Frame-by-Frame Analysis of a Commercially Available Artificial Intelligence Polyp Detection System in Full-Length Colonoscopies

Markus Brand\textsuperscript{a} Joel Troya\textsuperscript{a} Adrian Krenzer\textsuperscript{a,b} Costanza De Maria\textsuperscript{c,d} Niklas Mehlhase\textsuperscript{e} Sebastian Götze\textsuperscript{e} Benjamin Walter\textsuperscript{e} Alexander Meining\textsuperscript{a} Alexander Hann\textsuperscript{a}

\textsuperscript{a}Interventional and Experimental Endoscopy (InExEn), Department of Internal Medicine II, University Hospital Würzburg, Würzburg, Germany; \textsuperscript{b}Artificial Intelligence and Knowledge Systems, Institute for Computer Science, Julius-Maximilians-Universität, Würzburg, Germany; \textsuperscript{c}Department of Gastroenterology and Hepatology, Ente Ospedaliero Cantonale (EOC), Bellinzona, Switzerland; \textsuperscript{d}Department of Biomedical Science, University of Italian Switzerland (USI), Lugano, Switzerland; \textsuperscript{e}Department of Internal Medicine I, University Hospital Ulm, Ulm, Germany

Keywords
Colonoscopy · Deep learning · Computer-aided detection · Artificial intelligence

Abstract

\textbf{Introduction:} Computer-aided detection (CADe) helps increase colonoscopic polyp detection. However, little is known about other performance metrics like the number and duration of false-positive (FP) activations or how stable the detection of a polyp is. \textbf{Methods:} 111 colonoscopy videos with total 1,793,371 frames were analyzed on a frame-by-frame basis using a commercially available CADe system (GI-Genius, Medtronic Inc.). Primary endpoint was the number and duration of FP activations per colonoscopy. Additionally, we analyzed other CADe performance parameters, including per-polyp sensitivity, per-frame sensitivity, and first detection time of a polyp. We additionally investigated whether a threshold for withholding CADe activations can be set to suppress short FP activations and how this threshold alters the CADe performance parameters. \textbf{Results:} A mean of 101 ± 88 FPs per colonoscopy were found. Most of the FPs consisted of less than three frames with a maximal 66-ms duration. The CADe system detected all 118 polyps and achieved a mean per-frame sensitivity of 46.6 ± 26.6%, with the lowest value for flat polyps (37.6 ± 24.8%). Withholding CADe detections up to 6 frames length would reduce the number of FPs by 87.97% \((p < 0.001)\) without a significant impact on CADe performance metrics. \textbf{Conclusions:} The CADe system works reliable but generates many FPs as a side effect. Since most FPs are very short, withholding short-term CADe activations could substantially reduce the number of FPs without impact on other performance metrics. Clinical practice would benefit from the implementation of customizable CADe thresholds.

Introduction

Artificial intelligence (AI) is presumably a powerful tool in colorectal cancer prevention using colonoscopy, as several randomized controlled trials (RCTs) have shown that computer-aided detection (CADe) increases Markus Brand and Joel Troya contributed equally to this work.
adenoma detection rate (ADR) and decreases the miss rate of neoplastic lesions [1–6].

However, there are still many unanswered questions regarding CADe systems. For example, many false-positive (FP) activations of up to 8% of all frames occur during examination with CADe systems [7]. The number and duration of FP activations play an important role regarding the examiners comfort in using those systems, as these activations can affect the examiners attention leading to misinterpretation of normal mucosa [8]. Therefore, an international consensus conference has identified the analysis of FP activations as an important research focus [9]. Current studies on this topic include only small numbers of cases with about 40 colonoscopy examinations and mainly investigate the cause and the clinical impact of FP activations [10, 11]. However, specific data on the duration and pattern of FP activations are not available, although such information is necessary to better understand the operation of CADe systems in order to improve them. An example for improvement might be the reduction of FPs through customizable activation thresholds. In addition, previous RCTs only provide data on per-polyp sensitivity (PPS), i.e., whether a polyp was detected resulting in a yes or no answer. How stable the detection signal is over time, termed per-frame sensitivity, was not assessed as no frame-by-frame analysis of real full-length videos has been performed so far.

