# Artificial intelligence-assisted optical diagnosis for the resect-and-discard strategy in clinical practice: the Artificial intelligence BLI Characterization (ABC) study

Authors

Emanuele Rondonotti[1,*], Cesare Hassan[2,*], Giacomo Tamanini[1], Giulio Antonelli[2,3], Gianluca Andrisani[4], Giovanni Leonetti[2,5], Silvia Paggi[1], Arnaldo Amato[1], Giulia Scardino[1], Dhanai Di Paolo[1,6], Giovanna Mandelli[1], Nicoletta Lenoci[1], Natalia Terreni[1], Alida Andrealli[1], Roberta Maselli[7,8], Marco Spadaccini[7,8], Piera Alessia Galtieri[8], Loredana Correale[2], Alessandro Repici[7,8], Francesco Maria Di Matteo[4], Luciana Ambrosiani[9], Emanuela Filippi[9], Prateek Sharma[10,11], Franco Radaelli[1]

Institutions

1 Gastroenterology Unit, Valduce Hospital, Como, Italy
2 Digestive Endoscopy Unit, Nuovo Regina Margherita Hospital, Rome, Italy
3 Department of Translational and Precision Medicine, Sapienza University of Rome, Rome, Italy
4 Digestive Endoscopy Unit, Campus Bio-Medico, University of Rome, Rome, Italy
5 Endoscopy Unit, Casa di Cura Nuova Santa Teresa, Viterbo, Italy
6 Department of Gastroenterology and Hepatology, Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
7 Department of Biomedical Sciences, Humanitas University, Rozzano, Milan, Italy
8 Endoscopy Unit, Humanitas Clinical and Research Center IRCCS, Rozzano, Milan, Italy
9 Pathology Department, Valduce Hospital, Como, Italy
10 Department of Gastroenterology, Kansas City VA Medical Center, Kansas City, Missouri, USA
11 Department of Gastroenterology and Hepatology, University of Kansas Medical Center, Kansas City, Kansas, USA

**submitted** 15.10.2021
**accepted after revision** 13.5.2022
**published online** 13.5.2022

Bibliography
Endoscopy 2023; 55: 14–22
**DOI** 10.1055/a-1852-0330
**ISSN** 0013-726X

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Supplementary material
Supplementary material is available under https://doi.org/10.1055/a-1852-0330

Scan this QR-Code for the author commentary.

Corresponding author
Emanuele Rondonotti, MD, PhD, Gastroenterology Unit, Valduce Hospital, Via Dante 10, 22100, Como, Italy
ema.rondo@gmail.com

## ABSTRACT

**Background** Optical diagnosis of colonic polyps is poorly reproducible outside of high volume referral centers. The present study aimed to assess whether real-time artificial intelligence (AI)-assisted optical diagnosis is accurate enough to implement the leave-in-situ strategy for diminutive (≤ 5 mm) rectosigmoid polyps (DRSPs).

**Methods** Consecutive colonoscopy outpatients with ≥ 1 DRSP were included. DRSPs were categorized as adenomas or nonadenomas by the endoscopists, who had differing expertise in optical diagnosis, with the assistance of a real-time AI system (CAD-EYE). The primary end point was ≥ 90 % negative predictive value (NPV) for adenomatous histology in high confidence AI-assisted optical diagnosis of DRSPs (Preservation and Incorporation of Valuable endoscopic Innovations [PIVI-1] threshold), with histopathology as the reference standard. The agreement between optical- and histology-based post-polypectomy surveillance intervals (≥ 90 %; PIVI-2 threshold) was also calculated according to European Society of Gastrointestinal Endoscopy (ESGE) and United States Multi-Society Task Force (USMSTF) guidelines.

**Results** Overall 596 DRSPs were retrieved for histology in 389 patients; an AI-assisted high confidence optical diagnosis was made in 92.3 %. The NPV of AI-assisted optical diagnosis for DRSPs (PIVI-1) was 91.0 % (95 %CI 87.1 %–93.9 %). The PIVI-2 threshold was met with 97.4 % (95 %CI

* Joint first authors

95.7%–98.9%) and 92.6% (95%CI 90.0%–95.2%) of patients according to ESGE and USMSTF, respectively. AI-assisted optical diagnosis accuracy was significantly lower for nonexperts (82.3%, 95%CI 76.4%–87.3%) than for experts (91.9%, 95%CI 88.5%–94.5%); however, nonexperts quickly approached the performance levels of experts over time.

**Conclusion** AI-assisted optical diagnosis matches the required PIVI thresholds. This does not however offset the need for endoscopists' high level confidence and expertise. The AI system seems to be useful, especially for nonexperts.

## Introduction

The implementation of clinical strategies based on optical diagnosis of diminutive (≤5 mm) colorectal polyps may lead to a substantial saving of economic and financial resources [1,2]. Despite this, 84% of European endoscopists reported not using such strategies – also known as the "leave-in-situ" and "resect-and-discard" strategies [3] – in their practice owing to the fear of an incorrect optical diagnosis [4]. Indeed, the accuracy of optical diagnosis is operator-dependent, and values reported in community settings are below the safety thresholds proposed for its incorporation in clinical practice [5,6].

Because of its accuracy in discriminating different visual patterns, artificial intelligence (AI) has the potential to help endoscopists in distinguishing neoplastic from non-neoplastic polyps, making the characterization process more reliable and objective. Preliminary retrospective data in an artificial setting have shown high feasibility and accuracy levels of AI for optical diagnosis of colorectal polyps [7,8], especially when focusing on diminutive rectosigmoid polyps (DRSPs) [9–12]. Recently, AI-assisted endocytoscopy showed an adequate accuracy for ≤5-mm rectosigmoid polyps, achieving a ≥90% negative predictive value (NPV) for adenomatous histology in a real-time setting [13]. Despite this, endocytoscopy is not widely used in Western countries, and it requires dedicated equipment and special skills that are not available in community endoscopy.

