Thieme

White light computer-aided optical diagnosis of diminutive colorectal polyps in routine clinical practice

Authors

Emanuele Rondonotti[1], Irene Maria Bambina Bergna[2,1,3], Silvia Paggi[1], Arnaldo Amato[1,3], Alida Andrealli[1], Giulia Scardino[1], Giacomo Tamanini[1], Nicoletta Lenoci[1], Giovanna Mandelli[1], Natalia Terreni[1], SImone Rocchetto[1,2], Alessandra Piagnani[1,2], Dhanai Di Paolo[1], Niccolò Bina[1], Emanuela Filippi[4], Luciana Ambrosiani[4], Cesare Hassan[5,6], Loredana Correale[5], Franco Radaelli[1]

Institutions

1 Gastroenterology Unit, Valduce Hospital, Como, Italy
2 University of Milan, Milano, Italy
3 Gastroenterology and Digestive Endoscopy Unit, Alessandro Manzoni Hospital, Lecco, Italy
4 Pathology Department, Valduce Hospital, Como, Italy
5 Department of Gastroenterology, IRCCS Humanitas Research Hospital, Rozzano, Italy
6 Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Italy

Keywords

Endoscopy Lower GI Tract, Polyps / adenomas / ..., Diagnosis and imaging (inc chromoendoscopy, NBI, iSCAN, FICE, CLE...), Colorectal cancer

received 5.7.2023
accepted after revision 4.4.2024

Bibliography

Endosc Int Open 2024; 12: E676–E683
DOI 10.1055/a-2303-0922
ISSN 2364-3722

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

Corresponding author

Dr. Emanuele Rondonotti, Valduce Hospital, Gastroenterology Unit, via Dante 10, 22100 Como, Italy
ema.rondo@gmail.com

Supplementary Material is available at https://doi.org/10.1055/a-2303-0922

ABSTRACT

Background and study aims Artificial Intelligence (AI) systems could make the optical diagnosis (OD) of diminutive colorectal polyps (DCPs) more reliable and objective. This study was aimed at prospectively evaluating feasibility and diagnostic performance of AI-standalone and AI-assisted OD of DCPs in a real-life setting by using a white light-based system (GI Genius, Medtronic Co, Minneapolis, Minnesota, United States).

Patients and methods Consecutive colonoscopy outpatients with at least one DCP were evaluated by 11 endoscopists (5 experts and 6 non-experts in OD). DCPs were classified in real time by AI (AI-standalone OD) and by the endoscopist with the assistance of AI (AI-assisted OD), with histopathology as the reference standard.

Results Of the 480 DCPs, AI provided the outcome "adenoma" or "non-adenoma" in 81.4% (95% confidence interval [CI]: 77.5–84.6). Sensitivity, specificity, positive and negative predictive value, and accuracy of AI-standalone OD were 97.0% (95% CI 94.0–98.6), 38.1% (95% CI 28.9–48.1), 80.1% (95% CI 75.2–84.2), 83.3% (95% CI 69.2–92.0), and 80.5% (95% CI 68.7–82.8%), respectively. Compared with AI-standalone, the specificity of AI-assisted OD was significantly higher (58.9%, 95% CI 49.7–67.5) and a trend toward an increase was observed for other diagnostic performance measures. Overall accuracy and negative predictive value of AI-assisted OD for experts and non-experts were 85.8% (95% CI 80.0–90.4) vs. 80.1% (95% CI 73.6–85.6) and 89.1% (95% CI 75.6–95.9) vs. 80.0% (95% CI 63.9–90.4), respectively.

Conclusions Standalone AI is able to provide an OD of adenoma/non-adenoma in more than 80% of DCPs, with a high sensitivity but low specificity. The human-machine interaction improved diagnostic performance, especially when experts were involved.

