Received: 28 November 2020 | Accepted: 28 February 2021

DOI: 10.1002/ueg2.12108

REVIEW ARTICLE

UNITED EUROPEAN GASTROENTEROLOGY ueg journal WILEY

Current status and limitations of artificial intelligence in colonoscopy

Alexander Hann | Joel Troya | Daniel Fitting

Department of Internal Medicine II, Interventional and Experimental Endoscopy (InExEn), University Hospital Wuerzburg, Würzburg, Germany

Correspondence
Alexander Hann, Medizinische Klinik und Poliklinik II, Uniklinik Würzburg, Oberdürrbacher Straße 6, Würzburg 97080, Germany.
Email: hann_a@ukw.de

Abstract

Background: Artificial intelligence (AI) using deep learning methods for polyp detection (CADe) and characterization (CADx) is on the verge of clinical application. CADe already implied its potential use in randomized controlled trials. Further efforts are needed to take CADx to the next level of development.

Aim: This work aims to give an overview of the current status of AI in colonoscopy, without going into too much technical detail.

Methods: A literature search to identify important studies exploring the use of AI in colonoscopy was performed.

Results: This review focuses on AI performance in screening colonoscopy summarizing the first prospective trials for CADe, the state of research in CADx as well as current limitations of those systems and legal issues.

KEYWORDS
colonic polyps, colonoscopy, colorectal neoplasms, computer-assisted, deep learning, diagnosis, endoscopy, gastrointestinal

INTRODUCTION

The advance of artificial intelligence (AI) using deep learning (DL) methods opened a new scope in medical imaging analysis. Regarding possible applications in gastroenterology prevention of colorectal cancer (CRC) by screening colonoscopy has always been a main field of activity and research for endoscopists. An increase of adenoma detection rate (ADR) is expected to lead to a decrease of cancer risk but simultaneously burdens workload, costs, and risk of additional polypectomy to the patient.[1]

The use of computer-aided detection (CADe) and characterization (CADx) systems promise both augmented performance by increased ADR and higher efficiency by firm identification of non-adenomatous polyps. Thus, for example, hyperplastic polyps in the rectosigmoid can be left unresected. Yet physicians' sentiment toward AI is rather ambiguous. Strong interest and concerns about unnecessary workload, development of operator deskilling, and over-reliance on AI are facing one another.[2] Several commercial CAD systems have entered the market already (Table 1). The ESGE states that incorporation of these systems in clinical routine is possible as soon as high-quality multicenter studies demonstrated acceptable accuracy.[3] This review focuses on AI performance in screening colonoscopy. It provides an overview of the first prospective trials in CADe, the state of research in CADx as well as current limitations of those systems and legal issues.

POLYP DETECTION (CADe)

The main contribution of screening colonoscopy to CRC prevention is the ability to detect precancerous lesions and the possibility to remove them. The most intensively studied AI that was built using DL

TABLE 1 Overview of commercial CAD systems

Name	EndoBrain	GI Genius	Discovery	REiLI CAD EYE	ai4gi	EndoScreener
Company	Cybernet Systems Co.	Medtronic	Pentax Medical	Fujifilm	Imagia, Satis	Shanghai Wision AI Co
Integration	CF-H290ECI, Olympus Corp.	Different vendors possible	Pentax colonoscopes	Fujifilm's 700 series colonoscopes	Different vendors possible	Different vendors possible
Dual monitor in study	No	No	Yes	No	No	Yes
Study data	Mori Y et al. (2016), Misawa M et al. (2016)	Repici A et al. (2020)	n/a	Non	Byrne M et al. (2017)	Wang P et al. (2020)
Mode	CADe and CADx	CADe	CADe	CADe and CADx	CADe and CADx	CADe
Regulatory	Approved by PMDA 12/2018	CE-approved 07/2019	CE-approved 01/2020	CE-approved 02/2020	No	No

TABLE 2 Qualitative overview of AI capability for polyp detection (CADe)

		Polyp size (mm)			Location		Appearance		
	Adenoma detection rate improvement	<6	6–9	>9	Proximal	Distal	Flat	sessile	Pedunculated
AI benefit	+++	+++	+		++	+++	+++	+++	

Note: Increase in plus signs resembles higher number of studies supporting this feature.
Abbreviation: AI, artificial intelligence.

and that bridged the gap from an experimental tool tested on retrospectively collected colonoscopy videos to the clinical practice is the image-based polyp detection. Until now, six prospective randomized trials analyze the performance of such systems in clinical practice[4–9] (Table S1).

