Detecting colorectal polyps with use of artificial intelligence

Yuichi Mori, Shin-ei Kudo, Masashi Misawa

Digestive Disease Center, Showa University Northern Yokohama Hospital, Yokohama, Japan

Correspondence to: Yuichi Mori, MD, PhD. Digestive Disease Center, Showa University Northern Yokohama Hospital, 35-1 Chigasaki-chuo, Tsuzuki, Yokohama 224-8503, Japan. Email: ibusiginjp@gmail.com.

Comment on: Urban G, Tripathi P, Alkayali T, *et al.* Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. Gastroenterology 2018;155:1069-78.e8.

Received: 06 May 2019; Accepted: 20 May 2019; Published: 24 May 2019.

doi: 10.21037/jmai.2019.05.01

View this article at: http://dx.doi.org/10.21037/jmai.2019.05.01

Colorectal cancer (CRC) is a major cause of cancer-related mortality in most countries. Colonoscopy during which all neoplastic and pre-malignant polyps (e.g., adenomas) are eradicated is considered beneficial in decreasing the incidence of CRCs and their associated mortality (1,2). This concept has been supported by several large-scale prospective studies (3). The quality of the colonoscopy procedure, however, varies according to the expertise of the endoscopist. Adenoma detection rates (ADRs) during colonoscopy, which are inversely associated with the incidence of "interval CRC" (defined as cancer diagnosed between the screening and post-screening surveillance examinations) are reportedly 7–53% (4). Thus, many adenomas are missed during routine colonoscopy (5), posing a major barrier to the standardization of high-quality colonoscopy.

Artificial intelligence (AI) is currently attracting increasing attention as a means to address this issue. It could help bypass the problem of not recognizing adenomas during colonoscopy by informing the endoscopist of the presence and location of the polyp(s) in a real-time fashion (6-9). Several experimental studies focusing on automated polyp detection have been reported in engineering journals (10,11), but introducing AI technology into the field of colonoscopy has been thought to require more time because of both technological issues and limited evidence of its possibilities (12).

Urban *et al.* presented a clear solution to this situation. In the present commentary, we address the strength and limitation of their study from the clinical perspective, rather than an engineering perspective. They conducted a retrospective study using a much larger sample than in previous reports and showed excellent results supporting automated polyp detection (13). They prepared four datasets: [A] 8,641 colonoscopic images containing 4,088 images with polyps and 4,553 without polyps; [B] 1,330 colonoscopic images containing 672 images with polyps and 658 without polyps; [B] 9 colonoscopic videos containing 13,292 frames with visible polyps and 31,655 without visible polyps; [D] 11 challenging colonoscopic videos in which all the visible polyps were deliberately overlooked during withdrawal of the endoscope to mimic a missed-polyp scenario.

Using these four datasets, they conducted multiple experiments using the developed AI program whose algorithm was based on a convolutional neural network: (I) cross-validation on dataset [A]; (II) training on dataset [A] and testing on dataset [B]; (III) training on dataset [A] and testing on datasets [C] and [D]; (IV) training on datasets [A] and [C] and testing on dataset [D]. According to their results, experiment (I) (the cross-validation assessment) provided roughly 96% accuracy and an area under the curve (AUC) of 0.99. Experiment (II) provided a test accuracy of 96.4% with an AUC of 0.974. Experiment (III) provided sensitivity of 0.93 and specificity of 0.93 for dataset [C] and sensitivity of 0.93 at a frame-by-frame false-positive rate of 7% on dataset [D]. Experiment (IV) provided 0.92 sensitivity at a frame-by-frame false-positive rate of 5%.

These results appeared outstanding because sensitivity >90% was secured with a relatively low false-positive rate—considered an ideal, smooth performance in the clinical setting. This excellent ability of AI to detect polyps was also reported in another, recently published paper (14). Another interesting point of the present study is that the AI model worked better in terms of reduction of false-positive rate when video-based image frames were added to the learning material [experiment (III) versus experiment (4)]. So far, most of the similar researches have adopted only static images as learning material because acquisition and annotation of static endoscopic images are much easier and faster than using video-based image frames (6), however the present study suggests the potential benefit of using video-recordings as learning material for automated polyp detection.

A major strength of AI systems is that they are not susceptible to inter-observer and intra-observer variability and therefore could improve ADRs regardless of the endoscopist's experience and expertise. From this point of view, the AI developed by Urban *et al.* is likely to play a significant role in colonoscopy practice. The polyp "miss rate" is currently thought to be approximately 20% (5), which could be avoided with a highly sensitive means of detection such as the proposed AI system with its >90% sensitivity for polyp recognition. On the other hand, polyps which are invisible due to the structural issues of colon (e.g., behind folds, beneath cloudy fluid) cannot be detected even with such a smart AI system, which is another challenging issue to achieve high-quality colonoscopy.

