OPEN ACCESS

ORIGINAL ARTICLE

Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study

Pu Wang,[1] Tyler M Berzin,[2] Jeremy Romek Glissen Brown,[2] Shishira Bharadwaj,[2] Aymeric Becq,[2] Xun Xiao,[1] Peixi Liu,[1] Liangping Li,[1] Yan Song,[1] Di Zhang,[1] Yi Li,[1] Guangre Xu,[1] Mengtian Tu,[1] Xiaogang Liu[1]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2018-317500).

[1]Department of Gastroenterology, Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital, Chengdu, China
[2]Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA

Correspondence to
Xiaogang Liu, Department of Gastroenterology Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu, China; Gary.samsph@gmail.com

Received 30 August 2018
Revised 4 February 2019
Accepted 13 February 2019
Published Online First 27 February 2019

ABSTRACT

Objective The effect of colonoscopy on colorectal cancer mortality is limited by several factors, among them a certain miss rate, leading to limited adenoma detection rates (ADRs). We investigated the effect of an automatic polyp detection system based on deep learning on polyp detection rate and ADR.

Design In an open, non-blinded trial, consecutive patients were prospectively randomised to undergo diagnostic colonoscopy with or without assistance of a real-time automatic polyp detection system providing a simultaneous visual notice and sound alarm on polyp detection. The primary outcome was ADR.

Results Of 1058 patients included, 536 were randomised to standard colonoscopy, and 522 were randomised to colonoscopy with computer-aided diagnosis. The artificial intelligence (AI) system significantly increased ADR (29.1%vs20.3%, p<0.001) and the mean number of adenomas per patient (0.53vs0.31, p<0.001). This was due to a higher number of diminutive adenomas found (185vs102; p<0.001), while there was no statistical difference in larger adenomas (77vs58, p=0.075). In addition, the number of hyperplastic polyps was also significantly increased (114vs52, p<0.001).

Conclusions In a low prevalent ADR population, an automatic polyp detection system during colonoscopy resulted in a significant increase in the number of diminutive adenomas detected, as well as an increase in the rate of hyperplastic polyps. The cost–benefit ratio of such effects has to be determined further.

Trial registration number ChiCTR-DDD-17012221; Results.

Significance of this study

What is already known on this subject?

- Colorectal adenoma detection rate (ADR) is regarded as a main quality indicator of (screening) colonoscopy and has been shown to correlate with interval cancers. Reducing adenoma miss rates by increasing ADR has been a goal of many studies focused on imaging techniques and mechanical methods.
- Artificial intelligence has been recently introduced for polyp and adenoma detection as well as differentiation and has shown promising results in preliminary studies.

What are the new findings?

- This represents the first prospective randomised controlled trial examining an automatic polyp detection during colonoscopy and shows an increase of ADR by 50%, from 20% to 30%.
- This effect was mainly due to a higher rate of small adenomas found.
- The detection rate of hyperplastic polyps was also significantly increased.

How might it impact on clinical practice in the foreseeable future?

- Automatic polyp and adenoma detection could be the future of diagnostic colonoscopy in order to achieve stable high adenoma detection rates.
- However, the effect on ultimate outcome is still unclear, and further improvements such as polyp differentiation have to be implemented.

INTRODUCTION

Colorectal cancer (CRC) is the second and third-leading causes of cancer-related deaths in men and women respectively.[1] Colonoscopy is the gold standard for screening CRC.[2 3] Screening colonoscopy has allowed for a reduction in the incidence and mortality of CRC via the detection and removal of adenomatous polyps.[4–8] Additionally, there is evidence that with each 1.0% increase in adenoma detection rate (ADR), there is an associated 3.0% decrease in the risk of interval CRC.[9 10] However, polyps can be missed, with reported miss rates of up to 27% due to both polyp and operator characteristics.[11 12]

Unrecognised polyps within the visual field is an important problem to address.[11] Several studies have shown that assistance by a second observer increases the polyp detection rate (PDR), but such a strategy remains controversial in terms of increasing the ADR.[13–15]

Ideally, a real-time automatic polyp detection system, with performance close to that of expert endoscopists, could assist the endoscopist in detecting lesions that might correspond to adenomas in a more consistent and reliable way

Check for updates

To cite: Wang P, Berzin TM, Glissen Brown JR, *et al*. *Gut* 2019;**68**:1813–1819.

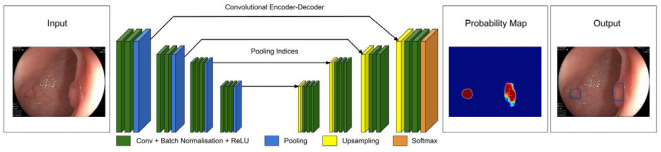

Figure 1 Deep learning architecture. The detection algorithm is a deep convolutional neural network (CNN) based on SegNet architecture. Data flow is from left to right: a colonoscopy image is sequentially warped into a binary image, with 1 representing polyp pixels and 0 representing no polyp in probability a map. This is then displayed, as showed in the output, with a hollow tracing box on the CADe monitor.

than a human assistant. Though several automatic polyp detection systems have been developed over the past decade,[16 17] evidence on the ability of this technology to locate and trace polyps in clinical practice during live colonoscopy is lacking.

