# Artificial Intelligence in

# *Gastrointestinal Endoscopy*

## Contents

### MINIREVIEWS

# Contents

## ABOUT COVER

Editorial Board Member of *Artificial Intelligence in Gastrointestinal Endoscopy*, Cihangir Akyol, FASCRS, FEBS, MD, Professor, Surgeon, Department of General Surgery, Ankara University, Ankara 06100, Turkey. cihangirakyol@gmail.com

## AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE*, *Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

   *AIGE* mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangio-pancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

## INDEXING/ABSTRACTING

There is currently no indexing.

## RESPONSIBLE EDITORS FOR THIS ISSUE

*MINIREVIEWS*

# Artificial intelligence in polyp detection - where are we and where are we headed?

Kristen E Dougherty, Vatche J Melkonian, Grace A Montenegro

**Kristen E Dougherty, Vatche J Melkonian, Grace A Montenegro,** Department of Surgery, St. Louis University Hospital, Saint Louis, MO 63110, United States

**Corresponding author:** Grace A Montenegro, FACS, MD, MS, Assistant Professor, Surgeon, Department of Surgery, St. Louis University Hospital, 1008 S. Spring Avenue, SLUCare Academic Pavilion, Saint Louis, MO 63110, United States. grace.montenegro@health.slu.edu

## Abstract

The goal of artificial intelligence in colonoscopy is to improve adenoma detection rate and reduce interval colorectal cancer. Artificial intelligence in polyp detection during colonoscopy has evolved tremendously over the last decade mainly due to the implementation of neural networks. Computer aided detection (CADe) utilizing neural networks allows real time detection of polyps and adenomas. Current CADe systems are built in single centers by multidisciplinary teams and have only been utilized in limited clinical research studies. We review the most recent prospective randomized controlled trials here. These randomized control trials, both non-blinded and blinded, demonstrated increase in adenoma and polyp detection rates when endoscopists used CADe systems *vs* standard high definition colonoscopes. Increase of polyps and adenomas detected were mainly small and sessile in nature. CADe systems were found to be safe with little added time to the overall procedure. Results are promising as more CADe have shown to have ability to increase accuracy and improve quality of colonoscopy. Overall limitations included selection bias as all trials built and utilized different CADe developed at their own institutions, non-blinded arms, and question of external validity.

**Key Words:** Neural networks; Computer aided detection; Artificial intelligence in colonoscopy and polyp detection; Artificial intelligence in adenoma detection

**Core Tip:** Use of computer aided detection (CADe) in colonoscopy has been shown to increase polyp and adenoma detection rates compared to standard high-definition colonoscopy with little added procedure time. Additionally, CADe have been built to increase quality of screening colonoscopy. These advantages and features have been

demonstrated in blinded and non-blinded randomized controlled trials.

## INTRODUCTION

Colorectal cancer (CRC) is the second leading cause of cancer death in the United States. Adenomas are the most common type of precancerous polyp. Colonoscopy remains the gold standard for identifying these precancerous polyps and is the only nonsurgical intervention capable of removing them. The National polyp study showed that up to 90% of CRCs are preventable with polyp removal[1]. The adenoma detection rate (ADR) represents the percent of colonoscopies in which at least 1 adenoma is found. ADR is regarded as the main quality indicator of colonoscopy and ideally ADR should equal adenoma prevalence, estimated to be greater than 50%[2]. Unfortunately, ADRs vary widely, with some endoscopists having ADRs as low as 7% [3]. It has been shown that for each 1% increase in ADR, the interval CRC rate was decreased by 3%-6%[3,4]. The main cause of interval CRC incidence is overlooked lesions due to failure of recognition or incomplete mucosal exposure due to suboptimal technique during the withdrawal phase of colonoscopy[5]. Artificial intelligence in colonoscopy was expected to address these issues in hopes to reduce polyp detection miss rates and subsequently interval CRCs[6,7].

## COMPUTER-AIDED DETECTION

The concept of computer-aided detection (CADe) in polyp detection was first described in the early 2000s where software was developed that utilized color and texture to identify polyps[8]. Polyp detection accuracy was as high as 95% however only applicable on static images due to high latency. Tajbakhsh *et al*[9] created CADe based on hybrid shape analysis. The system sensitivity reached close to 90% however proved un-competitive for real time video stream with high latency[9].