In the present study, we exploited a new computer-aided diagnosis system (CAD-EYE; Fujifilm Co., Tokyo, Japan) that provides real-time polyp characterization with standard endoscopy. The study was primarily aimed at prospectively evaluating whether the endoscopist assisted by AI could achieve a ≥90% NPV for adenomatous histology of DRSPs. Secondary study aims were to evaluate the optical diagnosis performances of the endoscopists alone and the AI system alone, and the agreement between the post-polypectomy surveillance intervals based on optical diagnosis and histology.

## Methods

### Centers and patients

This prospective cohort study was conducted in four open-access endoscopy centers in Italy (listed in **Appendix 1s**, see online-only Supplementary Material). The institutional review boards of all participating centers approved the protocol. All patients provided their written informed consent. The study is reported according to STROBE guidelines [14].

Consecutive adults (18–85 years) undergoing outpatient colonoscopy were considered for inclusion, with enrollment limited to those patients in whom at least one DRSP was detected. The exclusion criteria are listed in the **Appendix 2s**.

### Study outcomes

According to the Preservation and Incorporation of Valuable endoscopic Innovations (PIVI)-1 threshold, proposed by the American Society of Gastrointestinal Endoscopy [15], the primary end point was to assess whether AI-assisted optical diagnosis with a high degree of confidence achieved ≥90% NPV for adenomatous histology of DRSPs, having histopathology as the reference standard. The secondary aims were: (i) to calculate the performance measures of the endoscopist alone (endoscopist-alone optical diagnosis) and the AI system alone (AI-alone optical diagnosis); (ii) to evaluate whether the post-polypectomy surveillance interval based on optical diagnosis achieved ≥90% agreement (the PIVI-2 threshold) according to both the United States Multi-Society Task Force (USMSTF) and European Society of Gastrointestinal Endoscopy (ESGE) guidelines [16,17].

We also planned exploratory subgroup analyses on accuracy according to the level of expertise (i.e. expert vs. nonexpert), level of confidence (i.e. high vs. low), and polyp location (i.e. rectosigmoid vs. nonrectosigmoid polyps).

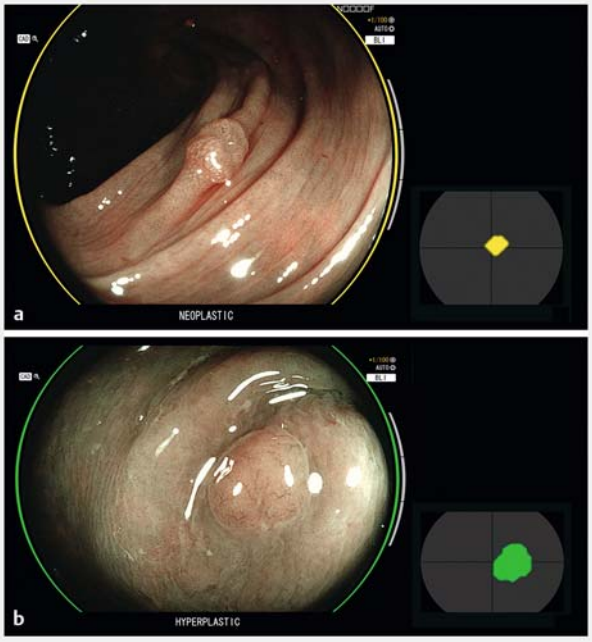
### Endoscopic procedures

All procedures were performed using the ELUXEO 7000 endoscopy platform (EC-760ZPV and EC-760RV endoscopes, ELUXEO VP-7000 videoprocessor, and ELUXEO BL-7000 light source; Fujifilm Co.).

The participating endoscopists were dichotomized as experts (had followed a dedicated training program, underwent periodic auditing and monitoring, and performed optical diagnosis on a regular basis, according to the ESGE curriculum [18,19]) and nonexperts in optical diagnosis. Regardless of their expertise, all endoscopists received formal training, which consisted of a 45-minute lecture on the principles of optical diagnosis, the blue-light imaging (BLI) system, the BLI Adenomas Serrated International Classification (BASIC) system [20], and the features of the AI system used in the present study.

### AI system

A real-time convolutional neural network-based AI system (CAD-EYE) was used for polyp characterization in BLI mode. The technical features of the system have been described else-




▶ **Fig. 1** Illustration of the output of the artificial intelligence system (CAD-EYE) showing color coding of the location map and similar coloring of side brackets around the endoscopic image with: **a** yellow dots and brackets for adenomatous polyps (plus "neoplastic" banner below the endoscopic image); and **b** green dots and brackets for nonadenomatous polyps (plus "hyperplastic" banner below the endoscopic image).

where [8]. Briefly, the AI system provides optical diagnosis through: (i) polyp identification in the "location map"; (ii) colored brackets surrounding the endoscopic image; and (iii) diagnostic labeling as "hyperplastic" or "neoplastic" (i. e. adenoma or nonadenoma, according to the manufacturer's indication). ▶ **Fig. 1** shows how the CAD-EYE output is provided.

## Steps involved in AI-assisted optical diagnosis

All polyps identified by the endoscopist had their size, location, and morphology (according to the Paris classification [21]) reported. They were resected and retrieved in separate jars and sent for pathology assessment.

All ≤5-mm polyps were characterized (as adenomas or nonadenomas) through a three-step sequential process. Every step of polyp optical diagnosis was performed with BLI and in real time during the endoscopic procedure.