The aim of this study was to investigate whether a high-performance real-time automatic polyp detection system[18] can increase polyp and adenoma detection rates in the real clinical setting.

MATERIALS AND METHODS

Study design

Real-time automatic polyp detection system

The real-time automatic polyp detection system (Shanghai Wision AI Co., Ltd.) was developed on a deep learning architecture (figure 1) with the aid of endoscopists and modelers.[18]

In a preliminary study from our group, the algorithm was validated and found to have a per-image sensitivity of 94.38% per-image specificity of 95.92% and an area under the receiver operating characteristic curve of 0.984. By using a multithreaded processing system, the system processed at least 25 frames per second with a latency of 76.80±5.60 ms in real-time video analysis.[18] Furthermore, the detection delay was hardly noticeable for endoscopists. The system monitor was fixed adjacent and parallel with the original endoscopy monitor.

Prospective comparative study

The prospective study was designed as a randomised controlled trial to investigate the impact of an automatic polyp detection system acting as an assistant to the endoscopist on PDR and ADR. This study was conducted in the Endoscopy Center of the Sichuan Provincial People's Hospital, China. Consecutive patients who underwent a colonoscopy from September 2017 to February 2018 were eligible for enrolment (figure 2). Routine bowel preparation consisted of 4 L of polyethylene glycol, given in split doses. Colonoscopies were performed with high definition colonoscopes (Olympus CF-Q260, CF-H260) and high-definition monitors. We excluded patients with a history of inflammatory bowel disease (IBD), CRC, colorectal surgery and patients with a contraindication for biopsy. Patients with prior failed colonoscopy and high suspicion of polyposis syndromes, IBD and typical advanced CRC were also excluded. Basic demographic characteristics including gender, age, indication for colonoscopy, procedure time (morning/afternoon), type of sedation, risk factors of colon polyps including diabetes mellitus, coronary artery disease, body mass index, family history of colon adenoma or cancer, use of aspirin or other nonsteroidal anti-inflammatory drugs, metformin, folic acid, calcium or hormone replacement therapy and alcohol and tobacco use were recorded before colonoscopy. Any complication during the procedure or recovery were also recorded by the staff assistant. Eight physicians from the division of gastroenterology participated in the study, including two senior endoscopists (>20 000 colonoscopies), two midlevel endoscopists (between 3000 and 10 000 colonoscopies) and four junior endoscopists (between 100 and 500 colonoscopies). Each patient was prospectively randomised into two groups by the staff assistant using a digital random number generator before the colonoscopy procedure. In the control group, a routine colonoscopy was performed. In the research group (computer-aided detection (CADe) group), the real-time automatic polyp detection system was used to assist the endoscopist. The system was connected to the endoscopy generator, and the video stream was captured synchronously. Furthermore, the system processed each frame and displayed the detected polyp

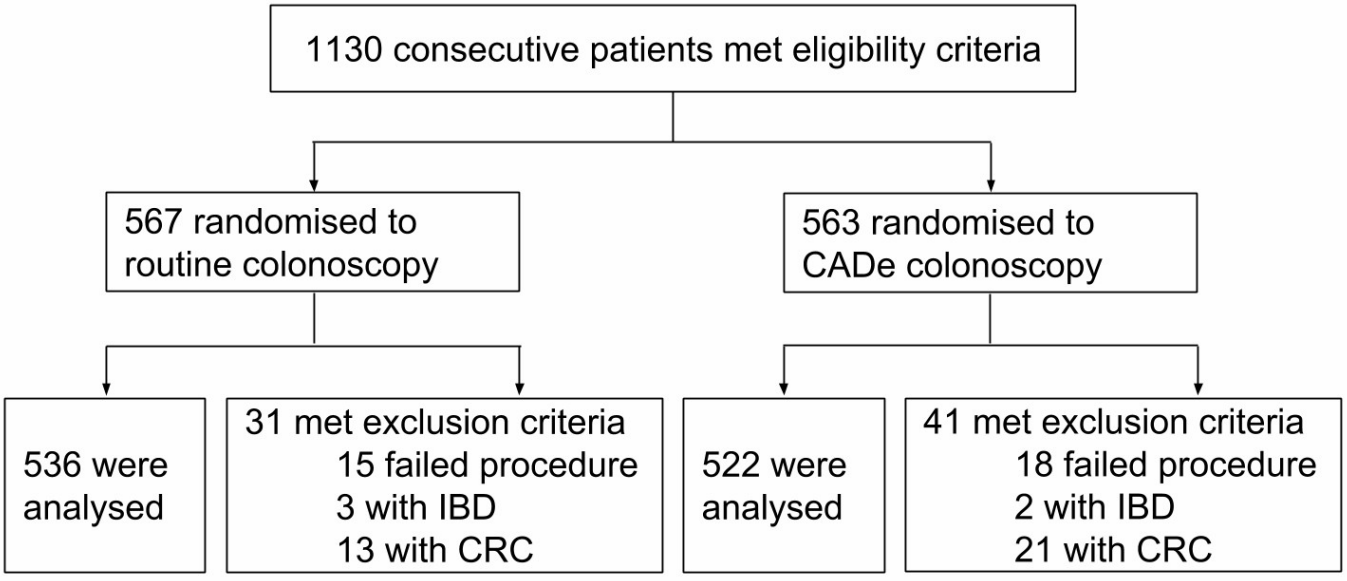

Figure 2 Flow Diagram of Enrollment. 1130 patients met the inclusion criteria, and 72 of them were exclude because of failed procedures, diagnosed IBD or CRC. CRC, colorectal cancer; IBD, inflammatory bowel disease.