CADe of polyps has evolved exponentially since 2012 when deep learning models began utilizing convolutional neural networks (CNNs) to identify polyp-specific features independent of human input.

CNNs utilize statistical pattern recognition algorithms to identify an object, in this case, a polyp. In brief, the computer recognizes an array of numbers, or picture variables, based on pixel analysis of the captured images. The input layer is then filtered through several hidden layers each acting as distinctive feature identifiers or recognizable features. For example, if the desired outcome is for the CNN to recognize a discrete face, hidden layers would include the nose, mouth, eyebrows, *etc.* The fully connected layer comes at the end of this neural network and analyzes the output from previous layers to determine which features correlate best to a certain class, *i.e.*, the probability of the image being Jack's face *vs* Joe's face. The higher the probability of identifying the image in effect strengthens the network[10].

CNNs are created and utilized by multiple disciplines including computer science, bioinformatics, machine learning/intelligent systems, and increasingly in healthcare and medicine. CNNs afford the ability to detect images, in this case polyps, in real time analysis.

Training these networks involves providing a groundwork of data sets or images. Urban *et al*[11] utilized five different data sets: First, data including over one million images of non-medical objects. Second, a set of over 8600 colonoscopy images containing over 4000 images of unique polyps of varying size and morphologies, as well as over 4500 images without polyps. Third, a separate set of 1330 colonoscopy images, half showing unique polyps and half showing other non-polyp images collected from different patients. Fourth, videos of colonoscopies, and fifth, a larger data set augmenting the original set of colonoscopy images. This model identified

polyps with a 96.4% accuracy rate and demonstrated the ability to work in real time conditions with a processing rate of one frame per 10 milliseconds (ms). It identified all polyps discovered by expert viewers (ADRs > 50%) as well as any additional polyps that were missed. The authors believe utilizing real-time live analysis with this model during colonoscopies will prompt increasingly careful inspection and lead to discovery of additional polyps that may have been missed[11].

Five randomized control trials (RCTs) utilizing independent CADe systems are reviewed here demonstrating significant improvement in ADR compared to standard colonoscopy (Table 1).

## RCT

In 2019, Wang *et al*[12] presented the first prospective single center RCT (Sichuan Provincial, China) investigating the influence of an automatic polyp detection method based on deep learning regarding the polyp detection rate and ADR. The study scheme was a non-blinded trial in which subjects who underwent diagnostic colonoscopy with or without assistance of a real-time automated polyp detection system. The primary outcome was ADR. The real-time automatic polyp detection system was based on SegNet architecture. The algorithm was authenticated and had a per image sensitivity of 94.4% and per image specificity of 95.9%. The system handled at least 25 frames per second with a dormancy of 76.8 ± 5.6 ms in simultaneous video analysis. The monitor was parallel with the original endoscopy monitor and it provided simultaneous visual notice and audible alarm when a polyp was detected. Subjects who had colonoscopy from September 2017 to February 2018 were suitable for enrollment and bowel preparation and high definition colonoscope's were standardized. Exclusion criteria included inflammatory bowel disease, CRC, previous unsuccessful colonoscopy, and high suspicion for polyposis syndromes.

Eight endoscopists participated in the study, half of which who were junior endoscopists. The experience was as follows – two seasoned endoscopists (> 20000 colonoscopies), two midlevel endoscopist (3000-10000), and four junior endoscopist (100-500).

Standard colonoscopy was completed in the control group. In the CADe the endoscopist was assisted by the real-time automatic detection system. The system captured the endoscopy video and displayed the polyp location with a blue box on a neighboring screen with a coinciding audible alert. The system was turned on during withdrawal only. The endoscopist was obligated to check every polyp location detected by the system devoid of assistance. A missed polyp was delineated as a polyp confirmed by the endoscopist but unobserved by the system. A false alarm was delineated as detected lesion which was interminably traced by the system deemed by the endoscopist not to be a polyp.

Five hundred thirty-six patients were randomized prospectively into the control group and 522 into the CADe group. There were no statistical differences in demographics, total time of colonoscopy, no polyp withdrawal time or withdrawal time excluding biopsy, bowel preparation and endoscopist experience. There was a statistical difference with withdrawal time of 6.39 min in the routine colonoscopy *vs* 6.89 minutes in the CADe group.