Five of those randomized controlled trials have ADR as the primary endpoint. One tandem trial analyses the adenoma miss rate (AMR) as a primary endpoint. The most important finding is that all trials powered for ADR comparison present a significant increase in ADR. Pooling the data from those trials present an increase in ADR of 6–15.2 percentage points (median 12.4), depending on the experience of the examiners and the inclusion criteria (Table 2). The range of ADR using AI rose in one multicenter study up to 54.8% in the CADe group compared to 40.4% in the control group.[4] Thus, extrapolating those results to the studies that correlated the ADR to the rate of interval cancer[10] and death due to interval cancer[1] would suggest that there is a huge benefit if this new technology is applied in clinical practice. A more detailed analysis of the randomized trials with ADR as a primary endpoint reveals that all report on a significant increase in detection of diminutive adenomas of 5 mm and lower.[11] Only one study reported a significant increase in the detection of adenomas of 6–9 mm.[4] None of those studies present an increase in the detection of larger adenoma.

Interestingly the withdrawal time without the duration of polyp removal did not significantly differ between examinations with CADe support or without in any of the studies.

Taken together, the trend toward an increase in ADR[12] will go on and will be pushed even further due to the introduction of CADe.

QUALITY CONTROL OF ENDOSCOPY (QC)

Key performance measures for lower gastrointestinal endoscopy aim to improve quality.[13] AI might help to objectively record key performance measures and thus contribute to an improved quality of the procedure.[10] One of the most important quality measures is the withdrawal time. It corresponds with the ADR[14] and with death due to interval cancer.[1] The recommended time for withdrawal is six minutes or more.[13,15] However, there is no recommendation on the time the examiner should spend per colon segment. In practice, some examiners move quicker in the ascending colon and then slow down after realizing that the withdrawal time will be too short resulting in a more thorough examination of the distal colon. The withdrawal speed and time spent per optically covered surface might be a more reliable measurement tool. In the study of Gong D et al., the authors could correlate similarity scores of images to each other and relate them to withdrawal speed.[16] Setting thresholds for withdrawal speed in this randomized trial resulted in a higher ADR of 16% compared to the control group with 8% obtained in a non-screening population. An additional feature of the system was to recognize endoscope slipping during withdrawal. To help resume colonoscopy at the point the endoscope slipped, it presented the last ten frames before slipping at the bottom of the screen.

Another important quality measurement is the adequacy of bowel cleansing. The Boston Bowel Preparation Scale (BBPS) became the most commonly used tool for this estimation.[17] The scale is applied during withdrawal after cleansing maneuvers including washing of mucosa and suction of residual fluids have been applied. Poor bowel preparation is associated with a reduced ADR.[18,19]

Thus, different European guidelines recommend a BBPS value of 6 and more as adequate with none of the three segments being below 2.[13,20,21] One disadvantage of this scale is that the examiners recall the scoring impression from their memory. A more detailed and precise subdivision might better reflect reality. Zhou J et al. trained an AI to estimate the BBPS every 30 s and thus present an overview of the bowel preparation.[22] A further step forward was performed by Thakkar et al. who developed an algorithm that displays quality of examination metrics comprising total percentage of visualized colonic surface area, colon preparation, clarity of the endoscopic view, and colon distension in real-time to the endoscopist.[23] After this proof-of-concept study, further efforts are needed to quantify the contribution of such AIs to colonoscopy performance. Additionally, the potential of distraction from the endoscopic image introduced by such digital assistants needs to be evaluated.

