Artificial Intelligence and the Promise of Expediting and Standardizing the Reviewing and Reporting of Capsule Endoscopy Videos

Michael Sey, MD, MPH

The invention of video capsule endoscopy at the turn of the millennium revolutionized imaging of the small bowel.¹ What was once considered an inaccessible portion of the gastrointestinal tract could now be viewed endoscopically in its entirety. Despite the advances made during the past 2 decades, capsule endoscopy remains hindered by 2 major factors: the highly subjective nature of capsule reading and the amount of time required to review a video. The first problem results from a lack of standardization in describing findings in the small bowel, whereas the latter is due to the amount of time required for the capsule to transit through the small intestine, which is often many hours. As a result, 2 capsule endoscopists reading the same video may offer differing reports, despite each spending considerable time reviewing the study. The solution may lie in artificial intelligence (AI), given its ability to perform time-consuming, repetitive tasks quickly in a standardized fashion.²

Xie et al³ report their findings on the use of AI-assisted small bowel capsule endoscopy. The investigators used a large registry of small bowel capsule endoscopy videos from 51 Chinese medical centers between 2012 and 2020 to form 2 groups: a training data set (2927 videos) and a validation data set (2898 videos). The former was used to develop a convolutional neural network algorithm to detect small bowel abnormalities and the latter was used to validate the algorithm. In the validation stage, videos were read in a conventional manner by a capsule endoscopist, followed by AI-assisted review some months later. When the findings from the conventional review and AI-assisted review were discordant, adjudication by a separate group of expert readers was used.

There were a total of 6084 findings among the 2898 videos in the validation cohort; conventional reading detected 4630 abnormalities (76.1%), whereas AI-assisted review detected 5834 abnormalities (95.9%). Most discordant videos were the result of a normal designation on conventional review, whereas AI assistance detected an abnormality (n = 278) rather than the other way around (n = 28). Given the higher detection rate in the AI-assisted group, this finding would suggest that the experts tended to agree with the AI. Not surprisingly, the AI-assisted group completed their review much faster than the conventional reading group (median time, 5.4 minutes vs 51.4 minutes; \( P < .001 \)).

Perhaps the greatest barrier to the adoption of AI in capsule endoscopy is the fear of missing a lesion without human review of the entire video. Xie et al³ specifically addressed this scenario in their study and reported a very low miss rate of 1.0% using the AI algorithm when compared with the conventional reading group. In contrast, the conventional reading group had a miss rate of 9.6% when compared with the AI-assisted group. Thus, although the AI algorithm is not perfect, neither are human readers.

An AI algorithm is only as good as its programming. An algorithm trained to identify an apple as an orange will do so consistently every time. In this case, the algorithm was trained to detect abnormalities in the small bowel using capsule endoscopy structured terminology (CEST).⁴ Although the desire to use CEST as a means to standardize the reporting of small bowel findings is desirable, it falls short in several ways. For one, CEST was developed nearly 2 decades ago, only 5 years after the invention of capsule endoscopy, and as such, some would consider it outdated. Perhaps more importantly, CEST provides a framework only for the vocabulary that can be used to report capsule endoscopy findings, not definitions. Recognizing this shortcoming, investigators have since worked toward standardizing the nomenclature and definitions for small bowel findings using a Delphi process.

Open Access. This is an open access article distributed under the terms of the CC-BY License.
process, with consensus reached among a group of international experts with a high degree of agreement.\textsuperscript{5,6} Thus, it would be important to validate the algorithm against the new consensus definitions and retrain the algorithm if necessary. Furthermore, because the algorithm was derived in a single country and validated at a single center, it would be important to externally validate the algorithm with a more geographically diverse group of patients and capsule endoscopists. Ultimately, if the algorithm performs well in these scenarios, then AI-assisted review may become the new standard for capsule reporting.

**ARTICLE INFORMATION**

Published: July 14, 2022. doi:10.1001/jamanetworkopen.2022.22006

Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2022 Sey M. JAMA Network Open.

**Corresponding Author:** Michael Sey, MD, MPH, Gastroenterology and Therapeutic Endoscopy, Small Bowel Endoscopy Program, London Health Sciences Centre-Victoria Hospital, 800 Commissioners Rd E, London, ON N6A 5W9, Canada (msey2@uwo.ca).

**Author Affiliations:** Schulich School of Medicine, Western University, London, Ontario, Canada; Gastroenterology and Therapeutic Endoscopy, London Health Sciences Centre, London, Ontario, Canada; Lawson Health Research Institute, London Health Sciences Centre, London, Ontario, Canada.

**Conflict of Interest Disclosures:** Dr Sey reported serving as a consultant for Medtronic; receiving educational grants from Cook Medical; and serving as a speaker for and receiving research grants from Pendopharm. No other disclosures were reported.

**REFERENCES**