Introduction

Detection and removal of polyps and adenomas during colonoscopy has been shown to reduce the risk of colorectal cancer [1, 2, 3, 4]. However, more than 90% of polyps detected during colonoscopy are small (6–9 mm) or diminutive (≤5 mm) and, especially in the latter group, the risk of cancer is negligible [5, 6, 7, 8, 9, 10, 11, 12, 13, 14]. Currently, all polyps identified during colonoscopy are removed by endoscopic resection and sent for histopathology; this policy is associated with significant cost, time, and potential adverse events.

In recent years, it has been consistently demonstrated that implementation of strategies based on optical diagnosis (OD) of diminutive colorectal polyps (DCPs) (i. e., leave-in-situ and resect-and-discard policies) reduces pathology- and polypectomy-related costs, leading to a significant economic benefit without an impact on efficacy and safety [15, 16]. The American Society for Gastrointestinal Endoscopy has set Preservation and Incorporation of Valuable endoscopic Innovations (PIVI) standards for safe and effective incorporation of such strategies in clinical practice [17, 18]. More recently, the European Society of Gastrointestinal Endoscopy (ESGE) developed simple, safe, and easy-to-measure competence standards for OD which could optimize OD decision-making in endoscopy [19, 20]. Although several studies have shown that image enhancement technologies in expert hands can meet the PIVI criteria [21, 22, 23, 24], integration of OD in routine clinical practice has failed due to the inability of general endoscopists to meet the established standards [21, 22, 23, 24, 25]. A recent international survey investigating clinical uptake of these strategies showed that 84% of endoscopists practicing across the world were not using leave-in-situ and/or resect-and-discard strategy [26].

In the last few years, deep learning allowed the development of artificial intelligence (AI) algorithms that are able to characterize polyps in real time, potentially making the OD process easy, quick, objective, and reproducible, regardless of endoscopist expertise or training in optical diagnosis. However, at present, most AI systems for polyp characterization involve proprietary image enhancement technologies (e. g., endocytoscopy [27], narrow-band imaging (NBI) [28, 29], blue light imaging (BLI) [30]) and are evaluated in artificial settings (e. g., by analyzing still images or short video segments). Recently a brand-new AI system for polyp characterization has been marketed (GI Genius Medtronic Inc, Dublin, Ireland). It works in white light unmagnified endoscopy, automatically activates when a polyp is detected, overlays a frame-by-frame live OD, and potentially can be suitable for all endoscopy platforms. This AI system has been initially validated on a prospectively acquired dataset with a multi-reader study design [31, 32],but clinical data about its clinical performances in a real-life setting are still limited [24, 32, 33].

The study aim was to prospectively evaluate the feasibility and clinical performance of AI (both standalone AI and AI-assisted) OD by using GI Genius module for characterization of diminutive colorectal polyp in a real-life setting.

Patients and methods

Center

This prospective observational study was conducted at the Endoscopy Unit of Ospedale Valduce Como, Italy. The institutional review board (Comitato Etico dell'Insubria, ASST Sette Laghi) approved the protocol (unique identifier number 41/2022- approval date 28/07/2022) which was registered with ClinicalTrial platform (NCT05492656) on 11/08/2022.

Patients and inclusion and exclusion criteria

Consecutive adult outpatients aged 18 to 80 years referred for screening, symptoms, or post-polypectomy surveillance colonoscopy in which at least one diminutive (≤ 5 mm) DCP was detected were included in the present study. Patients with increased risk of harboring adenomatous lesions (e. g., history of colorectal cancer, hereditary polyposis syndrome or hereditary non-polyposis colorectal cancer) were excluded. Furthermore, patients with newly diagnosed inflammatory bowel disease, polypectomy not performed because of ongoing anticoagulation, and urgent colonoscopy scheduled were also excluded from the present study. All patients provided their written informed consent.

Study outline

All patients undergoing outpatient colonoscopy at the Gastroenterology Unit of Valduce Hospital during the study timeframe were offered the opportunity to participate in the study, provided they met all the specified inclusion criteria and none of the exclusion criteria. When diminutive colon polyps were detected during colonoscopy, excluding polyps < 3 mm in the rectum with an obvious hyperplastic appearance, the GI Genius system in white light was employed for polyp characterization. Moreover, the AI system's ability to generate an output (adenoma, non-adenoma, or no prediction) and to provide clear-cut relevant information for subsequent polyp management (namely the adenoma or non-adenoma output) was measured. The same polyps underwent assessment by the endoscopist, who was aware of the AI system output (AI-assisted OD). The endoscopist then provided his/her evaluation (adenoma or non-adenoma) along with the confidence level. Further details regarding each step of the study are outlined in the following paragraphs.