A 1.89-fold increase was found in the mean number of polyps discovered between the two groups [95% confidence interval (CI): 1.63 to 2.192, $P < 0.001$]. The PDR of the control and CADe group were 0.29 and 0.45, respectively [odds ratio (OR), 1.995; 95%CI: 1.532-2.544, $P < 0.001$]. They found a 1.72-fold increase in the mean number of adenomas discovered. The ADR of the control and CADe groups were 0.20 and 0.29, respectively (OR, 1.61; 95%CI: 1.213 to 2.135, $P < 0.001$). The number of detected polyps was significantly higher in the CADe group when looking specifically at non-pedunculated polyps, polyps 0 cm to 1 cm in size and polyps in all portions of the colon. There was also a considerably higher number of adenomas detected in the CADe group when looking at non-pedunculated polyps, polyps smaller than 0.5 cm and polyps in all portions of the colon except for the cecum and ascending colon. There was a total of 39 false positives in the CADe group. Of discovered polyps in the CADe cohort, none were missed by the automatic system.

ADR in the CADe group showed a trend of 6% increase in the subgroup of patients with excellent bowel preparation. In addition, their results, including the mean number of detected adenomas, mean number of detected polyps and PDR, were significantly increased. However, this was not statistically significant given the small sample size.

**Table 1 Five randomized control trials utilizing independent computer aided detection systems**

| Ref. | Study design | Blinded? | Type of system | ADR control | ADR CADe | False alarms (per colonoscopy) | Missed polyps |
|---|---|---|---|---|---|---|---|
| Wang *et al*[12], 2019 | RCT | No | CADe | 0.2 | 0.29 ($P < 0.001$) | 0.075 | 0 |
| Repici *et al*[13], 2020 | RCT | No | CADe | 0.404 | 0.548 ($P < 0.001$) | - | - |
| Liu *et al*[14], 2020 | RCT | No | CADe | 0.23 | 039 ($P < 0.001$) | 0.071 | 0 |
| Wang *et al*[2], 2020 | RCT | Yes | CADe | 0.28 | 0.34 ($P = 0.03$) | 0.1 | 0 |
| Su *et al*[15], 2020 | RCT | No | CADe + Quality | 0.165 | 0.289 ($P < 0.001$) | 0.2 | 0 |

RCT: Randomized control trial; CADe: Computer aided detection.

Limitations of this study include the inability to blind the endoscopists of each arm. In addition, the adenoma and polyp detection rates in this study are substantially lower than what is reported in Western countries, and thus there is a question of whether this study is applicable in centers with higher ADRs at baseline[12].

Repici *et al*[13] published work on a separate CNN by the GI genius, Medtronic system in 2020. The system was trained and validated with 99.7% per lesion sensitivity and 0.9% false-positive frames. Using a series of videos of 2684 histologically confirmed polyps from 840 patients. They performed a multicenter randomized trial to assess the safety and efficacy of this CADe in detection of colorectal neoplasia during real-time colonoscopy. Like Wang *et al*[12], the operator was not blinded to the study arm assigned to each patient. Colonoscopies were performed by 6 experienced endoscopist to from each center with over 2000 screening colonoscopies; inexperienced endoscopists were not included. High definition colonoscopes were utilized. The CADe system would signal the endoscopist with a bounding box only when a target polyp was recognized in the image. Primary outcome was ADR. Secondary outcomes were proximal ADR, total number of polyps detected, sessile serrated lesions detection rate, mean number of adenomas per colonoscopy (APC), cecal intubation rate and withdrawal time.

Patient's undergoing colonoscopy from September to November 2019 were included. Colonoscopy requirements included colorectal screening or post poly-pectomy surveillance as well as work-up following FIT positivity or patients with appropriate signs and symptoms warranting further work up. Patients were excluded in the case of personal history of CRC or Inflammatory bowel disease (IBD), previous colon resection, or antithrombotic therapy precluding polyp resection.