In the first step (endoscopist-alone optical diagnosis), the endoscopist categorized the polyp as an adenoma or nonadenoma, using the BASIC classification, without AI assistance. The endoscopist's confidence level in the optical diagnosis (high vs. low) was recorded. Only DRSPs evaluated with high confidence were included in the analysis of endoscopist-alone performance.

In the second step (AI-alone optical diagnosis), the AI system was switched on and the output that was automatically provided by the AI system (adenoma vs. nonadenoma) was recorded, irrespective of the previous output and level of confidence of the endoscopist. The AI output was collected only when the system was able to provide it and it was technically reliable and stable over time. Further details about the CAD-EYE user interface and operation are reported in **Appendix 3 s**.

In the third step (AI-assisted optical diagnosis), the final diagnosis (adenoma vs. nonadenoma) provided by the endoscopist combining the results of the first two steps was reported, as well as the confidence level (high vs. low). Only DRSPs receiving a high confidence AI-assisted optical diagnosis were used for the computation, irrespective of the results of the previous steps.

## Pathology (reference standard)

Expert pathologists (at least one in each center), blinded to the optical diagnosis, evaluated all the resected polyps according to the Vienna classification [22]. For the present study, hyperplastic polyps, sessile serrated lesions, inflammatory polyps, or normal mucosal samples were all labeled as nonadenomas. Taking into account the dichotomy "adenoma" and "nonadenoma," if disagreement between the pathological diagnosis and high confidence AI-assisted optical diagnosis of DRSPs was disclosed, the pathology specimens were blindly reviewed by a second expert pathologist and the polyp was then reclassified by agreement. Adenomas with significant villous features (>25 %), size ≥1.0 cm, high grade dysplasia, or early invasive cancer were defined as advanced.

## Statistical analysis and sample size calculation

With an expected prevalence of rectosigmoid adenomas of 46.8 %, based on previous data collected in the centers participating in the present study, an NPV for AI-assisted diagnosis of >90 % implied a <0.11 likelihood ratio for negative results. Using the equations described by Simel et al. [23], we determined the minimum sample size required to test the primary hypothesis (at 5 % two-sided significance level and 80 % power) to be 235 adenomatous DRSPs.

With regard to the post-polypectomy surveillance intervals, the optical diagnosis-based strategy was calculated taking into account high confidence optical diagnosis of ≤5-mm polyps, along with the histopathological assessment of both polyps ≥6-mm in size and those of ≤5 mm that were evaluated with low confidence [6]. If only diminutive polyps were detected and evaluated with high confidence, the optical diagnosis-based post-polypectomy surveillance interval was provided at the end of the endoscopic procedure; otherwise, it was made as soon as the histopathology became available.

Comparisons of categorical variables were performed by two-sided chi-squared test or Fisher exact test, as appropriate. A *P* value of ≤0.05 was considered statistically significant for our primary outcome measure. All other outcome measures (e. g. diagnostic performance according to the level of endoscopist expertise) were treated as secondary in our study design. There was no need to adjust for multiplicity, as findings for secondary outcomes were considered subsidiary and exploratory, rather than confirmatory [24].

## Results

### Patients and polyps

From October 2020 to February 2021, 1134 adults referred for outpatient screening, surveillance, or diagnostic colonoscopy were evaluated. Of these, 745 were excluded from further analysis (study flow chart is shown in **Fig. 1 s**); therefore 389 patients (52.5 % men; mean [SD] age: 63.7 [10.4] years;) with 1031 polyps were identified (▶ **Table 1**). Of the polyps, 30 were not retrieved (30/1031: 2.9 %) and DRSPs accounted for 58.6 % (604/1031) of all polyps.

Because of disagreement between the high confidence AI-assisted optical diagnosis and the histology, 64 DRSPs (64/550; 11.6 %, 95 %CI 9.1 %–14.6 %) were blindly reviewed by a second expert pathologist and two of these (2/64; 3.1 %, 95 % CI 0.4 %–10.8 %), were reclassified. In one case the final diagnosis was in line with the AI-assisted high confidence diagnosis.

Out of 596 retrieved DRSPs, 259 were histologically classified as adenomas and 337 as nonadenomas. The features of the DRSPs are reported in **Table 1 s**.

### Accuracy of the optical diagnosis process

Optical diagnosis was performed with high confidence by the endoscopist alone (step one) in 540/596 DRSPs (90.6 %, 95 %CI 87.8 %–92.8 %). The AI-alone optical diagnosis (step two) was recorded in 541/596 DRSPs (90.8 %, 95 %CI 88.2 %–92.9 %): out of the 55 polyps excluded from the AI-alone optical diagnosis, the characterization was unstable in 47 (47/596; 7.9 %, 95 %CI 5.8 %–10.3 %) and characterization was not possible (the system did not provide any recordable outcome) in the remaining eight polyps (8/596; 1.3 %, 95 %CI 0.6–2.6 %). AI-assisted endoscopist diagnosis (step three) was performed with high confidence in 550/596 DRSPs (92.3 %, 95 %CI 89.8 %–94.3 %); 238 of these were adenomas (43.3 %, 95 %CI 39.1 %–47.5 %) and 312 had nonadenomatous histology (56.7 %, 95 %CI 52.5 %–60.9 %).

The NPV for adenomatous histology of DRSPs for AI-assisted optical diagnosis (step three) was 91.0 % (95 %CI 87.1 %–93.9 %), while the sensitivity, specificity, and accuracy were 88.6 % (95 % CI 83.7 %–91.4 %), 88.1 % (95 %CI 83.9 %–91.4 %), and 88.4 % (95 %CI 85.3 %–90.9 %), respectively (▶ **Table 2**). The same figures for endoscopist-alone optical diagnosis (step one) and AI-alone optical diagnosis (step two) are also detailed in ▶ **Table 2**. Out of 541 DRSPs in which AI output was predictable and stable, the outcomes provided by the endoscopist alone (step one) and by AI alone (step two) were divergent in 32 cases (5.9 %, 95 %CI 4.1 %–8.2 %).