Therefore, the objective of this study was to analyze the FP pattern of a commercial CADe system. This was done using a frame-by-frame analysis of full-length real-life videos to determine the effects of different CADe activation thresholds on FPs. Additionally, in a patient-based analysis, we examined performance parameters such as PPS or the mean number of polyps per colonoscopy (PPC).

Materials and Methods

Study Design

Videos from 244 routine colonoscopies performed in two tertiary centers (University Hospital Ulm and Würzburg) were retrospectively analyzed. Recording took place between March 2019 and April 2020. Those colonoscopies (raw signals) were recorded using the high-definition video signal of the endoscopy processor (Olympus CV-190). For the performance analysis of a commercially available CADe system (GI Genius, Medtronic Inc., Ireland, software version of March 2020), this raw video signal was introduced into the AI system, and the output signal (with visible CADe detections) was recorded. Accordingly, a video pair consisting of raw signal and CADe signal was assembled for video analysis of each colonoscopy.

Colonoscopies

Colonoscopies were performed using the colonoscopes CF-HQ190AL and CF H180AI/AL (Olympus Co., Tokyo, Japan). All patients were prepared for the colonoscopy using a standard split-dose regimen with 2L polyethylene glycol with ascorbic acid (Moviprep, Norgine Pharma; Harefield, England). Endoscopies were performed using nurse-assisted propofol sedation [12]. Polyps were removed upon detection by cold or hot snare technique if no contraindication for resection was present. The examiners were classified due to their experience in colonoscopy between junior and senior with 2,000 performed colonoscopies as a threshold.

Video Analysis

All videos were screened by a board-certified gastroenterologist and experienced endoscopist (MB) with over 4,000 performed colonoscopies. Examinations performed for screening reasons or post polypectomy surveillance were included in the analysis. For further analysis, the following exclusion criteria were defined: inflammatory bowel disease, active gastrointestinal bleeding, poor bowel preparation defined by a Boston Bowel Preparation Scale (BBPS) lower than 5, incomplete colonoscopies, advanced neoplasia, altered gut anatomy, endoscopy only performed for an extended resection and polyposis syndrome. Included colonoscopies were analyzed in a deep frame-by-frame manner using a custom-made annotation tool as previously described [13].

Analysis of Non-CADe Signal (Raw Videos)

The start and the end of withdrawal and polypectomies were annotated. Each polyp was counted for the analysis. Additionally, polyps were characterized using the Paris classification and size (<5 mm, 5–10 mm, 11–20 mm, >20 mm). In a frame-by-frame analysis, each frame with a partially or completely visible polyp was annotated as a polyp frame. Frames with even small parts of a polyp visible were regarded as a polyp-containing frame. Polyp annotation stopped at the beginning of the resection (first frame with a visible instrument in the image).

Analysis of CADe Signal (AI Videos)

All frames with visible bounding boxes resembling CADe detections were automatically identified by a custom-made application. Subsequently, each bounding box was classified by an experienced endoscopist (MB) as a true-positive (TP) or FP detection. It was considered TP if the bounding box had contact with the visible polyp, irrespective of how much area of the lesion was covered. Small hyperplastic polyps of the rectosigmoid were excluded from the analysis. The absence of a bounding box in a frame with a visible polyp was regarded as a false negative. The absence of a box in a frame without a polyp was considered a true negative. A FP detection was defined as a detected area that was not in contact with a polyp. In case of a FP detection in a frame with a visible polyp, the term distraction was used.

Endpoints

The primary endpoint of the study was the number of FP activations per colonoscopy and the duration of FP activations. For the secondary endpoints, we analyzed further CADe performance parameters, including mean number of PPC of the CADe System, PPS, per-frame sensitivity, and first detection time (FDT) of a polyp.
In addition, we investigated whether a threshold for withholding short CADe activations can be set to suppress FP activations and how this threshold alters CADe performance parameters such as PPC, PPS, or per-frame sensitivity.