A total of 685 patients were randomized, 341 in CADe arm and 344 in the control arm. There was no significant difference in terms of bowel preparation or cecal intubation rate. ADR in the CADe group was 54.8% *vs* 40.4% in the control group. After adjusting for age, gender, and indication the ADR was significantly higher in the CADe group compared to the control.

The CADe group identified more non-polypoid (26.6% *vs* 18.4%) and polypoid (37.3% *vs* 26.5%) lesions compared to control. The proportion of patients with < 10 mm adenomas was higher in the CADe group, 44.3%, *vs* in the control group, 32.3%. The difference between the 2 arms was significant for both ≤ 5 mm and 6 mm to 9 mm adenomas. Regarding location, the proportion of patients with proximal adenomas was higher in the CADe group then in the control group. This was also true for distal adenomas. Forty-five patients were diagnosed with advanced neoplasia in the CADe group compared with 36 in control group, demonstrating a detection rate for advanced neoplasia of 13.3% and 10.5% respectively.

Of the 460 patients who underwent polyp resection, 120 did not have histologically proven adenomas, sessile serrated lesions, or CRC. The non-neoplastic resection rate for CADe and control were 26% and 28.8%, respectively.

Repici *et al*[13] demonstrated that addition of real-time CADe to colonoscopy resulted in 30% and 46% relative increase in ADR and APC. Safety of the system was demonstrated by the lack of increase of both useless resections and withdrawal time. Computer aided detection efficacy appeared to be independent of morphology and location of neoplasia and was mainly explained by the additional detection of polyps that were less than 5 mm, or between 6 mm to 9 mm in size. Limitations of this study

were like Wang *et al*[12] in that the endoscopists in each arm were not blinded. In addition, they did not include inexperienced endoscopists in their study. They demonstrated the safety and efficacy of integrating a CADe with colonoscopy with a substantial improvement of ADR and adenoma per colonoscopy without increasing the removal of non-neoplastic lesions. This is likely to improve the quality of colonoscopy without affecting efficiency[13].

In 2020, Liu *et al*[14] published their work using yet another CNN or CADe system. Polyp positive videos (151) and polyp negative videos (384) were used to design the system. This system utilized spatiotemporal data to recognize polyps, which is presumed to be more suitable for video data sets. This was a prospective, single center, randomized control study (China) to demonstrate the effective of CADe on the detection rate of polyps and adenomas during colonoscopy. Bowel preparations and high definition colonoscope's were standardized. Exclusion criteria included inflammatory bowel disease, history of CRC surgery, history of radiotherapy and/or chemotherapy and biopsy contraindications. CADe was only utilized during withdrawal phase. The system processed each frame and displayed the detected polyp. When the lesion would appear on the screen a voice alarm would prompt endoscopist to view the system. This study was done without the assistance of nurses, trainees, or staff. Polyps identified by the endoscopist but not identified by the CADe system were deemed "missed polyps" and were documented. False alarms were defined as lesions detected and continuously tracked but were not identified as polyps by the system.

A total of 1026 patients were eligible: 518 in the control and 508 in the computer-aided detected group. The two groups were similar in demographics and risk factors. Total withdrawal time in the standard group *vs* the CADe group was 6.74 min and 6.82 min, respectively ($P < 0.001$).

A total of 734 polyps were detected. Three hundred and ninety-two of these were adenomas representing 53.41%, and 31 were sessile serrated adenomas representing 4.22%. In total 248 polyps were detected in the control group and 486 polyps in the CADe group, a rate of 33.79% *vs* 66.21%, respectively. The corresponded to 1.53 times increase in the average number of polyps detected in both (95%CI: 1.652-2.297, $P < 0.001$). The polyp detection rates in control and CADe were 0.28 and 0.44, respectively which corresponded to 1.51 times increase in number of adenomas detected (95%CI: 1.423-2.016, $P < 0.001$). The detection rates of adenomas in control *vs* CADe group were 0.23 and 0.39, respectively (CI: 1.201- 2.220, $P < 0.001$). The number of polyps detected in the CADe group was significantly higher than control group when looking at sessile polyps, polyps 0 cm to 1 cm in size, and polyps in all portions of the colon. The number of adenomas detected in the CADe group also increased significantly when considering sessile polyps, polyps ≤ 0.5 cm, and polyps in all parts of the colon excluding the cecum and ascending colon.