When comparing AI-assisted diagnosis with high confidence with that scored as low confidence by the AI-assisted endoscopist (adenomatous histology, 21/46; 45.6 %, 95 %CI 30.9 %–61.0 %), the accuracy and NPV were significantly lower in the low confidence group: accuracy 50.0 % (95 %CI 45.3 %–74.9 %) vs. 88.4 % (95 %CI 85.3 %–90.9 %); NPV 70.6 % (95 %CI 44.0 %–87.1 %) vs. 91.0 % (95 %CI 87.1 %–93.9 %) (**Fig. 2 s**).

When focusing our analysis of ≤ 5-mm lesions according to their location, the overall accuracy for high confidence diagnosis was similar for DRSPs and diminutive polyps proximal to the rectosigmoid tract (diminutive non-rectosigmoid polyps): 87.0 % (95 %CI 86.0 %–92.9 %) vs. 88.4 % (95 %CI 85.3 %–90.9 %). However, the NPV for adenomatous histology was significantly lower for diminutive non-rectosigmoid polyps compared with that for DRSPs: 72.4 % (95 %CI 58.8 %–82.9 %) vs. 91.0 %, (95 %CI 87.1 %–93.9 %) (**Fig. 2 s**). This was related to a higher relative prevalence of adenomatous histology (222/285; 77.9 %, 95 %CI 72.6 %–82.5 %) in diminutive non-rectosigmoid polyps compared with that among DRSPs (238/550; 43.3 %, 95 %CI 39.1 %–47.5 %; $P < 0.001$). In addition, the specificity of optical diagnosis was lower for diminutive non-rectosigmoid polyps than for DRSPs: 66.7 % (95 %CI 53.5 %–69.6 %) vs. 88.1 % (95 %CI 83.9 %–91.4 %).

### Agreement of post-polypectomy surveillance intervals

The optical diagnosis-based post-polypectomy surveillance interval was readily available at the end of colonoscopy (i. e. patients in whom only diminutive polyps were detected, with all evaluated with high confidence) in 280 patients (280/389; 72.0 %, 95 %CI 67.2 %–76.4 %).

The post-polypectomy surveillance interval based on AI-assisted optical diagnosis (step three) was correctly advised in 97.4 % of patients (95 %CI 95.7 %–98.9 %) within the ESGE framework and in 92.6 % (95 %CI 90.0 %–95.2 %) within the USMSTF framework (▶ **Table 3**).

The post-polypectomy surveillance interval agreement rate with both the ESGE and USMSTF recommendations, based on endoscopist-only optical diagnosis (step one) and on AI-only optical diagnosis (step two) are also reported in ▶ **Table 3**. The rates of patients receiving a delayed optical diagnosis-based surveillance colonoscopy according to the ESGE and USMSTF guidelines and to each step of optical diagnosis are shown in **Table 2 s**.

### Performance in AI-assisted optical diagnosis according to level of expertise

Of the 18 participating endoscopists, nine were experts and nine were nonexperts. Of the 596 DRSPs included in the analysis, 374 were evaluated by experts (62.7 %, 95 %CI 76.1 %–82.7 %) and 222 by nonexperts (37.3 %, 95 %CI 33.3 %–41.7 %). Among DRSPs, the rates of adenomas evaluated by experts and nonexperts were 165/374 (44.1 %, 95 %CI 39.0 %–49.3 %) and 94/222 (42.3 %, 95 %CI 35.7 %–49.1 %), respectively. AI-assisted optical diagnosis was performed with high confidence in 92.5 % (346/374; 95 %CI 89.4 %–94.9 %) and in 91.9 % (204/222; 95 %CI 87.5 %–95.2 %) of DRSPs by experts and nonexperts, respectively. The accuracy of AI-assisted optical diagnosis of DRSPs was higher for experts than nonexperts in optical diagnosis: 91.9 % (95 %CI 88.5 %–94.5 %) vs. 82.3 % (95 %CI 76.4 %–87.3 %).

To explore the potential impact of the use of AI according to the level of endoscopist expertise, the sensitivity, specificity, positive predictive value, NPV, and accuracy in endoscopist-alone (step one) and AI-assisted optical diagnosis (step three) for experts and nonexperts are reported in ▶ **Table 4**.

The rate of cases in which the outcomes provided by the endoscopist alone (step one) and by AI alone (step two) were divergent was similar for experts and nonexperts (22/344; 6.4 %,

▶ **Table 1** Baseline features of the 389 patients and their polyps according to endoscopist expertise.