POLYP CHARACTERIZATION (CADx)

The vast majority of polyps that are detected during screening colonoscopy are sized less than 5 mm. Those diminutive polyps rarely progress to advanced adenomatous histology.[24,25] Therefore removal and histopathological examination burden costs and workload to the health systems.[26]

Optical biopsy means distinguishing adenomatous and non-adenomatous lesions by real-time assessment of polyps using advanced endoscopy imaging techniques. Many classifications were introduced to distinguish adenoma's different grades of dysplasia or sessile serrated adenoma (SSA) from hyperplastic tissue. However, their consequent implementation in the clinical routine is rare.[27–30] Identification of hyperplastic polyps enables "resect-and-discard" and "diagnose and leave" strategies resulting in shorter procedure time, fewer adverse events, and lower costs.[26] To implement these strategies, quality directives of the American Society for Gastrointestinal Endoscopy PIVI have been determined for diminutive polyps in the rectosigmoid. For resect and discard polyps the threshold of ≥90% agreement in the assignment of post-polypectomy surveillance compared to histopathology has to be achieved. Leaving polyps requires a negative predictive value (NPV) ≥90% for adenomatous histology with high confidence.[31] These thresholds are not reached outside expert centers.[32–34] This leads to dissent of international (ASGE/ESGE) and national guidelines.[20,35]

CADx as a virtual assistant can overcome interobserver variability.[36] CADx systems do not rely on optical clinician-based classifications. They rely on histological findings although this is also prone to interobserver variability. Especially, pathologists often do not agree regarding the differentiation between SAA and hyperplastic polyps.

To evaluate practical approaches, we focus in this review on WLE and so-called "push the button technologies" (Table S2).

Renner et al. presented in a prospective study evaluating 100 images of polyps, an accuracy of 78% in differentiating between adenomatous and nonadenomatous histology. Predominantly diminutive polyps were analyzed, and the NPV was 88.2%.[37] Chen et al. added full magnification imaging modality and achieved a substantially higher specificity in a set of 384 diminutive polyps resulting in an NPV of 92.6%[36] in their validation set. The performance of the convolutional neural network (CNN) was equal to expert colonoscopists but superior to novices.

Using a training data set of >5000 retrospectively collected images, Zachariah et al. could reach an accuracy of 93.6% with an NPV of 92.6% with their CNN. There was no significant difference in using narrow-band-imaging or HD-WLE images for analysis.[38]

To implement CADx systems in clinical routine, real-time characterization of polyps detected during colonoscopy is a prerequisite. One CNN developed by Komeda et al. showed an accuracy of 75.1% in the characterization of 180 polyp images in real-time.[39] Byrne et al. validated their system with 125 videos[40] and achieved a high accuracy of 94%. However, 15% of the videos could not be characterized with high confidence.

In conclusion, the feasibility of CADx systems and thereby improvement of optical biopsy with accuracy >90% have been shown (Table 3). Confirmation of those results have to be shown in real life settings.

LIMITATIONS

Although there have been great advances of AI in GI-endoscopy, its contribution to improved CRC surveillance is still uncertain. Limitations include that existing data of CAD systems has exclusively been collected using in-hospital colonoscopy settings, usually resembling a mixture of inpatient and outpatient colonoscopy procedures. All randomized prospective studies for polyp detection except one were done in China. A comparison of the ADR in those studies performed in China with the one study done in Italy reveals a marked difference in baseline ADR.[4–9] A reason for this difference might have been different expertise of the participating endoscopists. Furthermore,

TABLE 3 Qualitative overview of AI capability for polyp characterization (CADx)

	Adenoma/Non-adenoma (mm)			SSA/Hyperplastic	Invasiveness depth CRC		
Differentiation	<6	6–9	>9	Any size	TA	HGD/T1	>=T2
AI benefit	+++	+	+	n/a	+		+

Note: Increase in plus signs resembles higher number of studies supporting this feature.

Abbreviations: AI, artificial intelligence; CRC, colorectal cancer; HGD, high-grade dysplasia; TA, tubular adenoma.