Similar to Wang *et al*[12], detection rate in the CADe was higher when intestinal preparation was deemed adequate. Insufficient sample size of the subgroup analysis failed to show statistical significance. There were 36 false alarms in the CADe group corresponding to an average of 0.071 false alarms per colonoscopy. Of all polyps detected in the CADe group no polyps were missed. In mirroring the results of previous RCTs, this study again demonstrated significantly higher detection rates of adenoma, and average number of polyps and adenomas by colonoscopy in the CADe group when contrasted to control groups. However, the overall rise in adenoma detection was mainly due to the rise in detection of small adenomas, less than 1 cm in most instances.

The study review shows that integration of computer aided detection systems can effectively detect polyps that were otherwise missed by the endoscopist, however there is a blind area with polyps that remain undetected which remains an unanswered problem. Similar to Wang *et al*[12] study limitations include non-blinded endoscopists and low ADRs of endoscopists, as compared to Western countries. In conclusion the study showed this CADe increases the detection rate of colorectal polyps and adenomas, therefore depicting its feasibility for detection of polyps and adenomas on colonoscopy[14].

The first single center, randomized, double-blind trial to evaluate the use of automatic polyp detection using the CAD system during colonoscopy was published in early 2020 by Wang *et al*[12]. They enrolled consecutive patients between September 2018 and January 2019. After all exclusion criteria, there were 484 patients in the CADe group and 478 in the sham group. All qualified patients were randomized 1:1 to either white light colonoscopy with CADe assistance or to the control group consisting of white light colonoscopy with a mock system. Patients were not notified of their assignment and blinding of the operating endoscopists was achieved by the mock

system. Four senior endoscopists participated, each with at least 5 years' experience, completing a minimum of 1000 colonoscopy procedures per year.

Endoscopists were told to perform all colonoscopy procedures with the aid of a CADe system, and they were unaware of the use of the mock system. Authors utilized the same CADe system as previously discussed in their non-blinded trial. A mock system was designed to appropriately mask the endoscopists. This system simulated alert boxes on polyp-like non polyp structures without tracking actual polyps during colonoscopy. The sham model was built using a portion of the images used to develop the CADe system as previously described, producing a much lower sensitivity and specificity. The grouping and outputs of the mock and computer aided detection systems were only observable to a senior endoscopist using a separate monitor acting as a second observer.

In both study groups, if the operating endoscopist did not recognize an abnormality within an alert box, the observer was tasked with informing the location of any alert box for the operating endoscopist using a laser pointer on the principal monitor. An alarm sounded to the observer *via* earpiece if an alert box was visible. The observer was not blinded to the intervention and aware which system was being used. All observable alert boxes were documented by the observer, however only the alert boxes that were on non-polyp structures were recorded as consistent false detections in the CADe, *i.e.*, false alarms. Consistent false detections of the CADe were recorded and believed not to be a polyp by the operating endoscopist. In the CADe group the observer also recorded any missed detections, defined as a polyp discovered by the working endoscopist and proved by histology but not alerted by the CADe.

After the clinical analysis, they analyzed the videos of polyps that were detected by CADe, but initially missed by the operating endoscopist. The video clips were then independently reviewed by an addition 3 skilled, experienced endoscopists who did not partake in the clinical trial. Endoscopists labeled each video when they identified a polyp, and they were limited to single viewing of the videos. Analysis of sensitivity and specificity on these easily overlooked polyps was performed.

The primary outcome was proportion of individuals who underwent a complete colonoscopy and had 1 or more adenomas detected. Secondary outcomes were the proportion of individuals undergoing a complete colonoscopy who had 1 or more polyps detected, the number of polyps per colonoscopy, and the number of APC, which was calculated by dividing the total number of polyps that are adenomas detected by the total number of colonoscopies done.

Four hundred seventy-eight patients were allocated to sham group and 484 to the CADe. No difference was in terms of demographics and adenoma detection probability features. There were no recorded untoward events with these procedures. There was a statistically significant increase in withdrawal time with the CADe group, 7.46 min *vs* 6.99 min ($P < 0.0001$).