Table 1 Baseline information

Characteristics	Routine colonoscopy (n=536)	CADe colonoscopy (n=522)	P* value
Age, mean (SD)	49.94 (±13.79)	51.07 (±13.15)	0.16
BMI, mean (SD)	23.02 (±3.14)	23.03 (±2.93)	0.99
Withdrawal time, mean (SD)	6.39 (±1.21)	6.89 (±1.79)	<0.001
Total time, mean (SD)	12.10 (±4.08)	12.52 (±4.38)	0.11
Insertion time, mean (SD)	5.71 (±3.90)	5.63 (±4.03)	0.75
No polyp withdrawal time, mean (SD)	6.13 (±1.07)	6.17 (±0.94)	0.58
Withdrawal time excluding biopsy, mean (SD)	6.07 (±1.11)	6.18 (±1.38)	0.15
Indication			0.74
Screening, n (%)	44 (8.21)	40 (7.66)	
Symptomatic, n (%)	492 (91.79)	482 (92.34)	
Sex			0.19
Female, n (%)	287 (53.54)	259 (49.62)	
Male, n (%)	249 (46.46)	263 (50.38)	
BMI category			0.5
<25, n (%)	408 (76.12)	403 (77.20)	
25<=BMI<30, n (%)	117 (21.83)	113 (21.65)	
≥30, n (%)	11 (2.05)	6 (1.15)	
Procedure time			0.61
AM, n (%)	280 (52.24)	264 (50.57)	
PM, n (%)	256 (47.76)	258 (49.43)	
Endoscope			0.6
CF-H260, n (%)	156 (29.10)	160 (30.65)	
CF-Q260, n (%)	380 (70.90)	362 (69.35)	
Anaesthesia†			0.95
No, n (%)	68 (12.69)	67 (12.84)	
Yes, n (%)	468 (87.31)	455 (87.16)	
Boston Score, mean (SD)	6.64 (1.25)	6.64 (1.21)	0.95
Boston Score Rank			0.72
Inadequate (sum <6.0 or anyone <2.0), n (%)	79 (14.74)	73 (13.98)	
Adequate (sum ≥6.0 and everyone ≥2.0), n (%)	457 (85.26)	449 (86.02)	
Endoscopist experience			0.33
Senior, n (%)	233 (43.47)	203 (38.89)	
Midlevel, n (%)	236 (44.03)	251 (48.08)	
Junior, n (%)	67 (12.50)	68 (13.03)	

No polyp withdrawal time means withdrawal time during those colonoscopies where no polyp was detected or removed.
*P value from χ^2 test (or Fisher's exact test, as appropriate) or t-test.
†Anaesthesia was administered with midazolam, fentanyl by an anaesthesiologist who monitored the patients for complications.
BMI, body mass index; NSAID, non-steroidal anti-inflammatory drug.

location with a hollow blue tracing box on an adjacent monitor with a simultaneous sound alarm (figure 1) (see online supplementary file 1). The system was turned on during withdrawal only. The endoscopist focused mainly on the main monitor during the procedure and was prompted to look at the system monitor by the sound alarm. The endoscopist was required to check every polyp location detected by the system. This was performed without the assistance of nurses, trainees or staff assistants. In both groups, the staff assistant recorded the type of colonoscope used (CF-H260/CF-Q260), the insertion time, withdrawal time and Boston Bowel Preparation Scale (BBPS) as described by the endoscopist. When a polyp was detected, the nurse assisted in performing cold forceps biopsy for histology and the staff-assistant recorded the location, size and morphological features according to the Paris classification. In the CADe group, missed polyps by the system and system false alarms were also recorded. A missed polyp was defined as a polyp confirmed by the endoscopist but undetected by the system. A false alarm was defined as a detected lesion, which was continuously traced by the system, deemed by the endoscopist not to be a polyp. It was registered with Chinese Clinical Trial Registry.

All authors had access to the study data and reviewed and approved the final manuscript.

Statistical analysis

We prospectively designed this study to allow for 80% power or more to detect a 10% difference (30% vs 20%), in adenoma detection rate, between colonoscopy procedures with a two-group χ^2 test with a two-sided α level of 0.05. A sample size of 702 participants was needed, and the overall participant enrolment goal was 1130 to allow for potential exclusions or dropouts.