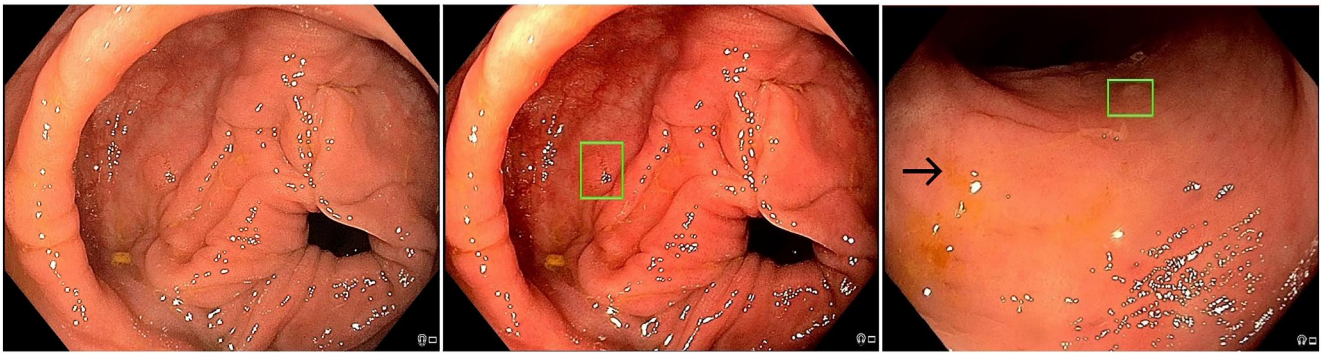

FIGURE 1 Example of images obtained using a commercial CADe system. Left image: Recorded raw video signal presenting a polyp that was undetected by the endoscopist and highlighted in the middle image by the CADe. Right image: Distraction of the endoscopist by the CADe due to the highlighting of a small mucosal protrusion instead of the arrow marked polyp

the studies performed in China were all single-center studies. These limitations regarding the studies' populations prevent the extrapolation of the current data to the outpatient CRC-surveillance cohorts in different regions. Multicenter studies in outpatient settings have to confirm the significant improvements in ADR.

Moreover, improvement of ADR in colonoscopy mainly by the identification of diminutive adenomas should be critically scrutinized to what extent it adds clinical benefit to the patient. CADe systems report an increase of ADR but no increase was shown for detection of advanced adenoma. One reason might be that the studies were not powered to detect an increase in recognition of advanced adenomas.[41] High-grade dysplasia is found in less than 1% of diminutive adenomas, defined as smaller than 6 mm.[42] Their natural history is not as clearly associated with interval cancer as in the case of adenomas 1 cm and greater in size.[42–44] The well-established link between improved ADR and decrease of interval cancer could be caused by enhanced thoroughness of inspection or enhanced mucosal visualization. CADe can only contribute to the first point. Thus, future studies are necessary to reveal if the usage of AI for polyp detection is associated with increased survival due to a lower rate of interval cancer.

The sole use of CADe system has the potential to increase endoscopists' workload due to unnecessary resection of additional detected hyperplastic polyps.[26] In particular, inexperienced examiners could be prone to that problem. Combined CADx systems can encounter this problem but their development is at an earlier stage. Existing study data mainly relies on the analysis of high-definition images, which have been in part retrospectively collected and reviewed for their quality.[36–38] Prospective multicenter studies analyzing data of real-world images in real time have to confirm these results. A further limitation is the lack of data for the differentiation of SAAs and hyperplastic polyps. SSAs were excluded in those studies because of the similarity of their surface pattern to hyperplastic polyps.

In particular, nonexpert colonoscopists are expected to benefit from support by CAD systems. Improved detection rates and confidence in the characterization of lesions are expected. However, the learning phase could potentially be hindered by the black-box nature of DL algorithms when using the CAD as a substitute for a skilled mentor who teaches patterns of polyp classification. Furthermore, concentration distraction (Figure 1) and over-reliance on AI systems are potential sequelae.[3] Studies involving examiners at different levels of experience have to evaluate this point.