More biopsies were performed for polyps in the CADe group than in the mock group. When omitting the time taken to do biopsies, the mean withdrawal time was not statistically significant between the groups. Overall, 809 polyps were detected, of which 38% were found in the mock group and 62% were found in the CADe group. Of these polyps, 57% were adenomas and 4% were sessile serrated adenomas. When considering shape (sessile) and size (0-5 mm), the CADe group had a significantly higher number of detected polyps and adenomas.

Notably, there was a 1.61-fold increase in polyps detected per colonoscopy between the 2 groups (95%CI: 1.39-1.85; $P < 0.0001$). The PDR was significantly higher in the CADe group then with the mock system, 52% *vs* 37%. A 1.53-fold increase in APC between the groups (95%CI: 1.27-1.85; $P < 0.0001$). The ADR was significantly higher with the CADe system compared to the mock system with 28% in the sham control group and 34% in the CADe group having an adenoma detected. Based on the observers' judgement, here were 48 false detections in the CADe group, averaging 0.1 per colonoscopy. Of all the detected polyps in the experimental group, none were missed by the CADe system.

An average of 0.17 adenomas and 0.33 polyps per patient were overlooked initially by the endoscopist in the CADe group. These polyps were small (mean adenoma size 3.89 mm), isochromatic, flat in shape, had unclear boundaries, were partly behind colon folds, and were on the edge of the visual field. The sensitivity and specificity of three skilled endoscopist during review of the endoscopy videos was 17% and 64%, respectively.

Once again, Wang *et al*[12] demonstrated a CADe system can effactually increase the number of polyps and adenomas detected with colonoscopy, and after controlling for operational bias. The CADe system had higher sensitivity and specificity for detection of easy to overlook polyps compared to evaluation based solely on the utilities of the

human endoscopist.

The main contribution of CADe in this system demonstrated a rise detection of diminutive and non-pedunculated, non-advanced adenomas and hyperplastic polyps. The CADe system is safe and effectual approach to increase ADR during colonoscopy [2].

In 2020, Su *et al*[15] published a prospective randomized control study comparing CADe with control, however, their CADe was built with quality control, specifically supervising withdrawal stability, from five different neural networks (AQCS). Similar to the other trials mentioned, AQCS was turned on during withdrawal. In addition to visual cue when polyps detected, the added features in this system were notifying the endoscopist to slow down withdrawal speed and re-examine colonic segments when unstable or blurry frames were detected continuously and prompting endoscopist to clean mucosa when inadequate score (Boston Prep Score < 2) was given by system. Study was single center in Qilu Hospital in China, from October 2018 to May 2019. Study included patients over 18yo who were able to give consent for screening colonoscopy. Exclusion criteria included history of IBD, CRC or colorectal surgery, patients with previously failed colonoscopy, or highly suspicious for polyposis syndromes, or patients whose colonoscopy could not be completed due to stenosis or large occupying lesions. Six endoscopists participated, each with 4-6 years of experience and colonoscopy volume of 5000 to 8000. Endoscopists were not blinded to randomization status, however, patients, data collection and study analyses were blinded.

A total of 659 patients were randomized, after exclusions, 308 in AQCS group and 315 in control. There were no differences in demographics between the two groups. They performed retrospective review of 459 normal colonoscopies without positive findings performed by participating colonoscopists and there was no significant difference in the mean withdrawal times between the two groups.

A total number of 169 adenomas were detected, 56 control and 113 AQCS group. ADR in control 16.51% *vs* 28.9% AQCS group (OR, 2.055; 95%CI: 1.397-3.024; $P <$ 0.001). Additionally, there were more adenomas detected in AQCS when non pedunculated, diminutive adenomas (≤ 5 mm), larger adenomas (> 5 mm), and adenomas in all segments with exception of cecum and rectum were considered.

Polyp detection rates were also higher in the AQCS group, 38.1% *vs* 25.4% in control (OR, 1.824; 95%CI: 1.296-2.569; $P = 0.001$).

Withdrawal time excluding biopsy time was significantly longer in the AQCS group than in the control group (7.03 ± 1.01 min *vs* 5.68 ± 1.26 min, $P < 0.001$). For the AQCS group there was a significant improvement in withdrawal time in this study compared to their retrospective withdrawal times before this study. Adequate bowel preparation rate was 87.3% in the AQCS group *vs* 80.6% in control group, (OR, 1.656; 95%CI: 1.070-2.564; $P = 0.023$).