Study aim and study endpoints

The primary aim was to assess, in routine clinical practice, the feasibility (i. e. the capability of AI to provide a clear outcome) of standalone AI and AI-assisted OD (see definitions in "Artificial Intelligence System and Optical Diagnosis process" section) in diagnosis of DCPs. For this purpose, we considered as primary endpoints the proportion of DCPs in which the standalone AI system was able to provide an output, such as "adenoma," "non-adenoma," or "no prediction" and the proportion of DCPs in which the standalone AI system was able to provide clinically relevant output (namely "adenoma" or "non-adenoma") which was potentially useful for further polyp management. We also evaluated feasibility of AI-assisted OD by calcu-

lating the proportion of diminutive polyps characterized by an endoscopist once they were aware of the AI output and the proportion of diminutive polyps characterized by an endoscopist, once they were aware of the AI output, with high confidence.

The secondary aim was to evaluate diagnostic accuracy. For this purpose, we calculated sensitivity, specificity, and positive and negative predictive value as well as overall accuracy of standalone AI and AI-assisted OD in characterizing DCPs with histopathology as reference standard (DCPs in which the histopathology was unavailable – e. g. non-retrieved polyps – were excluded from the analysis). For calculation of standalone AI OD accuracy variables, only DCPs characterized as adenomas or non-adenomas were included, whereas for AI-assisted OD accuracy, only DCPs characterized with high confidence were selected.

We also planned exploratory subgroup analyses for both standalone AI and AI-assisted OD on accuracy of performance according to level of expertise (i. e., expert vs. non-expert), level of confidence (i. e., high vs. low) and polyp location (i. e., proximal *vs.* distal; see "Endoscopic procedures" section for definitions).

Endoscopic procedures

All procedures were performed using the ELUXEO 7000 endoscopy platform (including video processor ELUXEO VP-7000 and light source ELUXEO BL-7000; Fujifilm Co., Tokyo, Japan). All endoscopists, before taking part in the study, received formal training, consisting of a 45-minute lecture about principles of OD and features of the AI system used in the present study. During the study period, neither periodical auditing nor monitoring of OD performance were performed in order to avoid potential bias. The participating endoscopists were dichotomized as experts (whether they followed dedicated training, participated in previous studies of OD, underwent periodical auditing and monitoring, and performed OD on a regular basis, according to ESGE curricula [19, 20]), and non-experts in OD.

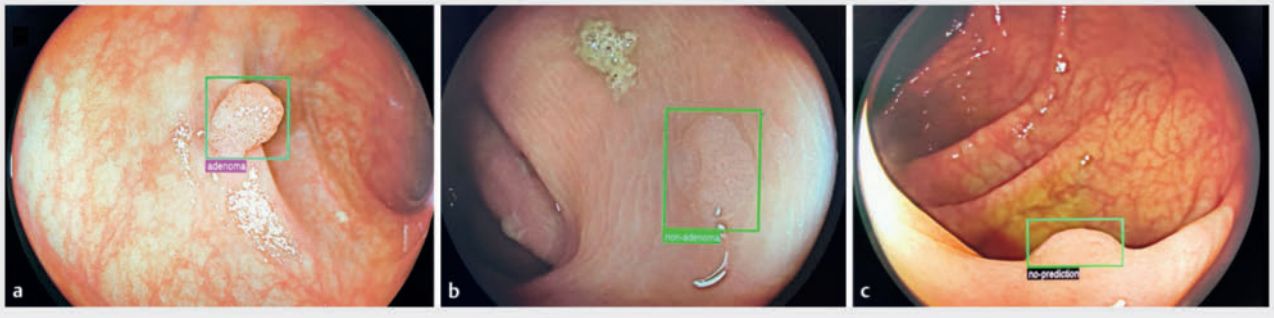
According to the current clinical practice in our center, all polyps except tiny ones (1–3 mm) located in the rectum with obvious hyperplastic appearance were removed, retrieved in separate jars, and sent for pathology assessment. Polyp size was estimated by comparison with an open snare or forceps; moreover, they were classified according to size as 1 to 5 mm (diminutive), 6 to 9 mm (small) or ≥10 mm. In the present paper, the analysis was limited to diminutive colonic polyps. According to current practice [34], diminutive polyps were defined as "proximal" (PDCPs) if located in the cecum, ascending colon, transverse colon, or splenic flexure; in the remaining cases (polyps located in the descending or sigmoid colon, or rectum) they were labeled as "distal" (DDCPs). Polyp morphology was described according to the Paris classification [35].