Evaluation of AI systems is of urgent need as several commercial systems have entered the market already. There is no standardized way of describing algorithm architecture, data sets, as well as training, validation, and testing of CAD systems. Technical details pose many challenges to clinicians.[45] Differing test data sets resembling various clinical challenges prevent direct comparability of the performance between different systems. Benchmarking data sets containing heterogeneous data, representing the natural variability of polyps, could overcome that problem if used to evaluate multiple systems the same way.

Until today, there is no standard definition for false positive (FP) or false negative (FN) detections. Some research groups defined a qualitative definition in which they considered a FP detection any activation continuously traced by the system or a consistent detection.[5,8,9,46] Other research groups consider FP detections as any activation, irrespectively of the number of frames.[47] By using a frame-by-frame approach, meaning counting an FP as any false activation and FN as any missed activation in a frame, the comparison between CADe systems quality would be more accurate. In this sense, Bernal et al. proposed “temporal coherence” on the method's response as a way to evaluate these features that would solve such a problem.[48] Standards on AI validation have to be established. Furthermore, clinical relevance of FP detections has not been established so far. Hassan et al. proposed a classification system to address this problem, which should be evaluated in further studies.[47]

COMMERCIAL CAD SYSTEMS AND REGULATIONS

Several commercial CAD systems have now entered the market (Table 1). Usage of any device in Europe requires a CE mark. From May 2021, a new Medical Devices Regulation (MDR) will be applied in Europe.[49] Devices will still be classified in different risk classes

(Class I, Class IIa, Class IIb and Class III) according to their intended purpose. Class I are noninvasive low-risk devices. Class IIa are low to medium risk, Class IIb are medium to high and Class III are high-risk devices.

The first established CADx system was EndoBRAIN from Cybernet Systems Co., Japan, that has been developed in collaboration with Kudo S and Mori K et al. It is based on endocytoscopy and a conventional machine learning algorithm. Analyzing images taken at 520-fold magnification the CADx accuracy in distinguishing neoplasms from non-neoplasms is about 98%.[50] An additional CADe system called EndoBRAIN-EYE was released in 2020.

GI Genius was the first commercial CADe system. It obtained CE approval in July 2019 and is marked by Medtronic, Ireland. The CADe system has been developed by Cosmo Pharmaceuticals and validated by Hassan C and Repici A et al. Its utilization in real-time colonoscopy increased significantly ADR in a multicenter trial.[4] DISCOVERY from Pentax Medical, Japan, is another CADe system that was CE approved in January 2020. Both systems were classified as Class I. In their case, the intended purpose is to work as an adjunct highlighting region such as colorectal polyps, but under no circumstances, their output can override that of the endoscopist. In such cases, the clinical evidence required is less stringent.

Under the new MDR, almost all software will be up-classified into higher risk classes.[51] Software intended to provide information that is used to make decisions with diagnosis or therapeutic purposes is classified as Class IIa, or even higher depending on the risk associated with this decision. Recently, United European Gastroenterology recommended that all AI-based systems for polyp detection should be classified as IIa-products and those used for differentiation as IIb-products.[52]

The corresponding system from Fujifilm is called CAD EYE. It supports polyp detection in white light endoscopy (WLE) and Linked color imaging as well as polyp characterization in Blue Light Imaging (BLI). The CADe system obtained a CE mark in February 2020.

Ai4gi was developed in collaboration with Byrne M and Rex D et al. It is commercialized by a joint venture between Satis Operations and Imagia, USA. The CADe system automatically switches to CADx mode when NBI light is detected. An accuracy of 97.6% for differentiating between NICE 1 and NICE 2 polyps has been shown recently.[40]

The EndoScreener from Shanghai Wision AI Co., China, is a CADe system developed in a joint work with Wang P et al. Although lacking CE approval, it has already been evaluated in prospective clinical trials in China. By its use, increased ADR for diminutive polyps and a significantly lower AMR compared to experienced colonoscopists was shown.[6,8]

Approval of future CAD systems will require a significantly higher level of evidence and they must provide post-market surveillance and prepare a periodic safety update report ('PSUR')51.