There were 62 false prompts (false positives), averaging 0.201 false prompts per colonoscopy and no missed prompts (false negatives) in the AQCS group.

Su *et al*[15] demonstrated significant improvement in ADR when utilizing a quality-controlled CADe system that supervises in withdrawal stability and prompts endoscopists to clean colonic mucosa when inadequate prep scores are recognized. The AQCS demonstrated significant increase in ADR and an increase detection of larger adenomas, compared to previously mentioned CADe systems.

Limitations of this study are similar to prior RCTs where endoscopists are not blinded to randomization. It is also a single center study with only experienced endoscopists participating. Authors mention that the system utilized 4 intra-procedural quality metrics, and these combined, improved ADR. They did not perform preliminary testing to evaluate whether 2 or 3 metrics would increase ADR to standard colonoscopy.

## CONCLUSION

The goal of computer aided detection of polyps and adenomas is to close the gap between ADR and adenoma prevalence and in turn reduce interval CRC rates. CADe systems could act as second observers and reduce miss-rates of polyps. Implementation of CNNs for image recognition has overhauled the playing field regarding artificial intelligence utilization in colonoscopy, as these networks are built to allow image recognition in real time. As mentioned above, multiple CADe systems are being built and programmed by multidisciplinary teams from bioinformatics, computer science, machine learning/intelligent systems and in medicine. A single system has

not been shown to be superior to others.

As demonstrated by randomized trials, the ability to integrate CADe with colonoscopy in real time has demonstrated overall ADR and polyp detection rates were significantly higher for CADe groups compared with control. These were most significant for small, diminutive polyps and adenomas ≤ 5 mm and those which were sessile in character. These CADe systems have been shown to be safe and efficient. The CADe systems mentioned here have scarce miss rates, if any, when it comes to polyp detection. Small adenomas have less probability for malignant transformation compared to larger, however the increase in total ADR may contribute to decreased risk of interval CRC. Conversely, increased detection and resection of diminutive lesions may represent additional unnecessary polypectomies and add to workload, cost, and pathology resource utilization. Wang *et al*[12] remarks that gaining the ability to identify small adenomas will provide the advantage of resecting small pre-malignant lesions along with distinguishing patients who are at higher risk for future adenomas and interval cancer. Sue *et al*[15]'s AQCS network demonstrated increase detection of larger adenomas. Their system was unique to others in that it was built with a quality control feature that essentially improved the quality of colonoscopy by improving withdrawal time and identifying inadequate exposure of mucosa. Detection of polyps and adenomas by CADe relies on exposure of the entirety of the colonic mucosal field by the endoscopist. Polyps that remain outside of the visual field still pose a major deficiency that have the potential to be addressed by this system.

Endoscopists are still responsible for proper execution of the colonoscopy procedure, including cecal intubation rate. Inexperienced endoscopists are likely to have suboptimal results in the technical exposure of colorectal mucosa and perhaps adding quality control to CADe is the answer as demonstrated by Su *et al*[15]. Additionally, the AQCS was not utilized by inexperienced endoscopists in that RCT.

Artificial intelligence in colonoscopy has certainly made strides over the last decade, specifically in real time detection. Currently, CADe systems based on CNN for the use of polyp and adenoma detection during colonoscopy are being built in single centers. This poses a risk of selection bias leading to difficulty implementing any one CADe system on a wide scale. Appropriately curated large scale data sets are needed to limit data set bias. Collection of image and video inputs should be broad and include unsampled or under-represented lesions. The added complexity of developing CADe to assist in withdrawal stability and identification of inadequate exposure elevates the technology of AI as these enhance the ability and accuracy of the endoscopist who remains the critical portion of the colonoscopy, for now[15-17]. In addition to polyp detection, models built to aid in diagnosis and classification of inflammatory bowel disease have been described[18,19]. Current systems should have controlled and practical set up, as not add to the workflow of standard colonoscopy. Ideally these systems should predict pathology and size and improve accuracy, minimizing unnecessary pathologic assessment and avoidable resection of non-neoplastic lesions. While it is expected that technologic cost will increase initially, when used effectively and efficiently, CADe systems should ultimately reduce cost. The review of RCTs demonstrates undeniable improvement of ADR when utilizing CADe compared to standard colonoscopy. Collectively they demonstrate CADe are safe and practical when used in real-time and more complex CADe systems have the potential to improve accuracy of the endoscopist improving quality of colonoscopy.