**Data Analysis and Statistics**

FP activations were counted in their number, with each contiguous sequence of FP frames counted as one activation. In addition, the duration of FP activations was measured in frames. Each frame had a duration of 33 ms. The mean number of PPC was calculated by dividing the number of detected polyps by the number of performed colonoscopies. PPS was defined as the number of polyps detected by the CADe system in at least one frame divided by the number of polyps annotated in the raw video data. The per-frame sensitivity, previously published as temporal coherence, was calculated by dividing the number of TP frames by the total number of frames where the polyp was visible in the raw signal (TP + false negative), as previously described by Zhou et al. [14]. Additionally, the per-lesion sensitivity, defined as the number of polyps in which more than half of each polyp’s frame were detected by the CADe, divided by the total number of polyps, was analyzed as previously described by Misawa et al. [15]. FDT of a polyp was defined as the time interval between the first appearance of a polyp in the raw video and the first frame containing a TP-CADe activation. If the polyp was not permanently visible during this time span, frames without a visible polyp were excluded. By this method, FDT included only frames with a visible polyp. The mean withdrawal time was determined using the recorded videos and defined as the time frame between the coecum and anal canal, excluding time spent for performing biopsies or snare resection [16].

Statistical analysis was performed using Python version 3.8. The χ² and Fisher’s exact tests were used to test for significant differences between categorical variables. Student’s t test and Mann-Whitney U test were applied for continuous variables depending on their distribution pattern. A p value of <0.05 indicated statistical significance.

**Results**

**Baseline Characteristics**

From 244 routine colonoscopies, 133 colonoscopies met the exclusion criteria. Thus, a total of 111 pairs of colonoscopy videos including the raw video signal and the CADe signal were analyzed (Fig. 1) in a deep frame-
by-frame manner. Most of the examinations (65.8%) were done by experienced investigators with over 2,000 performed colonoscopies. The mean BBPS score was 7.5, with the lowest value being 6. The mean withdrawal time was 8:58 min. A total of 118 polyps were identified and annotated in the 111 videos. Most of the polyps were diminutive with 1–5 mm in size (55.08%) with flat or sessile shape (Paris 0-Ip/Ia; 35.59%/59.32%). Baseline characteristics of the colonoscopies and detailed characterization of the polyps are shown in Table 1. In total, the 111 examinations analyzed contained 1,793,371 frames, including 173,959 frames (9.7%) with polyps and 1,619,412 frames (90.3%) without polyps. Three polyps were detected by the CADe but not perceived by the endoscopist.

**Primary Endpoint**

Rate of FP Detections and Distracting Detections

A total of 11,188 FP activations were detected in the 111 coloscopies (101 ± 88 FPs per colonoscopy). The mean duration of a FP activation was 135 ms. In relation to the withdrawal time, the FPs account for a mean of 2.48% resembling 13.61 s. Most of the FP detections consisted of one to two frames, corresponding to a period of max. 66 ms (Fig. 2). Only a minority of detections accounted for continuous detections consisting of 10 frames or more, resembling more than 330 ms. In the subgroup of colonoscopies with at least one polyp, we examined the frames with a FP CADe detection in an image with a visible polyp, termed distracting detection. Here we found that 1.6 ± 2.1% of the frames with polyps contain this distraction.

**Secondary Endpoints**

PPC and PPS

The CADe system detected all 118 polyps that were visible in the videos, resulting in a PPS of 100%. The mean number of PPC was 1.06.

Per-Frame Sensitivity and Per-Lesion Sensitivity

The mean per-frame sensitivity of the CADe system for all 118 polyps was 47.73 ± 26.5% (Table 2). The mean per-lesion sensitivity of the CADe system was 47.46%.
a subgroup analysis, we found that the per-frame sensitivity was significantly lower in flat polyps (Paris 0-IIa) compared to 0-Ip or 0-Is configuration (37.99 ± 24.65% vs. 74.86 ± 21.78%, p < 0.001 or 37.99 ± 24.65% vs. 60.10 ± 22.31%, p < 0.001). While polyp size did not influence the per-frame sensitivity, polyp localization in the right-sided colon segments was associated with lower mean per-frame sensitivity, compared to the left-sided colon segments (43.56 ± 25.04% vs. 53.43 ± 26.00, p = 0.017).