AI system and OD process

In the present study, the convolutional neural network system GI Genius v. 3.0.1 was used for polyp characterization. The system has been described in detail elsewhere [31, 32]. Briefly, it processes non-magnified white light images in nearly real-time. When a polyp is framed within the detection box, a three-option OD ("adenoma", "non-adenoma" or "no prediction") is immediately overlaid (AI-standalone OD). ▶**Fig. 1** shows how the AI output is provided. Because standalone AI OD output may change during observation (because of factors such as movements of the endoscope or patien and presence of debris), the output which was more consistent throughout observation (which was at least 30 seconds) was recorded as standalone AI OD output. After receiving the AI output, the endoscopist was free to shift from white to blue light (BLI) and to provide his own OD (AI-assisted OD). The level of confidence in AI-assisted OD (high vs. low) was also recorded.

Pathology (reference standard)

All resected polyps or biopsy specimens were fixed in a 10% buffered formalin solution and sent to pathology in separate jars. They were processed and stained for histopathology using standard methods and evaluated by expert pathologists. Participating pathologists were blinded to the OD, and evaluated all the resected polyp according to the Vienna classification and World Health Organization guidelines [36, 37]. For the present study, hyperplastic polyps, sessile serrated lesions with or without dysplasia, inflammatory polyps, and normal mucosa samples were all labeled as non-adenomas. Adenomas with significant villous features (>25%), size ≥1.0 cm, high-grade dysplasia, or early invasive cancer were defined as advanced.



▶**Fig. 1** This frame shows how the artificial intelligence (AI) system provides the optical diagnosis output; when the polyp is framed within the detection box, an optical diagnosis (OD) label is overlaid (a adenoma, b non-adenoma, c no prediction).

Statistical analysis and sample size calculation

Normally and non-normally distributed data were presented using mean and standard deviation (SD) and median and interquartile ranges (IQRs), respectively.

For all the estimates reported in the present study, 95% confidence intervals (95% CIs) were calculated. To evaluate paired nominal data (e. g. standalone AI and AI-assisted feasibility) a two-tailed McNemar test was used. P <0.05 was considered statistically significant. In assessing the diagnostic accuracy variables of the standalone AI and AI-assisted approaches, it is important to note that although the population samples were not identical, they exhibited significant overlap, almost to the point of being nearly identical. Consequently, the McNemar test and Chi-square test were not entirely suitable for this context. Thus, we presented absolute numbers along with their corresponding 95% CIs. Significance was attributed to differences only when there was no overlap in the 95% CIs [38]. There was no need to adjust for multiplicity because the analysis of secondary outcomes was considered subsidiary and exploratory.

Satistical analysis was performed with Microsoft Excel (Microsoft Co., Redmonton, Washington, United States) and MedCalc Software package (MedCalc Software Ltd., Ostend, Belgium).