Nevertheless, we must bear in mind several weak points of the Urban *et al.* study. The most important issue is that their study was not prospective, nor did it use AI in real time. Instead, it had a retrospective design and used unaltered images/videos. Retrospective studies always contain selection bias to some extent, which possibly shades the results in favor of AI—although Urban *et al.* attempted to avoid this problem in their study. As the authors pointed out, prospective trials conducted in multiple facilities are mandatory to prove the efficacy of AI in real clinical practice. So far, two prospective, single-center studies have been published in this area (15,16), both of which showed favorable results, suggesting the potential for AI to increase the ADR.

Based on the data of other prospective studies and the present article by Urban *et al.*, we might conclude that AI could be an indispensable modality with which to improve the quality of colonoscopy and probably decrease the number of interval CRCs. Before introducing this technology into widespread colonoscopy practice, however, we must wait for more robust validation via multicenter prospective studies conducted in facilities where AI is not developed (i.e., completely external validation). Using such prospective studies, we could simultaneously evaluate AI's efficacy (e.g., increments of ADRs, suppression of interval CRCs) and drawbacks (e.g., prolonged procedure time (17), potential adverse events caused by the endoscopist's distracted concentration, deskilling endoscopists' ability). When we obtain positive results from such strictly designed trials, AI will be recognized as a "game changer" of colonoscopy practice.

Acknowledgments

We thank Nancy Schatken, BS, MT (ASCP), from Edanz Group (www.edanzediting.com/ac), for editing a draft of this manuscript.

Funding: This work is supported by the Japan Society for the Promotion of Science (No. 19K08403).

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Journal of Medical Artificial Intelligence*. The article did not undergo external peer review.

Conflict of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jmai.2019.05.01). Y Mori, S Kudo, and M Misawa have received speaking honoraria from Olympus Corporation.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. N Engl J Med 2012;366:687-96.
2. Winawer SJ, Zauber AG, Ho MN, et al. Prevention

of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993;329:1977-81.
3. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. N Engl J Med 2013;369:1095-105.
4. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014;370:1298-306.
5. van Rijn JC, Reitsma JB, Stoker J, et al. Polyp miss rate determined by tandem colonoscopy: a systematic review. Am J Gastroenterol 2006;101:343-50.
6. Mori Y, Kudo SE, Berzin TM, et al. Computer-aided diagnosis for colonoscopy. Endoscopy 2017;49:813-9.
7. Ahmad OF, Soares AS, Mazomenos E, et al. Artificial intelligence and computer-aided diagnosis in colonoscopy: current evidence and future directions. Lancet Gastroenterol Hepatol 2019;4:71-80.
8. Kudo SE, Mori Y, Misawa M, et al. Artificial intelligence and colonoscopy: Current status and future perspectives. Dig Endosc 2019. [Epub ahead of print].
9. Vinsard DG, Mori Y, Misawa M, et al. Quality Assurance of Computer-Aided Detection and Diagnosis in Colonoscopy. Gastrointest Endosc 2019. [Epub ahead of print].
10. Tajbakhsh N, Gurudu SR, Liang J. Automatic polyp detection using global geometric constraints and local intensity variation patterns. Med Image Comput Comput Assist Interv 2014;17:179-87.
11. Bernal J, Tajkbaksh N, Sanchez FJ, et al. Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results From the MICCAI 2015 Endoscopic Vision Challenge. IEEE Trans Med Imaging 2017;36:1231-49.
12. Misawa M, Kudo SE, Mori Y, et al. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. Gastroenterology 2018;154:2027-9.e3.
13. Urban G, Tripathi P, Alkayali T, et al. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. Gastroenterology 2018;155:1069-78.e8.
14. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nature Biomedical Engineering 2018;2:741-8.
15. Klare P, Sander C, Prinzen M, et al. Automated polyp detection in the colorectum: a prospective study (with videos). Gastrointest Endosc 2019;89:576-82.e1.
16. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut 2019. [Epub ahead of print].
17. Mori Y, Kudo SE, Misawa M, et al. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. Ann Intern Med 2018;169:357-66.

doi: 10.21037/jmai.2019.05.01
Cite this article as: Mori Y, Kudo S, Misawa M. Detecting colorectal polyps with use of artificial intelligence. J Med Artif Intell 2019;2:11.