FDT of a Polyp
FDT was available for each of the 118 polyps. The mean FDT was 1,692 ± 2,052 ms with a wide range from 33.3 to 12,033 ms. In a subgroup analysis, we found the highest FDT in the polyp size group 11–20 mm with mean 2,179 ± 3,174 ms (Table 3). However, this was not significant when compared to size groups 1–5 mm and 6–10 mm. In contrast, we found a significantly higher FDT in Paris 0-IIa polyps in comparison to 0-Ip or 0-Is polyps (2,068 ± 2,413 ms vs. 522 ± 216 ms or 1,233 ± 1,247 ms, p = 0.023 and p = 0.046).

Impact of Different CADe Activation Thresholds on FPs and CADe Performance Parameters
To estimate the effect of withholding CADe activations of a defined frame length on the FPs and the CADe performance, a subgroup analysis was performed using only activations of a defined frame length or longer. Figure 3 shows graphically how withholding short activations of 1–10 frames significantly reduces the rate of FPs while having little effect on the per-frame sensitivity of

**Table 3. Time to first detection of a visible polyp by the CADe system**

| Characteristic     | Mean, ms ± SD (n) |
|--------------------|-------------------|
| Time to first detection | 1,692.66±2,052.73 (118) |
| Paris classification |                   |
| 0-Ip               | 522.22±216.71 (6)  |
| 0-Is               | 1,233.33±1,246.96 (42) |
| 0-IIa              | 2,068.57±2,413.86 (70) |
| Size               |                   |
| 1–5 mm             | 1,621.03±1,792.53 (65) |
| 6–10 mm            | 1,474.44±1,420.34 (30) |
| 11–20 mm           | 2,179.71±3,174.02 (23) |

SD, standard deviation; ms, milliseconds; CADe, computer-aided detection.
on the impact of different thresholds on PPC, PPS, and per-frame sensitivity are shown in Table 4. In addition, online supplementary Video 1 (for all online suppl. material, see www.karger.com/doi/10.1159/000525345) shows an example of how a threshold of 6 frames (no significant changes in PPC, PPS, or per-frame sensitivity) affects FP activations in the endoscopic view.

**Discussion**

The development of an AI system for polyp detection using deep learning techniques applied on a larger dataset was first described by Wang et al. [17]. Subsequently, several commercially available CADe systems have been developed for colonoscopy. In prospective RCTs, CADe systems showed a significantly higher ADR compared to expert colonoscopists [1–4, 7, 18–21]. Moreover, a recently published meta-analysis found a significant increase of ADR [6]. While prospective studies have extensively evaluated the ADR of various CADe systems, little is known about the detailed performance of CADe systems, e.g., FP rate, FP duration, or per-frame sensitivity, especially in a real-life scenario. Only a few studies about CADe systems include a single-frame analysis. However, these studies used single polyp frames, short video sequences, or videos consisting of less than 160,000 frames [14, 15, 17]. Thus, we present the largest frame-by-frame dataset, to our knowledge, with 111 full-length videos consisting of over 170,000 polyp frames and a total of over 1,700,000 frames. Additionally, to the best of our knowledge, our study is the first evaluating CADe performance in a frame-by-frame analysis in real-life videos.