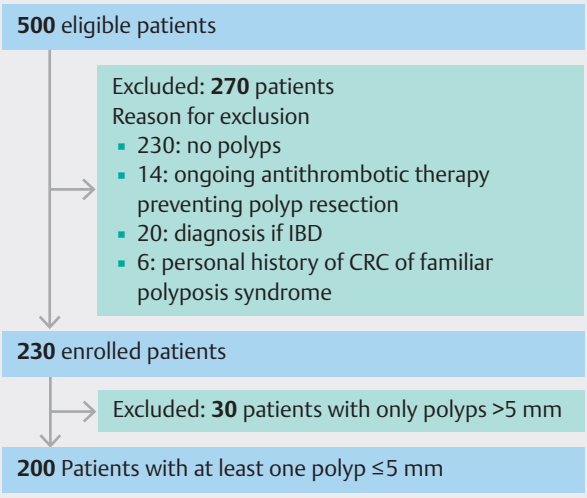
Taking into account the lack of studies evaluating the clinical feasibility and diagnostic accuracy of the AI system when the current study was designed, we considered our trial as a preliminary pilot study. Thus, we arbitrarily planned to include 500 consecutive patients overall. Moreover, taking into account the polyp detection rate in our center in 2001 (46.0%) in patients undergoing colonoscopy for mixed indication, we expected to collect data from at least 230 patients with at least one polyp. Because DCPs represents about 95% of all detected polyps and they are often multiple, we expected to collect about 450 DCPs.

Results

Patients and polyps

From August to November 2022, 500 outpatients referred for screening, surveillance, or diagnostic colonoscopy were evaluated. Of them, 270 were excluded (reasons for exclusion are detailed in ▶ **Fig. 2**) and 30 had only non-diminutive polyps (> 5 mm) and were also excluded from further analysis. Finally, 200 patients (median age 66.0 years, IQR 57–73, 55.0% male) with at least one DCP were included in the analysis. Patient flow is summarized in ▶ **Fig. 2**. In these patients, 551 polyps were identified. Baseline features of patients and polyps are summarized in **Supplementary Table 1**. The median number of polyps per patients was two (IQR 1–4) (**Supplementary Table 2**).

Of 551 resected polyps, 480 were DCPs (480/551, 87.1%). Twenty (4.2%; 95% CI 2.6–6.4) were not retrieved. Histological diagnosis, therefore, was available or 460 DCPs, of which 298 (64.8%; 95% CI 64.8–69.2) were adenomas and 162 (35.2%; 95%CI 30.9–39.8) non-adenomas, namely 99 hyperplastic polyps, eight sessile serrated lesions, and 55 inflammatory tissue or normal mucosa. Of 460 retrieved DCPs, 188 (188/460, 40.9%; 95% CI 36.3–45.5) were DDCPs (**Supplementary Table 3** summarizes the features of retrieved DCPs).



▶ **Fig. 2** Patient flow.

Overall, 11 endoscopists participated in the present study; of them, five were experts and six non-experts. Additional details regarding endoscopist selection and features are reported in the supplementary material. Expert endoscopists resected 231 DCPs (231/480, 48.1%; 95% CI 43.6–52.7) and retrieved 221 of them (221/460, 48.0%; 95% CI 43.4- 52.7); the median number of DCPs evaluated by each expert was 11 (IQR 3–107). Non-experts resected 249 DCPs (249/480, 51.9%; 95% CI 47.3–56.4) and retrieved 239 of them (239/460, 52.0%; 95% CI 47.3–56.6); the median number of DCPs evaluated by each non-expert was 12 (IQR 3–120).

Feasibility of AI-standalone and AI-assisted OD

AI output was obtained for all 480 diminutive polyps identified. The output of the AI system was “adenoma” or “non-adenoma” in 341 (341/480, 71.0 %; 95% CI 66.8–75.1) and in 50 DCPs (50/480, 10.4%; 95% CI 7.8–13.5) respectively. The AI output was “no prediction” in the remaining 89 diminutive polyps (89/480, 18.5%; 95% CI 15.2–22.3).

Of 480 DCPs, the endoscopist was able to provide an outcome in all of them. However, the DCPs evaluated with high confidence by the endoscopist with the assistance of the AI system were 392/480, 81.7% (95% CI 77.9–85.0). The rate of polyps evaluated as adenoma/non-adenoma by AI alone and those evaluated as adenoma/non-adenoma with high confidence by the endoscopist, once they were aware of the AI output, according to polyp location (proximal vs. distal) and expertise (expert vs. non-expert), are detailed in ▶ **Table 1**.