Statistical analysis was performed with R studio V.3.4.0 or higher. Comparison of baseline clinical and demographic characteristics between the CADe and the control group were performed using the χ^2 test for categorical variables and using the two-sample t-test for continuous variables. Regarding the

Table 2 Polyp and adenoma characteristics

Characteristics	Routine colonoscopy (n=269)	CADe colonoscopy (n=498)	P value*
Pathology			0.173
Carcinoma, n (%)	0 (0.00)	0 (0.00)	1.000
Sessile serrated adenoma/polyp (SSAP), n (%)	14 (5.20)	17 (3.41)	0.541
Adenoma, n (%)			
Advanced adenoma, n (%)	16 (5.95)	17 (3.41)	0.803
Others, n (%)	144 (53.53)	245 (49.20)	<0.001
Benign lesions, n (%)			
Hyperplastic and Inflammatory	94 (34.94)	217 (43.57)	<0.001
Hamartoma	0 (0.00)	0 (0.00)	1.000
Normal colon mucosa	1 (0.37)	2 (0.40)	0.557
Polyp location			<0.001
Caecum, n (%)	3 (1.12)	5 (1.00)	0.462
Ascending, n (%)	47 (17.47)	97 (19.48)	<0.001
Transverse, n (%)	45 (16.73)	110 (22.09)	<0.001
Descending, n (%)	26 (9.67)	75 (15.06)	<0.001
Sigmoid, n (%)	79 (29.37)	130 (26.10)	<0.001
Rectum, n (%)	69 (25.65)	81 (16.27)	0.254
Polyp shape			0.076
Pedunculated, n (%)	38 (14.13)	49 (9.84)	0.194
Sessile, n (%)	98 (36.43)	176 (35.34)	<0.001
Flat, n (%)	133 (49.44)	273 (54.82)	<0.001
Polyp size (mm), mean (SD)	5.03 (3.72)	4.51 (2.55)	0.050
Polyp size category			0.096
0–5 mm, n (%)	198 (73.61)	399 (80.12)	<0.001
6–10 mm, n (%)	61 (22.68)	83 (16.67)	0.047
>10 mm, n (%)	10 (3.72)	16 (3.21)	0.218
Adenoma location			0.397
Caecum, n (%)	1 (0.62)	3 (1.15)	0.330
Ascending, n (%)	39 (24.38)	47 (17.94)	0.325
Transverse, n (%)	36 (22.50)	72 (27.48)	<0.001
Descending, n (%)	20 (12.50)	44 (16.79)	0.003
Sigmoid, n (%)	41 (25.62)	64 (24.43)	0.018
Rectum, n (%)	23 (14.37)	32 (12.21)	0.192
Adenoma shape			0.277
Pedunculated, n (%)	33 (20.62)	39 (14.89)	0.413
Sessile, n (%)	64 (40.00)	105 (40.08)	0.001
Flat, n (%)	63 (39.38)	118 (45.04)	<0.001
Adenoma size (mm), mean (SD)	5.56 (3.54)	5.26 (3.08)	0.372
Adenoma size category			0.169
0–5 mm, n (%)	102 (63.75)	185 (70.61)	<0.001
6–10 mm, n (%)	50 (31.25)	61 (23.28)	0.238
>10 mm, n (%)	8 (5.00)	16 (6.11)	0.097

*P value from $\chi 2$ test (or Fisher's exact test, as appropriate).

adenoma and polyp detection rate, a logistic regression was performed to evaluate the effect of computer assisted diagnosis for colonoscopy on the adenoma/polyp detection rate. The response variable was the binary outcome of whether an adenoma/polyp was detected. The covariate was the group variable indicating whether the patient belonged to the computer assisted group. Regarding the number of detected adenomas and polyps, a Poisson regression was applied to evaluate the effect of computer assisted diagnosis for colonoscopy. A two-sided p value of 0.05 was used as the threshold for statistical significance. In the event of any baseline clinical and demographic characteristics showing a statistically significant difference between the two comparison groups, additional covariate adjusted logistics/Poisson regression models were built to address the possible confounding effect by adding those significant characteristics into the models as covariates. The primary outcome was ADR. The secondary outcomes were PDR, the mean number of polyps detected per colonoscopy, the mean number of adenomas detected per colonoscopy and the rate of false positives and false negatives.

RESULTS

Patient enrolment and baseline data

A total of 1130 consecutive patients were eligible for enrolment. Among these patients, 72 patients (31 in routine group, 41 in CADe group) were excluded during colonoscopy due to meeting exclusion criteria (figure 2). A total of 1058 eligible patients were analysed, with 536 patients randomised prospectively into the control group and 522 into the CADe group. Baseline characteristics are presented in table 1 (see online supplementary file 2). There were no statistically significant differences between the two groups in terms of demographic data and adenoma detection risk factors. There were no complications reported. The total withdrawal time in the control and CADe group were 6.39 min and 6.89 min (p<0.001), respectively. Two hundred and twenty-nine more biopsies were performed in the CADe group. Withdrawal times when biopsy time was excluded from analysis were 6.07 min and 6.18 min in the control and CADe groups, respectively (p=0.15).

Polyp characteristics, mean number of polyps per case and PDR

A total of 767 polyps were detected. There were 422 (55.02%) adenomas and 31 (4.04%) sessile serrated adenomas. Overall, 269 polyps (35.07%) were found in the control group and 498 (64.93%) in the CADe group (table 2). The mean number of polyps detected per colonoscopy in the control and the CADe group were 0.51 and 0.97, respectively (p<0.001). There was a 1.89-fold increase in the mean number of polyps detected between the two groups (95% CI 1.63 to 2.192, p<0.001) (table 3). The PDR of the control and CADe group were 0.29 and 0.45, respectively (OR=1.995, 95% CI 1.532 to 2.544, p<0.001) (table 3). There was no statistically significant difference between the two groups in terms of baseline clinical and demographic variables. Thus, covariate adjusted models were not considered to address the potential confounding effect.

