Tech Carilion Research Institute, Gastroenterology, Roanoke, United States
moenkemueller@yahoo.com

Artificial intelligence (AI) has the potential to facilitate the work of the practicing endoscopist and improve patient care in many aspects, including diminishing work time, detecting and preselecting lesions for review by the endoscopist, and by making a specific diagnosis [1]. Because reading results from a capsule endoscopy (CE) study is quite time-consuming, with times ranging between 30 and 120 minutes, AI may be a solution for improving the efficiency of CE by removing redundant images and identifying suspicious abnormalities [2]. In addition, AI may be capable of establishing a clear-cut diagnosis. Despite promising data showing that AI has excellent sensitivity and specificity to find abnormal pathology, there are still no real-world clinical studies evaluating whether AI can diminish reading times [2, 3].

Interestingly, there are now commercial CE systems with AI technology that includes redundancy deletion, lesion detection, and classification software. The OMOM HD CE system (Jinshan Science & Technology Group, Yubei, China) introduced in 2020 has AI technology that includes reporting software with a convolutional neural network-based computer aided detection (CADe) algorithm that can identify small bowel abnormalities and filter the video files for them [4]. The system also allows for operator-based or AI-assisted reading, the latter displaying only the filtered images selected by the CADe algorithm. Moreover, the filtered images are played in a video format familiar to experienced CE readers, or as a collection of still images with suggested findings [4].

O'Hara et al from Ireland present the first real-world clinical study looking into whether AI-assisted CE can really diminish reading time [4]. In this single-center, retrospective study, 40 patient studies performed using the OMOM capsule were analyzed first with standard reading (SR) methods and later using AI-assisted reading [3]. The aims were to compare reading time, pathology identified, intestinal landmark identification, and assessment of bowel preparation. The authors found that overall diagnosis ("per-patient") correlated 100% between the two reading methods. However, in a "per-lesion" analysis of 1293 images of significant lesions AI-assisted was much better than SR (98.1% versus 86.2%). Nonetheless, the most important finding of the study was that the mean reading time went from 29.7 minutes with SR to 2.3 minutes with AI-assisted reading. Also, AI and SR had a high concordance in quantifying the quality of small bowel preparation. Finally, SR clearly surpassed AI in one aspect: AI failed to clearly identify landmarks, especially the cecum [3].

The authors are to be congratulated for performing the first real-world study using a commercially available CE system with AI capabilities. The investigators put great effort into designing the study, setting the best possible gold standard for comparing both the SR and AI readings, and trying to eliminate subjective interpretations.

What are the main take-home messages of this study?

First, this can be considered a landmark study demonstrating that AI can reduce reading times. AI decreased the reading time by a median of 30 minutes. This has implications for clinical practice, as it enhances the productivity of endoscopists. Until now, there were concerns that AI might even increase reading times by selecting too many images for review. Indeed, in this study, AI had a very high false-positive rate, with only 8% of selected images having significant pathology. What we can

deduce is that the efficiency of CE became better with interaction by AI and human experience combined with intelligence.

Second, when comparing selected pathological images, AI clearly beat the human eye. However, when push comes to shove, i.e. looking at the diagnostic ability on an individual patient, both AI and human reading were 100% concordant. Indeed, the most important aspect of reading CE is the per-patient diagnosis, because it does not matter if one or 100 images showed the pathology, as long as the human eye or AI finds the lesion and reaches a diagnosis.

However, two key aspects that negatively surprised us from this AI system were its inability to properly determine key anatomic landmarks such as the cecum and its inability to properly determine the quality of bowel preparation. Both are essential elements for quality endoscopic metrics. Also, why could AI not clearly identify landmarks such as the cecum? This intriguing finding deserves further study. It is possible that most AI systems have been trained to analyze the small bowel and pathologies of the small bowel, and that less interest has been given to teaching AI how to recognize colon structures. Based on the results of this study, many more images of the cecum and colon should be utilized in CADe algorithms. Lastly, AI missed significant lesions such as angiodysplasia, circumferential cecal ulcers, lymphangiectasias, and clips placed during antegrade enteroscopy, as shown in Table 4 of the article. A possible explanation was poor bowel preparation. Nevertheless, we believe that AI systems should be trained with a wide spectrum of images and videos with a various types of bowel preparation.

There are some potential deficits of this study that merit discussion. First, its retrospective design makes it prone to the usual biases and inaccuracies of retrospective analyses. Nevertheless, endoscopic databases are quite objective and can be used for objective research. However, we wonder how the AI system compared with the original reports. Was any new diagnosis reached or discarded during the reanalysis of the capsule endoscopic readings? Was there any interobserver variability between the original SRs and SRs done for study purposes? Did the authors use any AI during the clinical part of the study, i.e. during the pre-retrospective data interpretation? Also, what is the intra-observer and interobserver reliability of AI? Did the study using AI and repeated SR change any of the original diagnosis or uncover new lesions? We raise these questions with the aim of assisting future investigators doing similar studies and making it easier for the practicing endoscopist to get a grasp on the potential uses of AI in clinical practice.

In summary, it is evident that AI decreased CE reading times significantly but did not surpass an experienced endoscopist in reaching a final diagnosis on an individual patient. This study shows that AI is here to complement and assist the practicing endoscopist in becoming more efficient at reading CE. At present, AI appears to have a supporting effect, but it still has significant shortcomings, and thus, cannot replace an experienced human examiner. Furthermore, AI is unlikely to be a substitute for the human eye, because humans do ultimately hold responsibility for the final endoscopic report.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Soffer S, Klang E, Shimon O et al. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. Gastrointest Endosc 2020; 92: 831–839 doi:10.1016/j.gie.2020.04.039

[2] Rondonotti E, Pennazio M, Toth E et al. How to read small bowel capsule endoscopy: a practical guide for everyday use. Endosc Int Open 2020; 8: E1220–E1224 doi:10.1055/a-1210-4830

[3] Messmann H, Bisschops R, Antonelli G et al. Expected value of artificial intelligence in gastrointestinal endoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy 2022; 54: 1211–1231 doi:10.1055/a-1950-5694

[4] O'Hara Fintan John et al. Capsule endoscopy with artificial intelligence-assisted technology: Real-world usage of a validated AI model for capsule image review. Endoscop Int Open 2023; 11: E970–E975 doi:10.1055/a-2161-1816