| | Endoscopist experience | | All |
|---|---|---|---|
| | Expert | Nonexpert | |
| **Patient features** | | | |
| Number of patients | 235 | 154 | 389 |
| Age, mean (SD), years | 63.2 (10.1) | 64.9 (10.6) | 63.7 (10.4) |
| Sex, male, % | 53.6 | 50.6 | 52.5 |
| Indication for colonoscopy, n (%) | | | |
| ▪ Symptoms | 63 (26.8) | 34 (22.2) | 97 (25.0) |
| ▪ Surveillance | 88 (37.4) | 52 (34.0) | 140 (36.1) |
| ▪ FIT-positive screening | 39 (16.6) | 30 (19.6) | 69 (17.8) |
| ▪ Primary screening | 45 (19.2) | 38 (24.2) | 83 (21.1) |
| **Polyp features** | | | |
| Size, n (%) | | | |
| ▪ 1–5 mm | 562 (92.0) | 381 (90.7) | 943 (91.5) |
| ▪ 6–9 mm | 31 (5.1) | 21 (5.0) | 52 (5.0) |
| ▪ ≥ 10 mm | 18 (2.9) | 18 (4.3) | 36 (3.5) |
| Shape, n (%) | | | |
| ▪ Ip (pedunculated) | 16 (3.8) | 17 (2.8) | 33 (3.2) |
| ▪ Is (sessile) | 290 (69.1) | 476 (77.9) | 766 (74.3) |
| ▪ Isp (semipedunculated) | 6 (1.4) | 18 (2.9) | 24 (2.3) |
| ▪ IIa (flat-raised) | 97 (23.1) | 31 (5.1) | 128 (12.4) |
| ▪ IIb (flat) | 11 (2.6) | 69 (11.3) | 80 (7.7) |
| ▪ IIc (depressed) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| Location, n (%) | | | |
| ▪ Cecum | 33 (5.8) | 32 (6.9) | 65 (6.3) |
| ▪ Ascending colon | 95 (16.8) | 95 (20.5) | 190 (18.4) |
| ▪ Transverse colon | 55 (9.7) | 36 (7.6) | 91 (8.8) |
| ▪ Descending colon | 23 (4.0) | 8 (1.6) | 31 (3.1) |
| ▪ Sigmoid colon | 227 (40.1) | 211 (45.7) | 438 (42.5) |
| ▪ Rectum | 133 (23.6) | 83 (17.7) | 216 (20.9) |
| Histology, n (%) | | | |
| ▪ Low risk adenoma | 287 (47.0) | 216 (51.5) | 503 (48.9) |
| ▪ High risk adenoma | 45 (7.4) | 12 (2.8) | 57 (5.5) |
| ▪ Hyperplastic polyp | 214 (35.0) | 143 (34.0) | 357 (34.6) |
| ▪ Sessile serrated lesion | 22 (3.6) | 13 (3.2) | 35 (3.4) |
| ▪ Inflammatory/normal mucosa | 26 (4.2) | 23 (5.5) | 49 (4.7) |
| ▪ Not retrieved | 17 (2.8) | 13 (3.0) | 30 (2.9) |

FIT, fecal immunochemical test.

▶ **Table 2** Accuracy parameters (95 %CIs) for optical diagnosis of diminutive rectosigmoid polyps at each step of the optical diagnosis process.

| | Optical diagnosis process* | | |
|---|---|---|---|
| | **Endoscopist alone (step one)** | **AI alone (step two)** | **AI assisted (step three)** |
| Sensitivity | 88.6 % (83.6 %–92.2 %) | 81.9 % (76.2 %–86.5 %) | 88.6 % (83.7 %–91.4 %) |
| Specificity | 88.8 % (84.5 %–91.9 %) | 88.7 % (84.4 %–91.9 %) | 88.1 % (83.9 %–91.4 %) |
| Positive predictive value | 86.1 % (80.8 %–90.0 %) | 84.4 % (78.8 %–88.7 %) | 85.1 % (79.8 %–89.1 %) |
| Negative predictive value | 90.9 % (86.8 %–93.7 %) | 86.7 % (82.3 %–90.1 %) | 91.0 % (87.1 %–93.9 %) |
| Accuracy | 88.7 % (85.7 %–91.2 %) | 85.8 % (82.5 %–88.6 %) | 88.4 % (85.3 %–90.9 %) |

AI, artificial intelligence.
* No statistically significant differences were found on comparison of each accuracy parameter for the three steps of optical diagnosis ($P > 0.05$ for each comparison).

▶ **Table 3** Agreement (95 %CIs) between the histology-based and optical diagnosis-based post-polypectomy surveillance intervals for each step of the optical diagnosis process according to the different recommendations*.

| | | International scientific society recommendations | |
|---|---|---|---|
| | | **ESGE** | **USMSTF** |
| Optical diagnosis process | Endoscopist alone (step one) | 97.1 % (95.4 %–98.8 %) | 92.6 % (90.0 %–95.2 %) |
| | AI alone (step two) | 96.8 % (95.0 %–98.6 %) | 92.1 % (89.4 %–94.8 %) |
| | AI assisted (step three) | 97.4 % (95.7 %–98.9 %) | 92.6 % (90.0 %–95.2 %) |

ESGE, European Society of Gastrointestinal Endoscopy; USMSTF, United States Multi-Society Task Force; AI, artificial intelligence.
* No statistically significant differences were found on comparison of the post-polypectomy surveillance interval agreement calculated for the three steps of optical diagnosis, according to both the ESGE and USMSTF recommendations ($P > 0.05$ for each comparison).

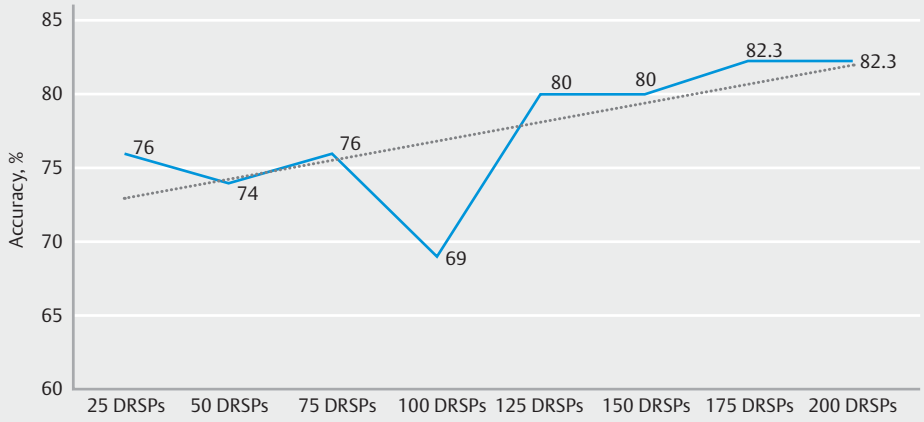
▶ **Table 4** Accuracy parameters (95 %CIs) for optical diagnosis of diminutive rectosigmoid polyps in step one (endoscopist-alone optical diagnosis) and step three (AI-assisted optical diagnosis) of the optical diagnosis process, according to endoscopist expertise in optical diagnosis.