Adenoma characteristics, mean number of adenomas per case and ADR

A total of 422 adenomas were detected (table 2). The mean number of adenomas detected per colonoscopy in the control and CADe group were 0.31 and 0.53, respectively (p<0.001). There was a 1.72-fold increase in the mean number of adenomas detected between the experimental and control groups (95% CI 1.419 to 2.084, p<0.001) (table 3). The ADR of the control and the CADe groups were 0.20 and 0.29, respectively (OR=1.61, 95% CI 1.213 to 2.135, p<0.001) (table 3).

The number of detected polyps was significantly higher in the CADe group as compared with the control group when considering non-pedunculated polyps, polyps ranging in size from 0 cm to 1 cm and polyps in all segments of the colon. The number of detected adenomas was also significantly higher in CADe group when considering non-pedunculated polyps, polyps smaller than

Table 3 Polyp and adenoma detection

	Routine colonoscopy (n=536)	CADe colonoscopy (n=522)	P value*	FC/OR	95% CI
PDR	0.291	0.4502	<0.001	1.995†	1.532 to 2.544
ADR	0.2034	0.2912	<0.001	1.61†	1.213 to 2.135
Mean number of detected polyp	0.5019	0.954	<0.001	1.89‡	1.63 to 2.192
Mean number of detected adenoma	0.3097	0.5326	<0.001	1.72‡	1.419 to 2.084

*P value from χ^2 test (or Fisher's exact test, as appropriate) or t-test.
†OR.
‡FC.
ADR, adenoma detection rate; FC, fold change; PDR, polyp detection rate.

0.5 cm and polyps in all segments of the colon with the exception of the caecum and the ascending colon (table 3).

Outcomes in excellent bowel preparation (BBPS ≥7)

In the situation of excellent bowel preparation, ADR in the CADe group showed a trend of 6% increase superior to that of the routine group. However, due to the inadequate sample size of the subgroup analysis, it failed to show a statistically significant difference. Other outcomes, including the mean number of detected adenomas, mean number of detected polyps and PDR were all significantly increased in the CADe group (table 4).

False positives with the automatic polyp detection system

There was a total of 39 false alarms (false positives) in the CADe group, averaging at 0.075 false alarms per colonoscopy (table 5).

Of all the detected polyps in CADe group, none was missed by the automatic polyp detection system (table 5).

DISCUSSION

CRC is a major public health issue given its high incidence and mortality rate. Furthermore, a recently published study reported a marked increase in the annual percentage change in the incidence of CRC among young adults.[19] The mortality rate and incidence of CRC in adults have drastically decreased (by 51% and 32%, respectively) in the last half century, mainly as a result of CRC screening and the removal of adenomatous polyps.[20] Screening colonoscopies have also allowed for an increase in 5-year survival rate in CRC, a consequence of early detection as well as removal of precancerous adenomas.[20] Given the impact of colonoscopy on CRC incidence and mortality, technological advances such as full spectrum colonoscopy, retrograde viewing accessories, balloon colonoscope and endocuff-assisted colonoscopy have been introduced to improve the ADR by expanding the visual field.[21–25] Methods aimed at improving the quality of colonoscopies such as minimal withdrawal time, split-dose bowel preparation and retroflexion in the right colon have also increased the ADR.[26–30] Despite these advancements, polyps can still be missed. Studies have also reported that some polyps are missed by the endoscopist despite being within the visual field.[31 32] Several hypotheses have been proposed to explain the mechanism by which polyps may be missed. These include differences in endoscopist skill level, differences in endoscopist tracking patterns, 'inattentional blindness', wherein an observer fails to process an image on the screen due to distraction, and 'change blindness', wherein changes are missed during interruptions in visual scanning or during eye movements13.[33–37] Distraction caused by fatigue or emotional factors may also contribute. A second party such as a nurse or a trainee observing may improve PDR. While several studies have shown that this increases PDR, controversy remains regarding ADR.[13–15] It is likely that adding additional human observers beyond one would not completely overcome these limitations.

With recent breakthroughs in artificial intelligence, especially with deep learning in computer vision, computer-aided diagnosis (CADx) for polyps during colonoscopy has drawn increased attention and has been shown to allow histological classification of colon polyps.[38 39] Though optical biopsy remains a promising field, tissue biopsy remains the gold standard, and the accuracy of AI diagnosis in optical biopsy depends on how much surface microstructures can reflect the histologic features of a lesion. In the case of a missed lesion, no further diagnosis can be made, accounting for the ongoing research in the field of computer-aided detection (CADe) in gastrointestinal (GI) endoscopy, which aims to increase adenoma detection automatically during white light endoscopy in real time.