## REFERENCES

1    **Winawer SJ**, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, Waye JD, Schapiro M, Bond JH, Panish JF. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med* 1993; **329**: 1977-1981 [PMID: 8247072 DOI: 10.1056/NEJM199312303292701]

2    **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. 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**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

3    **Corley DA**, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 1298-1306 [PMID: 24693890 DOI: 10.1056/NEJMoa1309086]

4    **Kaminski MF**, Wieszczy P, Rupinski M, Wojciechowska U, Didkowska J, Kraszewska E, Kobiela J, Franczyk R, Rupinska M, Kocot B, Chaber-Ciopinska A, Pachlewski J, Polkowski M, Regula J. Increased Rate of Adenoma Detection Associates With Reduced Risk of Colorectal Cancer and Death.

*Gastroenterology* 2017; **153**: 98-105 [PMID: 28428142 DOI: 10.1053/j.gastro.2017.04.006]

5    **Zhao S**, Wang S, Pan P, Xia T, Chang X, Yang X, Guo L, Meng Q, Yang F, Qian W, Xu Z, Wang Y, Wang Z, Gu L, Wang R, Jia F, Yao J, Li Z, Bai Y. Magnitude, Risk Factors, and Factors Associated With Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology* 2019; **156**: 1661-1674.e11 [PMID: 30738046 DOI: 10.1053/j.gastro.2019.01.260]

6    **Hassan C**, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, Antonelli G, Yu H, Areia M, Dinis-Ribeiro M, Bhandari P, Sharma P, Rex DK, Rösch T, Wallace M, Repici A. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021; **93**: 77-85.e6 [PMID: 32598963 DOI: 10.1016/j.gie.2020.06.059]

7    **Bisschops R**, East JE, Hassan C, Hazewinkel Y, Kamiński MF, Neumann H, Pellisé M, Antonelli G, Bustamante Balen M, Coron E, Cortas G, Iacucci M, Yuichi M, Longcroft-Wheaton G, Mouzyka S, Pilonis N, Puig I, van Hooft JE, Dekker E. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. *Endoscopy* 2019; **51**: 1155-1179 [PMID: 31711241 DOI: 10.1055/a-1031-7657]

8    **Maroulis DE**, Iakovidis DK, Karkanis SA, Karras DA. CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames. *Comput Methods Programs Biomed* 2003; **70**: 151-166 [PMID: 12507791 DOI: 10.1016/s0169-2607(02)00007-x]

9    **Tajbakhsh N**, Gurudu SR, Liang J. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information. *IEEE Trans Med Imaging* 2016; **35**: 630-644 [PMID: 26462083 DOI: 10.1109/TMI.2015.2487997]

10   **Lee H**, Grosse R, Ranganath R, Ng A.   Convolutional Deep Belief networks for scalable Unsupervised learning of hierarchical representations. ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning. Association for Computing Machinery, 2009: 609-616 [DOI: 10.1145/1553374]

11   **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078.e8 [PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037]

12   **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

13   **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology* 2020; **159**: 512-520.e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

14   **Liu WN**, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, Huang J. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; **26**: 13-19 [PMID: 31898644 DOI: 10.4103/sjg.SJG_377_19]

15   **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. 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**: 415-424.e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

16   **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

17   **Karnes WE**, Ninh A, Dao T, Requa J, Samarasena JB. Real-time identification of anatomic landmarks during colonoscopy using deep learning. *Gastrointest Endosc* 2018; **87**: AB252 [DOI: 10.1016/j.gie.2018.04.447]

18   **Mossotto E**, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep* 2017; **7**: 2427 [PMID: 28546534 DOI: 10.1038/s41598-017-02606-2]

19   **Karnes WE**, Ninh A, Dao T, Requa J, Samarasena JB. Unambiguous real-time scoring of bowel preparation using artificial intelligence. *Gastrointest Endosc* 2018; **87**: AB258 [DOI: 10.1016/j.gie.2018.04.461]