The PPS of 100% highlights the effectiveness of CADe systems in clinical practice; however, the number of FP activations is not negligible and is higher than the previously published values [10, 11]. While previous studies analyzed the cause and clinical relevance of FPs, the use of frame-by-frame analysis allowed us to determine the exact duration and distribution pattern of FPs. As shown, most FPs were shorter than 330 ms, hence they are perceived by the endoscopist only as a brief flashing of the bounding box. However, it is not yet clear whether the short activations do or do not affect the normal mucosa visualization pattern of endoscopists. Some retrospective studies suggest that FP activations result in the negligible increase of the total withdrawal time, as most of them are immediately discarded by the endoscopists [10, 11]. Other studies using for example eye-tracking glasses suggest that CADe and FPs activations might have an impact on the visualization pattern of the endoscopists [8, 22]. Therefore, further studies using eye tracking technology during endoscopic examinations in a prospective manner should be performed in order to analyze the influence of short FP activations on the examiner and the withdrawal time. Nevertheless, many short FPs may impair the endoscopist’s concentration in the long run; certainly, they reduce the comfort of the CADe application.

| Threshold, frames | FP reduction % (p value) | PPC, value (p value) | PPS, value (p value) | Per-frame sensitivity, mean ± SD (p value) |
|-------------------|--------------------------|---------------------|---------------------|------------------------------------------|
| None              | 0%                       | 1.06                | 1                   | 47.73±26.50                              |
| 1                 | 34.25 (<0.001)           | 1.06 (1.00)         | 1 (1.00)            | 47.29±26.54 (0.82)                       |
| 2                 | 54.78 (<0.001)           | 1.05 (0.99)         | 0.99 (1.00)         | 46.76±26.63 (0.70)                       |
| 3                 | 67.96 (<0.001)           | 1.04 (0.91)         | 0.97 (0.25)         | 46.10±26.72 (0.59)                       |
| 4                 | 76.16 (<0.001)           | 1.03 (0.83)         | 0.96 (0.13)         | 45.43±26.76 (0.45)                       |
| 5                 | 81.46 (<0.001)           | 1.02 (0.83)         | 0.96 (0.13)         | 44.61±26.97 (0.34)                       |
| 6                 | 85.54 (<0.001)           | 1.02 (0.83)         | 0.96 (0.07)         | 43.86±27.06 (0.24)                       |
| 7                 | 87.92 (<0.001)           | 1 (0.75)            | 0.94 (0.02)         | 43.29±27.21 (0.19)                       |
| 8                 | 89.97 (<0.001)           | 1 (0.75)            | 0.94 (0.02)         | 42.73±27.23 (0.14)                       |
| 9                 | 91.42 (<0.001)           | 1 (0.75)            | 0.94 (0.02)         | 42.25±27.2 (0.10)                        |
| 10                | 92.29 (<0.001)           | 0.99 (0.71)         | 0.93 (0.01)         | 41.63±27.4 (0.07)                        |

The threshold value indicates that activations of the value or shorter have been eliminated. SD, standard deviation; PPS, per-polyp sensitivity; PPC, polyps per colonoscopy. p values represent comparisons to the CADe signal without a threshold.
An option to reduce the FP rate, especially for short FP, could be withholding of short CADe activations. As shown, withholding short detections up to 10 frames length reduced the number of FP by up to 92.79% without having a significant effect on per-frame sensitivity. However, above a threshold of 3 frames representing 100 ms, this is at the expense of a few missed polyps, especially those with a flat shape. Another effect to consider should be the impact that the withholding of short CADe activations could have on FDT. Unfortunately, there are no studies that demonstrate the effect of different FDTs on the detection of polyps. However, considering the big effect in the reduction of FP activations and since there was no significant change in PPC or PPS up to a threshold of 6 frames (200 ms), an appropriate threshold for optimization of the CADe system could be in this range.

Besides PPC, PPS, and FP rate, per-frame sensitivity is another important performance parameter of CADe systems, particularly since the temporal stability of polyp detection indicates how well CADe detection works for different polyp types. The per-frame sensitivity determined in our study is lower than in previous publications [14, 15, 17]. However, in previous studies, only several single images of polyps or selected video sequences were used to evaluate the self-developed systems. For example, the study by Misawa et al. [15] analyzed video clips with a total of 152,560 frames. In our study, full-length real-life videos containing 1,793,371 frames were used, so the conditions for CADe detection may have been more challenging, yet more realistic. Another important reason is that small hyperplastic polyps in the rectosigmoid were excluded in our study due to clinical irrelevance, whereas these polyps, which can often be reliably identified, were included in the evaluation in previous studies.