High fidelity and consistent automatic colon polyp detection has been an attractive research topic for the past decade, with the aim of an increased ADR. However, to our knowledge, the current technology has yet to yield a sufficient diagnostic performance in order to be considered for clinical application.[40 41] In order for an automatic polyp detection system to be considered for real-world clinical application, it must have a very high sensitivity and specificity, a sufficient real-time standard processing time and an onscreen alerting system.[41 42] An inadequate

Table 4 Outcomes in excellent bowel preparation (BBPS ≥7)

	Routine colonoscopy (n=293)	CADe colonoscopy (n=292)	P value*	FC/OR	95% CI
PDR	0.2935	0.4144	0.002	1.703†	1.209 to 2.399
ADR	0.2082	0.2637	0.12	1.356†	0.924 to 1.991
Mean number of detected polyp	0.5017	0.8116	<0.001	1.618‡	1.317 to 1.987
Mean number of detected adenoma	0.2867	0.4212	0.007	1.464‡	1.11 to 1.932

*P value from χ^2 test (or Fisher's exact test, as appropriate) or t-test.
†OR.
‡FC.
ADR, adenoma detection rate; FC, fold change; PDR, polyp detection rate.

Table 5 Statistical results of false alarms and missed polyp

	CADe colonoscopy, n (%)
False alarm	39 (100.00)
Bubble	13 (33.33)
Faeces	5 (12.82)
Undigested debris	4 (10.26)
Wrinkled mucosa	7 (17.95)
Local inflammation	5 (12.82)
Local bleeding	1 (2.56)
Rounded drug capsules	4 (10.26)
Other (circular blood vessel, scar, diverticulum and so on)	0 (0.00)
Missed polyp	0 (0.00)

specificity would create a number of false positives. Conversely, an inadequate sensitivity would not increase the PDR. Moreover, for real-time detection to be efficient, the time of analysis must be fast, with no noticeable delay to the endoscopist. As a result of these prerequisites, most current studies on automatic polyp detection are small-scale, non-clinical investigations, though with rapidly increasing interest in the field and with the emergence of deep learning, dramatic advances are expected in the coming years.[41]

In this study, a significant increase in ADR, PDR and mean number of polyps and adenomas per colonoscopy were found in the CADe group as compared with the control group; however, the increase in overall adenoma detection was mainly due to an increase in diminutive adenomas. Most diminutive adenomas detected by the CADe system were, however, smaller, which supports the conventional view that small polyps are more likely to be missed within the visual field rather than bigger and more prominent polyps. Although diminutive adenomas confer less risk for malignancy compared with larger adenomas, the increase in overall adenoma detection rate may eventually contribute to a decreased risk of interval CRC. Further studies should address the role of CADe on decreasing interval cancer, which is the main goal of any screening colonoscopy.[9]

The results also showed a major increase in the detection of diminutive hyperplastic polyps, which may represent additional unnecessary polypectomies and add to workload. In the future, this CADe system may be combined with a CADx system to support a detect, diagnose and disregard[43] strategy to avoid excessive workload.

High reproducibility, fidelity and uniformity of such a CADe system are advantages compared with human assistance. However, direct comparison between the automatic polyp detection system and medical staff assistance of differing experience levels is also worthy of further investigation.

The reason why the system may have failed to assist the endoscopist in detecting more adenomas in the caecum and ascending colon could be due to the higher instability of the colonoscope in those areas, thus reducing the visual field. Moreover, there is also no significant difference in rectum possibly due to the good visualisation and stability of colonoscope in this segment.

This study is the first prospective, randomised controlled trial using a high-performance CADe system based on deep learning to assist endoscopists on detecting colon polyps with a large number of enrolled patients. The results indicate that previously unrecognised polyps may effectively be addressed by an artificial intelligence system. However, polyps that remain outside the visual field are still a major issue not addressed by the current CADe system. Both unrecognised and non-visualised polyps may be addressed together with a combination of different technologies in the future.

This study has several limitations. First, given the inability to blind the endoscopist, the exact contribution of the system may be difficult to assess. The very act of being observed may have also affected ADR in the experimental group because of 'competitive spirit'.[13] This mechanism may explain a potential confounder in the CADe group in that the endoscopist may have been more attentive in the setting of known observation. In this study, we subtracted the time of biopsy procedures from each corresponding withdrawal time as an indirect marker of attentiveness (table 1). The withdrawal time was similar (6.07 min vs 6.18 min, p=0.15) between the two groups, which may represent similar observation attentiveness. False alarms of the CADe system, 0.075 instances per colonoscopy, did not contribute to longer withdrawal time.

In the future, double-blind studies could be designed to investigate the exact contribution of this system in the increased adenoma detection rate. Such a study may also help determine whether a polyp is detected simultaneously by both the endoscopist and the system or initially missed by the endoscopist, a question that the current study was not designed to address.

The second limitation is the lack of external validity. The baseline adenoma and polyp detection rates of this study were not as high as reported from Western countries.[13 44–48] Multiple factors may contribute to this finding, including genetic, dietary, lifestyle and habitus differences between Chinese and Western populations, as well as differences in the morbidity of colon polyps/adenomas among the two groups. Therefore, the results of this study may not be generalisable to areas of the world where baseline ADR is higher. Further studies are needed to investigate the adaptability and effectiveness of this system in such areas.

Third, though false-positives rates were low, some false positives were unexpected by the designers of the system and occurred due to detection of medication capsules, of local sites of bleeding or of undigested debris, causing potential distraction during the procedure. This might be corrected by adding sufficient training data to the current system.