A White Paper on AI was released by the European Commission to present policy options.[53] A High-Level Expert Group published nonbinding risk-based Guidelines on trustworthy AI. They conclude that the future regulation should deal with the following: data and record-keeping to allow traceback and verification, robustness and accuracy in results, gender and racial bias, and personal data and privacy protection.

OUTLOOK

AI in polyp detection and characterization is on the verge of clinical application. CADe already implied its potential use in prospective trials. Further efforts are needed to take CADx to the next level of development. The prediction of clinical impact in CRC surveillance is limited by the current study data. Future devices will encounter more challenging requirements for approval and will be categorized into higher MDR classes.

CONFLICT OF INTEREST

The authors Alexander Hann and Daniel Fitting receive public funding from the state government of Baden-Württemberg, Germany (Funding cluster "Forum Gesundheitsstandort Baden-Württemberg") to research and develop artificial intelligence applications for polyp detection in screening colonoscopy.

AUTHOR CONTRIBUTIONS

Conception and design: all authors; drafting of the article: all authors; critical revision of the article for important intellectual content: all authors; final approval of the article: all authors.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Alexander Hann https://orcid.org/0000-0001-8035-3559

REFERENCES

1. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, et al. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med. 2014;370:1298-306.
2. Wadhwa V, Alagappan M, Gonzalez A, Gupta K, Brown JRG, Cohen J, et al. Physician sentiment toward artificial intelligence (AI) in colonoscopic practice: a survey of US gastroenterologists. Endosc Int Open. 2020;8:E1379-84.
3. Bisschops R, East JE, Hassan C, Hazewinkel Y, Kamiński MF, Neumann H, et al. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) guideline—update 2019. Endoscopy. 2019;51:1155-79.
4. Repici A, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, et al. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. Gastroenterology. 2020;159:512-20.
5. Liu W-N, Zhang Y-Y, Bian X-Q, Wang L-J, Yang Q, Zhang X-D, et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. Saudi J Gastroenterol. 2020;26:13-9.
6. Wang P, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, et al. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. Lancet Gastroenterol Hepatol. 2020;5:343-51.