| | Expert endoscopists[1] | | Nonexpert endoscopists[2] | |
|---|---|---|---|---|
| | **Endoscopist-alone optical diagnosis (step one)** | **AI-assisted optical diagnosis (step three)** | **Endoscopist-alone optical diagnosis (step one)** | **AI-assisted optical diagnosis (step three)** |
| Sensitivity | 90.6 (84.4–94.5) | 90.1 (83.8–94.1) | 81.8 (71.8–88.9) | 86.2 (76.7–92.3) |
| Specificity | 92.1 (87.0–95.3) | 93.3 (88.6–96.2) | 83.3 (74.9–89.4) | 79.5 (70.8–86.1) |
| PPV | 90.0 (83.7–94.1) | 91.3 (85.2–95.0) | 79.1 (69.0–86.6) | 75.7 (65.9–83.5) |
| NPV | 92.5 (87.5–95.7) | 92.4 (87.5–95.5) | 85.6 (77.3–91.2) | 88.6 (80.5–93.6) |
| Accuracy | 91.4 (87.9–94.1) | 91.9 (88.5–94.5) | 82.7 (76.7–87.6) | 82.3 (76.4–87.3) |

AI, artificial intelligence; PPV, positive predictive value; NPV, negative predictive value.
[1] No significant differences were found on comparison of each accuracy parameter for optical diagnosis step one vs. step three among the expert endoscopists ($P > 0.05$ for all comparisons).
[2] Trends toward increased sensitivity and negative predictive value were observed when comparing step one to step three among nonexpert endoscopists ($P = 0.053$ and $P = 0.058$, respectively).

95 %CI 4.0 %–9.5 % vs. 10/197; 5.1 %, 95 %CI 2.4 %–9.1 %). In detail, in these 32 divergent cases, experts and nonexperts "disregarded" the AI diagnosis and kept their initial diagnosis as well as the level of confidence in 17 cases (17/22; 77.3 %, 95 %CI 54.6 %–92.2 %) and seven cases (7/10; 70.0 %, 95 %CI 34.7 %–93–3 %), respectively. The experts modified either the diagnosis or the level of confidence in three and two DRSPs, respectively. The same figures for nonexperts were one and two DRSPs, respectively. Among divergent cases, by comparing endoscopist-alone optical diagnosis (step one) with AI-assisted optical diagnosis (step three), the accuracy increased from 87.5 % (95 %CI 61.6 %–98.4 %) to 92.8 % (95 %CI 66.1 %–99.8 %) and from 50.0 % (95 %CI 6.7–93.2 %) to 66.6 % (95 %CI 9.4–99.1 %) for experts and nonexperts, respectively.

To evaluate the potential learning curve, we compared the performance of AI-assisted optical diagnosis calculated from




▶ **Fig. 2** The trend (dotted line) in accuracy through the series of consecutive diminutive rectosigmoid polyps evaluated by nonexperts. DRSP, diminutive rectosigmoid polyp.

the first 50 DRSPs to be evaluated with that calculated from last 50 DRSPs, according to level of expertise. As far as accuracy was concerned, no differences were observed for experts (92.0 %, 95 %CI 80.7 %–97.7 % vs. 90.0 %, 95 %CI 66.2 %–89.9 %), whereas a trend toward an increase was observed for nonexperts (74.0 %, 95 %CI 59.6 %–85.3 % vs. 88.0 %, 95 %CI 75.6 %–95.4 %). Interestingly, the AI-assisted NPV of the last 50 DRSPs evaluated by nonexperts met the PIVI-1 threshold (95.2 %, 95 %CI 76.2 %–99.85 %) and was similar to the NPV calculated for the last 50 DRSPs evaluated by experts (93.9 %, 95 %CI 79.7 %–99.2 %; (**Fig. 3 s**). The trend of accuracy through the series of consecutive DRSPs evaluated by nonexperts is shown in ▶ **Fig. 2**.

**Table 3 s** summarizes the NPVs of endoscopist-alone optical diagnosis and AI-assisted optical diagnosis of DRSPs for each participating endoscopist.

With regards to the post-polypectomy surveillance interval, **Table 4 s** summarizes the agreement between the optical diagnosis-based and histology-based post-polypectomy surveillance intervals, according to endoscopist expertise and according to both the ESGE and USMSTF guidelines. Interestingly, the 90 % threshold (PIVI-2) was met regardless of the level of expertise or the use of AI.

## Discussion

In a clinical setting, real-time AI-assisted optical diagnosis without magnification was feasible in more than 90 % of cases and was effective in reaching the thresholds required for the implementation of cost-saving strategies, namely a ≥ 90 % NPV for ≤ 5-mm rectosigmoid polyps for the leave-in-situ strategy, and ≥ 90 % agreement rate on post-polypectomy intervals for the resect-and-discard strategy. Level of confidence and expertise were inversely associated with the overall accuracy of AI-assisted optical diagnosis, which was, in contrast, unaffected by the location of the lesion.