Fourth, this study did not control the fatigue level of participating endoscopists, which could be an independent factor that affects ADR. Future studies are needed to investigate the effectiveness of this CADe system on different fatigue levels.

Fifth, because of the inadequate sample size of colonoscopies performed by junior endoscopists, further studies are needed to show the role and effectiveness of this CADe system in different levels of training.

Lastly, the study was conducted using Olympus colonoscopy equipment. Thus, the adaptability of the system on equipment manufactured by other companies should also be explored.

In conclusion, this study shows that a real-time CADe system based on deep learning led to significant increases in both colorectal polyp and adenoma detection rates in a low prevalent ADR region. Given its high accuracy, fidelity and stability, the current CADe system is potentially applicable in current clinical practice for better detection of colon polyps.

Contributors PW and XL contributed to study concept and design. PW, XL, LL, PL, XX, YS, DZ, GX, MT and YL contributed to acquisition of data. PW, TMB, JRGB, SB and AB contributed to analysis, interpretation of data, drafting of the manuscript and statistical analysis. All authors read and approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Disclaimer The CADe system was developed by Shanghai Wision AI Co, Ltd., and was provided free-of-charge for the purposes of this study. Employees in the

company were not involved in the clinical trial in any way, including in study design, statistical analysis or manuscript writing