Since flat polyps (Paris 0-IIa) and sessile serrated adenomas have higher miss rates, the effect of CADe systems on polyp detection could be substantial if these lesions were reliably detected [23]. However, our data show that in clinical practice, per-frame sensitivity and FDT tend to be worse in these polyps. These findings are consistent with previous data, reporting lower per-frame sensitivity for laterally spreading tumors and sessile serrated adenoma, showing that there is an urgent need for improvement in this point [14].

There are several limitations to our study. Since this is a retrospective analysis of previously stored videos, histologic differentiation of colonic polyps was not possible. In order to increase the relevance of the detected polyps, we excluded hyperplastic polyps in the rectosigmoid. Due to the exclusion of examinations with a BBPS score of <6 points, the mean BBPS score is 7.5 points, which is relatively high [24]. However, recently published papers on CADe performance metrics describe similarly high BBPS values [10, 11]. To shorten the time-consuming deep frame analysis, we have dispensed with a detailed analysis of the FPs with respect to their cause. However, Hassan et al. [10] performed such an analysis using the same CADe system – they found bubbles, stool, and colonic folds to be the main reasons for FP activation. We also did not manually annotate each polyp-containing frame with bounding boxes. Thus, subsequent analysis of, for example, intersection over the union of the CADe boxes with ground truth was not performed.

Conclusion

This commercially available CADe system is a powerful tool to facilitate polyp detection even under daily clinical conditions, but at the expense of many FP activations. Through a frame-by-frame video analysis, we were able to show that many of these FPs are of very short duration. Withholding short-term CADe detections could substantially reduce the number of FP activations, but at higher thresholds at the expense of a few missed polyps. This applies in particular to flat polyps, which generally have poorer per-frame sensitivity values. Since we could not detect any significant change in the mean number of PPC and PPS up to a threshold of 6 frames, an appropriate threshold for optimization of the CADe system could be in this range. Nevertheless, further detailed analysis of CADe systems is needed to better understand the strengths and weaknesses of this promising technology and to further optimize the systems. A customizable CADe detection threshold that can be adjusted to the needs of the examiner would be useful in clinical practice.

Statement of Ethics

This study protocol involving retrospective analysis of data was reviewed and approved by the Ethics Committee of the University Hospital Würzburg, approval number 2021032901. According to the Ethics Committee of the University Hospital Würzburg, patients were not required to give informed consent for this retrospective analysis.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.
Funding Sources
Alexander Hann receives 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. Alexander Meining receives funding from the IZKF Würzburg and the Bavarian Center for Cancer Research (BZKF) for further implementation and development of artificial intelligence for detection of (pre)neoplastic lesions.

Author Contributions
Markus Brand, Joel Troya, Alexander Meining, and Alexander Hann: study concept and design, interpretation of results, and drafting of the manuscript. Joel Troya: statistical analysis. Markus Brand: annotation of videos. Joel Troya, Adrian Krenzer, Costanza De Maria, Niklas Mehlhase, Sebastian Götzte, and Benjamin Walter: acquisition of data. Markus Brand, Joel Troya, Adrian Krenzer, Costanza De Maria, Niklas Mehlhase, Sebastian Götzte, Benjamin Walter, Alexander Meining, and Alexander Hann: critical revision of the article for important intellectual content and final approval of the article.

Data Availability Statement
The data underlying this article will be shared on reasonable request to the corresponding author.