The clinical relevance of our study is the analysis of the interaction between AI and human factors in the decision-making process as expected in a clinical setting. Differently from previous studies [9, 25], we did not only analyze the value of AI as an independent reader, but sequentially incorporated the output of AI prediction into the diagnostic process of the endoscopist. This likely represents the most realistic scenario, as the endoscopist will ultimately be responsible for the incorporation of AI information into clinical practice. Of note, this deals not only with AI-assisted accuracy, but also with the level of confidence.

Regarding accuracy parameters, diagnoses made both by the endoscopist alone and with AI assistance were able to achieve the PIVI criteria. This was not unexpected as approximately two-thirds of optical diagnoses were performed by expert endoscopists, who have been shown to match such criteria, irrespective of their use of AI [3, 26, 27]. However, the equivalent accuracy between pre- and post-AI endoscopist diagnoses should not be underestimated.

First, before our study, a possible detrimental effect of AI on endoscopist diagnosis could not be ruled out. AI-assisted optical diagnosis might become the new standard of care; hence our study is reassuring with regard to the lack of harm from such an AI-based strategy. Second, our study was underpowered to show a possible benefit of AI for expert endoscopists as a sample size of a different magnitude would have been required to show an absolute difference within a 5 % range. Third, when segregating our study population according to the level of experience, our exploratory analysis showed favorable trends in terms of NPV and the learning curve for nonexpert endoscopists. This possible role of AI on optical diagnosis training of nonexperts deserves to be urgently addressed in future studies.

Fourth, there are several barriers against the implementation of optical diagnosis-based cost-saving strategies that may be addressed by AI, irrespective of its effect on endoscopist accuracy. For instance, other stakeholders, such as health systems and patients, could be reassured in replacing histology with an AI-documented optical diagnosis. Such documentation may be easily recorded and traced, as much as histological diagnosis.

Similarly, the endoscopist may be more confident in incorporating the possible risk of litigation related with an optical diagnosis, when this is independently supported by an AI machine.

Furthermore, the step-by-step design of our study allowed assessment of the AI output (AI-alone optical diagnosis) as an independent reader and compared it with the previous artificial validation done on a sequence of still-frame high quality selected images [8], which provided reassurance on the performance of AI in a real practice setting [8, 25]. Of note, we chose this study design because it closely mimics the endoscopist's real-life decision-making process (polyp detection under white-light endoscopy and initial evaluation by the endoscopist, with subsequent activation of virtual chromoendoscopy systems and artificial intelligence).

Interestingly, the absence of substantial differences in the rate of polyps assessed in a stable and reliable way by experts or nonexperts suggests that this system is easy to use, regardless of the endoscopists' expertise. However, we acknowledge that the performance of AI alone, although close to the settled thresholds, was somewhat suboptimal. Nevertheless, we used the first available CAD-EYE release and further improvements are expected with the second-generation software, which is going to be launched in Europe.

Unlike in a previous endocytoscopy-based study [9], where a much lower accuracy was shown for proximal as compared with distal lesions, we observed a similar accuracy between distal and proximal predictions. However, the apparent drop in NPV when passing from distal to proximal lesions was exclusively related to the higher relative prevalence of adenomatous histology in the latter location, confirming the inapplicability of the leave-in-situ strategy for proximal diminutive lesions.

The main limitation of our study was its inherent psychological bias, as endoscopists were aware that all the polyps would be sent to histology, irrespective of their prediction, so this may have led to a higher than expected rate of high confidence diagnoses.

Second, the feasibility of the resect-and-discard strategy must be confirmed in more dedicated studies because we excluded patients without distal lesions as the presence of a DRSP was an inclusion criterion. Furthermore, in the present study we labelled sessile serrated lesions (SSLs) as nonadenomatous. Although the natural history of these lesions remains poorly defined, some authors suggest they have a behavior and risk profile closer to adenomas than to hyperplastic polyps. However, in our study the rate of SSLs among DRSPs was 2.6 % and none of them showed dysplasia. This marginalizes the clinical relevance of the matter. In fact, even when SSLs were reclassified as adenomatous polyps, the results were very similar. However, we acknowledge that the detailed histopathological assessment could be an issue when focusing on right-sided polyps of > 5 mm.

Another possible limitation concerns the multicenter design of the study, as it may introduce some potential bias related to the number of patients, endoscopists, and pathologists involved in each center. To minimize these biases, we asked each center to involve a similar number of experts and nonexperts and we provided an exploratory subanalysis, according to optical diagnosis expertise, rather than a per-center or per-endoscopist analysis. Concerning pathological diagnosis, a centralized histology evaluation or an external review would have increased the strength of our study.

Lastly, we decided per-protocol to perform optical diagnosis without magnification. Although magnification can further improve the endoscopist optical diagnosis process [20], not all the participating centers had magnifying endoscopes when the study was planned and previous validation studies on optical diagnosis using BLI were performed without magnification [26]. Interestingly, in a recently published paper [25], in which optical diagnosis was performed with CAD-EYE, the authors did not show any significant difference in the performance of optical diagnosis according to the use of zoom. Finally, the sequential step architecture of our study is subject to possible cognitive biases, such as "confirmation bias" and "anchoring bias" [28, 29].

According to our study, real-time AI-assisted optical diagnosis without optical magnification appears to be feasible in clinical practice, allowing the optical diagnosis of > 90 % of DRSPs. Compared with optical diagnosis made by the endoscopists alone, the additional benefit of the AI system in terms of the correct diagnosis seems to be marginal for experts, but it might help nonexperts to meet the thresholds required for incorporation of optical diagnosis in clinical practice.