7. Su J-R, Li Z, Shao X-J, Ji C-R, Ji R, Zhou R-C, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). Gastrointest Endosc. 2020;91:415-24.
8. Wang P, Liu P, Glissen Brown JR, Berzin TM, Zhou G, Lei S, et al. Lower adenoma miss rate of computer-aided detection-assisted colonoscopy vs routine white-light colonoscopy in a prospective tandem study. Gastroenterology. 2020;159:1252-61. https://doi.org/10.1053/j.gastro.2020.06.023
9. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. Gut. 2019;68:1813-9.
10. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med. 2010;362:1795-803.
11. Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, et al. Performance of artificial intelligence for colonoscopy regarding adenoma and polyp detection: a meta-analysis. Gastrointest Endosc. 2020;93:77-85. https://doi.org/10.1016/j.gie.2020.06.059
12. Brenner H, Altenhofen L, Kretschmann J, Rösch T, Pox C, Stock C, et al. Trends in adenoma detection rates during the first 10 Years of the German screening colonoscopy program. Gastroenterology. 2015;149:356-66.
13. Kaminski MF, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) quality improvement initiative. United European Gastroenterol J. 2017;5:309-34.
14. Barclay RL, Vicari JJ, Doughty AS, Johanson JF, Greenlaw RL. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med. 2006;355:2533-41.
15. Rembacken B, Hassan C, Riemann JF, Chilton A, Rutter M, Dumonceau J-M, et al. Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE). Endoscopy. 2012;44:957-68.
16. Gong D, Wu L, Zhang J, Mu G, Shen L, Liu J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol. 2020;5:352-61.
17. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel preparation scale. Gastrointest Endosc. 2010;72:686-92.
18. Cohen LB. Split dosing of bowel preparations for colonoscopy: an analysis of its efficacy, safety, and tolerability. Gastrointest Endosc. 2010;72:406-12.
19. Radaelli F, Paggi S, Hassan C, Senore C, Fasoli R, Anderloni A, et al. Split-dose preparation for colonoscopy increases adenoma detection rate: a randomised controlled trial in an organised screening programme. Gut. 2017;66:270-7.
20. Pox C, Wesselmann S, Thurn J, Giuliani A, Schmiegel W. S3-Leitlinie Kolorektales Karzinom, Langversion 2.1,2019, AWMF Registrierungsnummer: 021/007OL. 2019. https://www.dgvs.de/wp-content/uploads/2019/01/LL_KRK_Langversion_2.1.pdf. Accessed 1 June 2020.
21. Denzer U, Beilenhoff U, Eickhoff A, Faiss S, Hüttl P, In der Smitten S, et al. [S2k guideline: quality requirements for gastrointestinal endoscopy, AWMF registry no. 021-022]. Z Gastroenterol. 2015;53:E1-227.
22. Zhou J, Wu L, Wan X, Shen L, Liu J, Zhang J, et al. A novel artificial intelligence system for the assessment of bowel preparation (with video). Gastrointest Endosc. 2000;91:428-35.
23. Thakkar S, Carleton NM, Rao B, Syed A. Use of artificial intelligence-based analytics from live colonoscopies to optimize the quality of the colonoscopy examination in real time: proof of concept. Gastroenterology. 2020;158:1219-21.
24. Butterly LF, Chase MP, Pohl H, Fiarman GS. Prevalence of clinically important histology in small adenomas. Clin Gastroenterol Hepatol. 2006;4:343-8.
25. Ponugoti PL, Cummings OW, Rex DK. Risk of cancer in small and diminutive colorectal polyps. Dig Liver Dis. 2017;49:34-7.
26. Hassan C, Pickhardt PJ, Rex DK. A resect and discard strategy would improve cost-effectiveness of colorectal cancer screening. Clin Gastroenterol Hepatol. 2010;8:865-9.
27. Kudo S, Tamura S, Nakajima T, Yamano H-o, Kusaka H, Watanabe H. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. Gastrointest Endosc. 1996;44:8-14.
28. Hewett DG, Kaltenbach T, Sano Y, Tanaka S, Saunders BP, Ponchon T, et al. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. Gastroenterology. 2012;143:599-607.
29. Tischendorf JJW, Wasmuth HE, Koch A, Hecker H, Trautwein C, Winograd R. Value of magnifying chromoendoscopy and narrow band imaging (NBI) in classifying colorectal polyps: a prospective controlled study. Endoscopy. 2007;39:1092-6.
30. East JE, Vleugels JL, Roelandt P, Bhandari P, Bisschops R, Dekker E, et al. Advanced endoscopic imaging: European Society of Gastrointestinal Endoscopy (ESGE) technology review. Endoscopy. 2016;48:1029-45.
31. Rex DK, Kahi C, O'Brien M, Levin TR, Pohl H, Rastogi A, et al. The American Society for Gastrointestinal Endoscopy PIVI (preservation and incorporation of valuable endoscopic innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc. 2011;73:419-22.
32. Patel SG, Schoenfeld P, Kim HM, Ward EK, Bansal A, Kim Y, et al. Real-time characterization of diminutive colorectal polyp histology using narrow-band imaging: implications for the resect and discard strategy. Gastroenterology. 2016;150:406-18.
33. Ladabaum U, Fioritto A, Mitani A, Desai M, Kim JP, Rex DK, et al. Real-time optical biopsy of colon polyps with narrow band imaging in community practice does not yet meet key thresholds for clinical decisions. Gastroenterology. 2013;144:81-91.
34. Kuiper T, Marsman WA, Jansen JM, van Soest EJ, Haan YCL, Bakker GJ, et al. Accuracy for optical diagnosis of small colorectal polyps in nonacademic settings. Clin Gastroenterol Hepatol. 2012;10:1016-20.
35. Ferlitsch M, Moss A, Hassan C, Bhandari P, Dumonceau J-M, Paspatis G, et al. Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) clinical guideline. Endoscopy. 2017;49:270-97.
36. Chen P-J, Lin M-C, Lai M-J, Lin J-C, Lu HH-S, Tseng VS. Accurate classification of diminutive colorectal polyps using computer-aided analysis. Gastroenterology. 2018;154:568-75.
37. Renner J, Phlipsen H, Haller B, Navarro-Avila F, Saint-Hill-Febles Y, Mateus D, et al. Optical classification of neoplastic colorectal polyps—a computer-assisted approach (the COACH study). Scand J Gastroenterol. 2018;53:1100-6.
38. Zachariah R, Samarasena J, Luba D, Duh E, Dao T, Requa J, et al. Prediction of polyp pathology using convolutional neural networks achieves "resect and discard" thresholds. Am J Gastroenterol. 2020;115:138-44.
39. Komeda Y, Handa H, Watanabe T, Nomura T, Kitahashi M, Sakurai T, et al. Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience. Oncology. 2017;93:30-4.
40. Byrne MF, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, et al. Real-time differentiation of adenomatous and hyperplastic

diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. Gut. 2019;68:94-100.
41. Bretthauer M, Kaminski MF, Løberg M, Zauber AG, Regula J, Kuipers EJ, et al. Population-based colonoscopy screening for colorectal cancer: a randomized clinical trial. JAMA Intern Med. 2016;176:894-902.
42. Vleugels JLA, Hassan C, Senore C, Cassoni P, Baron JA, Rex DK, et al. Diminutive polyps with advanced histologic features do not increase risk for metachronous advanced colon neoplasia. Gastroenterology. 2019;156:623-34.
43. Tanaka H, Oka S, Tanaka S, Inagaki K, Okamoto Y, Matsumoto K, et al. Can surveillance colonoscopy be discontinued in an elderly population with diminutive polyps?. J Anus Rectum Colon. 2019;3:128-35.
44. Cai B, Liu Z, Xu Y, Wei W, Zhang S. Adenoma detection rate in 41,010 patients from Southwest China. Oncol Lett. 2015;9:2073-7.
45. van der Sommen F, de Groof J, Struyvenberg M, van der Putten J, Boers T, Fockens K, et al. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. Gut. 2020;69:2035-45.
46. Barua I, Vinsard D, Jodal H, Løberg M, Kalager M, Holme Ø, et al. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. Endoscopy. 2020;53:277-84. https://doi.org/10.1055/a-1201-7165
47. Hassan C, Badalamenti M, Maselli R, Correale L, Iannone A, Radaelli F, et al. Computer-aided detection-assisted colonoscopy: classification and relevance of false positives. Gastrointest Endosc. 2020;92:900-4.
48. Bernal J, Tajkbaksh N, Sanchez FJ, Matuszewski BJ, Chen H, Yu L, et al. Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. IEEE Trans Med Imag. 2017;36:1231-49.
49. Regulation (EU) 2020/561 of the European Parliament and of the Council of 23 April 2020 amending Regulation (EU) 2017/745 on medical devices, as regards the dates of application of certain of its provisions (Text with EEA relevance). 32020R0561. 2020. http://data.europa.eu/eli/reg/2020/561/oj/eng, Accessed 18 Nov 2020.
50. Mori Y, Kudo S, Chiu P, Singh R, Misawa M, Wakamura K, et al. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. Endoscopy. 2016;48:1110-8.
51. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance.). 32017R0745. 2017. http://data.europa.eu/eli/reg/2017/745/oj/eng. Accessed 18 Nov 2020.
52. https://ueg.eu/files/1468/cec6f62cfb44b1be110b7bf70c8362d8.pdf
53. European Commission. White Paper on Artificial Intelligence: a European approach to excellence and trust, https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Accessed 18 Nov 2020.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Hann A, Troya J, Fitting D. Current status and limitations of artificial intelligence in colonoscopy. *United European Gastroenterol J*. 2021;9:527–533. https://doi.org/10.1002/ueg2.12108