References
1 Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Beccq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomized controlled study. Gut. 2019 Oct;68(10):1813–9.
2 Liu WN, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. Saudi J Gastroenterol. 2020 Jan-Feb;61(1):13–9.
3 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 Aug;159(2):512–20 e7.
4 Su JR, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). Gastroint Endosc. 2020 Feb;91(2):415–24 e4.
5 Deliwa SA, Hamid K, Barbarawi M, Lakshman H, Zayed Y, Kandel P, et al. Artificial intelligence (AI) real-time detection vs. routine colonoscopy for colorectal neoplasia: a meta-analysis and trial sequential analysis. Int J Colorectal Dis. 2021 Nov;36(11):2291–303.
6 Hassan C, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. Gastroint Endosc. 2021 Jan;93(1):77–85.e6.
7 Pfeifer L, Neufert C, Leppkes M, Waldner MJ, Hafner M, Beyer A, et al. Computer-aided detection of colorectal polyps using a newly generated deep convolutional neural network: from development to first clinical experience. Eur J Gastroenterol Hepatol. 2021 Dec 1;33:e662–9.
8 Troya J, Fitting D, Brand M, Sudarevic B, Kather JN, Meining A, et al. The influence of computer-aided polyp detection systems on reaction time for polyp detection and eye gaze. Endoscopy. 2022 Feb 14. Epub ahead of print.
9 Ahmad OF, Mori Y, Misawa M, Kudo SE, Anderson JT, Bernal J, et al. Establishing key research questions for the implementation of artificial intelligence in colonoscopy: a modified Delphi method. Endoscopy. 2021 Sep;53(9):893–901.
10 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. Gastroint Endosc. 2020 Oct;92(4):980–4.e4.
11 Spadaccini M, Hassan C, Allarone L, Da Rio L, Maselli R, Carrara S, et al. Comparing the number and relevance of false activations between 2 artificial intelligence computer-aided detection systems: the NOISE study. Gastroint Endosc. 2022 May;95(5):975–81.e1.
12 Riphaus A, Wehrmann T, Hausmann J, Weber B, von Delius S, Jung M, et al. S3-guidelines “sedation in gastrointestinal endoscopy” 2014 (AWMF register no. 021/014). Z Gastroenterol. 2015 Aug;53(8):802–42.
13 Krenzer A, Makowski K, Hekalo A, Puppe F. Semi-automated machine learning video annotation for gastroenterologists. Stud Health Technol Inform. 2021 May 27;281:484–5.
14 Zhou G, Xiao X, Tu M, Liu P, Yang D, Liu X, et al. Computer aided detection for laterally spreading tumors and sessile serrated adenomas during colonoscopy. PLoS One. 2020;15(4):e0231880.
15 Misawa M, Kudo SE, Mori Y, Hotta K, Ohitsuuka K, Matsuda T, et al. Development of a computer-aided detection system for colonoscopy and a publicly accessible large colonoscopy video database (with video). Gastroint Endosc. 2021 Apr;93(4):960–7.e3.
16 Kaminski MF, Thomas-Gibson S, Bugaiksi M, Brethauer 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 Apr;5(3):309–34.
17 Wang P, Xiao X, Glissen Brown JR, Berzin TM, Tu M, Xiong F, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng. 2018 Oct;2(10):741–8.
18 Klare P, Sander C, Prinz M, Haller B, Nowack S, Abdelhafez M, et al. Automated polyp detection in the colorectum: a prospective study (with videos). Gastroint Endosc. 2019 Mar;89(3):576–82.e1.
19 Luo Y, Zhang Y, Liu M, Lai Y, Liu P, Wang Z, et al. Artificial intelligence-assisted colonoscopy for detection of colon polyps: a prospective, randomized cohort study. J Gastroint Surg. 2021 Aug;25(8):2011–8.
20 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 Oct;159(4):1252–61.e5.
21 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(4):343–51.
22 Zhao SB, Yang W, Wang SL, Pan P, Wang RD, Chang X, et al. Establishment and validation of a computer-assisted colonic polyp localization system based on deep learning. World J Gastroenterol. 2021 Aug 21;27(31):5323–46.
23 Hereshbach D, Barrioz T, Lapalus MG, Cウンマラダ T, Bauret P, Potier P, et al. Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies. Endoscopy. 2008 Apr;40(4):284–90.
24 Flaqeye H, Laguier R, Nahon S, Coulom P, Isnard-Bagnis C, Albert-Marty A. Real-life conditions of use of sodium phosphate tablets for colon cleansing before colonoscopy. Endosc Int Open. 2015 Aug;3(4):E346–53.