## Conflict of interest

E. Rondonotti has received speaker's honoraria from Fujifilm Co., is a member of the Fujifilm Co. expert group, and has provided consultancy to Medtronic Co. C. Hassan has equipment on loan from Fujifilm Co. S. Paggi, A. Amato, and R. Maselli have received speaker's honoraria from Fujifilm Co. A. Repici has provided consultancy and has received research grants, not related to the present study, from Fujifilm Co., Medtronic, and Boston Scientific Co. and has provided consultancy to Cosmo Pharmaceuticals S.p.A. and Erbe Elektromedizin GmbH. F. Radaelli has received speaker's honoraria and a research grant, not related to the present study, from Fujifilm Co. The remaining authors declare that they have no conflict of interest.

## Clinical trial

Trial Registration: ClinicalTrials.gov | Registration number (trial ID): NCT04607083 | Type of study: Prospective, Multicenter study

## References


[1] Kessler WR, Imperiale TF, Klein RW et al. A quantitative assessment of the risks and cost savings of forgoing histologic examination of diminutive polyps. Endoscopy 2011; 43: 683–691

[2] Mori Y, Kudo SE, East JE et al. Cost savings in colonoscopy with artificial intelligence-aided polyp diagnosis: an add-on analysis of a clinical trial (with video). Gastrointest Endosc 2020; 92: 905–911

[3] Ignativic A, East JE, Suzuki N et al. Optical diagnosis of small colorectal polyps at routine colonoscopy (Detect InSpect ChAracterise Resect and Discard; DISCARD trial): a prospective cohort study. Lancet Oncol 2009; 10: 1171–1178

[4] Willems P, Djinbachian R, Ditisheim S et al. Uptake and barriers for implementation of the resect and discard strategy: an international survey. Endosc Int Open 2020; 8: E684–E692

[5] Rees CJ, Rajasekhar PT, Wilson A et al. Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. Gut 2017; 66: 887–895

[6] Abu Dayyeh BK, Thosani N. ASGE Technology Committee. et al. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2015; 81: 502.e1–502.e16

[7] Xu Y, Ding W, Wang Y et al. Comparison of diagnostic performance between convolutional neural networks and human endoscopists for diagnosis of colorectal polyp: A systematic review and meta-analysis. PLos One 2021; 16: e0246892

[8] Weigt J, Repici A, Antonelli G et al. Performance of a new integrated computer-assisted system (CADe/CADx) for detection and characterization of colorectal neoplasia. Endoscopy 2022; 54: 180–184

[9] Chen PJ, Lin MC, Lai MJ et al. Accurate classification of diminutive colorectal polyps using computer-aided analysis. Gastroenterology 2018; 154: 568–575

[10] 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: 94–100

[11] Ozawa T, Ishihara S, Fujishiro M et al. Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. Therap Adv Gastroenterol 2020: doi:10.1177/1756284820910659

[12] Zachariah R, Samarasena J, Luba D et al. Prediction of polyp pathology using convolutional neural networks achieves "resect and discard" thresholds. Am J Gastroenterol 2020; 115: 138–144

[13] 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: 357–366

[14] Strobe checklists. Available from: Accessed: 20 April 2021. https://www.strobe-statement.org/index.php?id=available-checklists

[15] Rex DK, Kahi C, O'Brien M et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. Gastrointest Endosc 2011; 73: 419–422

[16] Shaukat A, Kaltenbach T, Dominitz JA et al. Endoscopic recognition and management strategies for malignant colorectal polyps: Recommendations of the US Multi-Society Task Force on Colorectal Cancer. Am J Gastroenterol 2020; 115: 1751–1767

[17] Hassan C, Antonelli G, Dumonceau JM et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline – Update 2020. Endoscopy 2020; 52: 687–700

[18] Bisschops R, East JE, Hassan C et al. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) guideline – Update 2019. Endoscopy 2019; 51: 1155–1179

[19] Dekker E, Houven BBSL, Puig I et al. Curriculum for optical diagnosis training in Europe: European Society of Gastrointestinal Endoscopy (ESGE) position statement. Endoscopy 2020; 52: 899–923

[20] Bisschops R, Hassan C, Bhandari P et al. BASIC (BLI Adenoma Serrated International Classification) classification for colorectal polyp characterization with blue light imaging. Endoscopy 2018; 50: 211–220

[21] Endoscopic Classification Review Group. Update on the Paris classification of superficial neoplastic lesions in the digestive tract. Endoscopy 2005; 37: 570–578

[22] Schlemper MJ, Riddell RH, Kato Y et al. The Vienna classification of gastrointestinal epithelial neoplasia. Gut 2000; 47: 251–255

[23] Simel DL, Samsa GP, Matchar DB. Likelihood ratios for continuous test results–making the clinicians' job easier or harder? J Clin Epidemiol 1993; 46: 85–93

[24] Feise RJ. Do multiple outcome measures require p-value adjustment? BMC Med Res Methodol 2002; 2: 8

[25] Yoshida N, Inoue K, Tomita Y et al. An analysis about the function of a new artificial intelligence, CAD EYE with the lesion recognition and diagnosis for colorectal polyps in clinical practice. Int J Colorectal Dis 2021; 36: 2237–2245

[26] Rondonotti E, Hassan C, Andrealli A et al. Clinical validation of BASIC classification for the resect and discard strategy for diminutive colorectal polyps. Clin Gastroenterol Hepatol 2020; 18: 2357–2365

[27] Houwen BSSL, Hassan C, Coupé VHM et al. Definition of competence standards for optical diagnosis of diminutive colorectal polyps: European Society of Gastrointestinal Endoscopy position statement. Endoscopy 2022; 54: 88–99

[28] Elston DM. Confirmation bias in medical decision making. J Am Acad Dermatol 2020; 82: 572

[29] Glick M. Believing is seeing. J Am Dental Assoc 2017; 148: 131–132