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the Institutional Review Board of Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital. It was registered with Chinese Clinical Trial Registry (ChiCTR) under identifier ChiCTR-DDD-17012221.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 American Cancer SocietyAmerican Cancer Society. *Cancer Facts and Figures: 2017*. Atlanta, Georgia, 2017.
2 Bibbins-Domingo K, Grossman DC, Curry SJ, *et al*. Screening for colorectal cancer: US preventive services task force recommendation statement. *JAMA* 2016;315:2564–75.
3 Rex DK, Boland CR, Dominitz JA, *et al*. Colorectal cancer screening: recommendations for physicians and patients from the u.s. multi-society task force on colorectal cancer. *Am J Gastroenterol* 2017;112:1016–30.
4 Brenner H, Chang-Claude J, Jansen L, *et al*. Reduced risk of colorectal cancer up to 10 years after screening, surveillance, or diagnostic colonoscopy. *Gastroenterology* 2014;146:709–17.
5 Doubeni CA, Corley DA, Quinn VP, *et al*. Effectiveness of screening colonoscopy in reducing the risk of death from right and left colon cancer: a large community-based study. *Gut* 2018;67:291–8.
6 Zauber AG, Winawer SJ, O'Brien MJ, *et al*. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med* 2012;366:687–96.
7 Winawer SJ, Zauber AG, Ho MN, *et al*. Prevention of colorectal cancer by colonoscopic polypectomy. The national polyp study workgroup. *N Engl J Med* 1993;329:1977–81.
8 Doubeni CA, Weinmann S, Adams K, *et al*. Screening colonoscopy and risk for incident late-stage colorectal cancer diagnosis in average-risk adults: a nested case-control study. *Ann Intern Med* 2013;158:312–20.
9 Corley DA, Jensen CD, Marks AR, *et al*. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014;370:1298–306.
10 Kaminski MF, Regula J, Kraszewska E, *et al*. Quality indicators for colonoscopy and the risk of interval cancer. *N Engl J Med* 2010;362:1795–803.
11 Mahmud N, Cohen J, Tsourides K, *et al*. Computer vision and augmented reality in gastrointestinal endoscopy. *Gastroenterol Rep* 2015;3:179–84.
12 Ahn SB, Han DS, Bae JH, *et al*. The miss rate for colorectal adenoma determined by quality-adjusted, back-to-back colonoscopies. *Gut Liver* 2012;6:64–70.
13 Aslanian HR, Shieh FK, Chan FW, *et al*. Nurse observation during colonoscopy increases polyp detection: a randomized prospective study. *Am J Gastroenterol* 2013;108:166–72.
14 Lee CK, Park DI, Lee SH, *et al*. Participation by experienced endoscopy nurses increases the detection rate of colon polyps during a screening colonoscopy: a multicenter, prospective, randomized study. *Gastrointest Endosc* 2011;74:1094–102.
15 Buchner AM, Shahid MW, Heckman MG, *et al*. Trainee participation is associated with increased small adenoma detection. *Gastrointest Endosc* 2011;73:1223–31.
16 Tajbakhsh N, Gurudu SR, Liang J. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Trans Med Imaging* 2016;35:630–44.
17 Misawa M, Kudo SE, Mori Y, *et al*. Artificial intelligence-assisted polyp detection for colonoscopy: initial experience. *Gastroenterology* 2018;154:2027–9.
18 Wang P, Xiao X, Glissen Brown JR, *et al*. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2018;2:741–8.
19 Bailey CE, Hu CY, You YN, *et al*. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. *JAMA Surg* 2015;150:17–22.
20 Burke C, Kaul V, Pohl H. Polyp resection and removal procedures: insights from the 2017 digestive disease week. *Gastroenterol Hepatol* 2017;13:1–24.
21 Rex DK, Chadalawada V, Helper DJ. Wide angle colonoscopy with a prototype instrument: impact on miss rates and efficiency as determined by back-to-back colonoscopies. *Am J Gastroenterol* 2003;98:2000–5.
22 Adler A, Aminalai A, Aschenbeck J, *et al*. Latest generation, wide-angle, high-definition colonoscopes increase adenoma detection rate. *Clin Gastroenterol Hepatol* 2012;10:155–9.
23 Halpern Z, Gross SA, Gralnek IM, *et al*. Comparison of adenoma detection and miss rates between a novel balloon colonoscope and standard colonoscopy: a randomized tandem study. *Endoscopy* 2015;47:238–44.
24 Leufkens AM, DeMarco DC, Rastogi A, *et al*. Third Eye Retroscope Randomized Clinical Evaluation [TERRACE] Study Group. Effect of a retrograde-viewing device on adenoma detection rate during colonoscopy: the TERRACE study. *Gastrointest Endosc* 2011;73:480–9.
25 Chin M, Karnes W, Jamal MM, *et al*. Use of the Endocuff during routine colonoscopy examination improves adenoma detection: A meta-analysis. *World J Gastroenterol* 2016;22:9642–9.
26 Cohen J, Grunwald D, Grossberg LB, *et al*. The effect of right colon retroflexion on adenoma detection: a systematic review and meta-analysis. *J Clin Gastroenterol* 2017;51:818–24.
27 Barclay RL, Vicari JJ, Doughty AS, *et al*. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *N Engl J Med* 2006;355:2533–41.
28 Sanchez W, Harewood GC, Petersen BT. Evaluation of polyp detection in relation to procedure time of screening or surveillance colonoscopy. *Am J Gastroenterol* 2004;99:1941–5.
29 Siddiki HA, Fath AR, Ravi S, *et al*. Comparison of ADR and SSP detection rate (SSP-DR) between small volume split dose preparation (SDP) and large volume SDP [DDW abstract 553]. *Gastrointest Endosc* 2017;85.
30 Horton N, Garber A, Hasson H, *et al*. Impact of single- vs. split-dose low-volume bowel preparations on bowel movement kinetics, patient inconvenience, and polyp detection: a prospective trial. *Am J Gastroenterol* 2016;111:1330–7.
31 Rex DK, Cutler CS, Lemmel GT, *et al*. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology* 1997;112:24–8.
32 Rex DK. Maximizing detection of adenomas and cancers during colonoscopy. *Am J Gastroenterol* 2006;101:2866–77.
33 Rogart JN, Siddiqui UD, Jamidar PA, *et al*. Fellow involvement may increase adenoma detection rates during colonoscopy. *Am J Gastroenterol* 2008;103:2841–6.
34 Wolfe JM, Reinecke A, Brawn P. Why don't we see changes?: The role of attentional bottlenecks and limited visual memory. *Vis cogn* 2006;14:749–80.
35 Memmert D, Unkelbach C, Ganns S. The impact of regulatory fit on performance in an inattentional blindness paradigm. *J Gen Psychol* 2010;137:129–39.
36 Simons DJ, Chabris CF. Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception* 1999;28:1059–74.
37 Simons DJ, Rensink RA. Change blindness: past, present, and future. *Trends Cogn Sci* 2005;9:16–20.
38 Chen PJ, Lin MC, Lai MJ, *et al*. Accurate classification of diminutive colorectal polyps using computer-aided analysis. *Gastroenterology* 2018;154:568–75.
39 Byrne MF, Chapados N, Soudan F, *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:gutjnl-2017-314547:94–100.
40 Bernal J, Tajkbaksh N, Sanchez FJ, *et al*. Comparative validation of polyp detection methods in video colonoscopy: results from the MICCAI 2015 endoscopic vision challenge. *IEEE Trans Med Imaging* 2017;36:1231–49.
41 Mori Y, Kudo SE, Berzin TM, *et al*. Computer-aided diagnosis for colonoscopy. *Endoscopy* 2017;49:813–9.
42 Alagappan M, Brown JRG, Mori Y, *et al*. Artificial intelligence in gastrointestinal endoscopy: the future is almost here. *World J Gastrointest Endosc* 2018;10:239–49.
43 Mori Y, Kudo SE, Misawa M, *et al*. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Ann Intern Med* 2018;169:357–66.
44 Jia H, Pan Y, Guo X, *et al*. Water Exchange Method Significantly improves adenoma detection rate: a multicenter, randomized controlled trial. *Am J Gastroenterol* 2017;112:568–76.
45 Bai Y, Fang J, Zhao SB, *et al*. Impact of preprocedure simethicone on adenoma detection rate during colonoscopy: a multicenter, endoscopist-blinded randomized controlled trial. *Endoscopy* 2018;50:128–36.
46 Xu Y, Chen K, Xu L, *et al*. Diagnostic yield is not influenced by the timing of screening endoscopy: morning versus afternoon. *Scand J Gastroenterol* 2018;53:365–9.
47 Cai B, Liu Z, Xu Y, *et al*. Adenoma detection rate in 41,010 patients from Southwest China. *Oncol Lett* 2015;9:2073–7.
48 Zhou H, Shen Z, Zhao J, *et al*. [Distribution characteristics and risk factors of colorectal adenomas]. *Zhonghua Wei Chang Wai Ke Za Zhi* 2018;21:678–84.