Standalone AI OD diagnostic performances

A total of 374 DCPs were retrieved for pathological evaluation in which standalone AI OD was “adenoma” or “non-adenoma”. For DCPs the standalone AI sensitivity, specificity, negative, and positive predictive value as well as overall accuracy were 97%

▶ **Table 1** Rate of polyps evaluated as adenoma/non-adenoma by AI alone (AI-standalone) and evaluated as adenoma/non-adenoma with high confidence by the endoscopist once aware of AI output (AI-assisted).

	All endoscopists			Expert endoscopists			Non-experts endoscopists		
	Standalone AI	AI-assisted	*P* value	Standalone AI	AI-assisted	*P* value	Standalone AI	AI-assisted	*P* value
DCP (95% CI)	81.4% (77.5–84.6)	81.7% (77.9–85.0)	0.940	79.7% (73.9–84.7)	85.7% (73.9–84.7)	0.146	83.1% (77.9–87.6)	77.9% (72.2–82.9)	0.223
Proximal (95% CI)	82.1% (77.1–86.4)	78.9% (73.6–83.5)	0.443	82.7% (75.4–88.6)	84.2% (77.0–89.8)	0.882	81.1% (73.7–87.2)	73.4% (65.4–80.5)	0.214
Distal (95% CI)	80.1% (73.9–85.4)	85.6% (73.6–83.5)	0.228	75.0% (64.9–83.5)	89.0% (80.7–94.6)	0.059	85.9% (77.7–91.9)	84.0% (75.6–90.4)	0.859

AI, artificial intelligence; DCP, diminutive colorectal polyp; CI, confidence interval.

▶ **Table 2** Diagnostic performance in AI-standalone and AI-assisted optical diagnosis of all diminutive colorectal polyps.

	AI-standalone	AI-assisted
Number of polyps	374	376
Sensitivity (95% CI)	97.0% (94.0–98.6)	94.8% (91.1–97.1)
Specificity (95% CI)	38.1% (28.9–48.1)	58.9% (49.7–67.5)
Positive predictive value (95% CI)	80.1% (75.2–84.2)	82.4% (77.4–86.5)
Negative predictive value (95% CI)	83.3% (69.2–92.0)	84.9% (75.2–91.4)
Accuracy (95% CI)	80.5% (68.7–82.8)	83.0% (78.8–86.6)

AI, artificial intelligence; CI, confidence interval.

▶ **Table 3** Diagnostic performance in standalone AI and AI-assisted optical diagnosis according to polyp location and endoscopist expertise.

	Proximal diminutive polyps		Distal diminutive polyps		Experts		Non-experts	
	Standalone AI	AI-assisted	Standalone AI	AI-assisted	Standalone AI	AI-assisted	Standalone AI	AI-assisted
Number of polyps	222	213	152	163	176	190	199	186
Sensitivity (95% CI)	97.1% (93.1–98.9)	96.9% (92.4–98.8)	96.8% (90.3–99.2)	91.4% (83.3–95.9)	96.1% (90.7–98.6)	96.1% (90.6–98.5)	97.9% (93.4–99.4)	93.6% (87.4–97.0)
Specificity (95% CI)	31.9% (19.5–47.3)	51.9% (38.0–65.5)	43.1% (30.4–56.7)	64.3% (51.9–75.1)	36.0% (23.1–51.5)	65.0% (51.9–76.4)	39.7% (27.3–53.4)	52.5% (39.4–65.2)
PPV (95% CI)	84.2% (78.2–88.0)	85.6% (79.4–90.2)	73.4% (64.6–80.7)	77.2% (68.1–84.5)	80.5% (73.2–86.3)	84.7% (77.6–90.0)	79.8% (72.3–85.3)	80.1% (72.5–86.1)
NPV (95% CI)	75.0% (50.6–90.4)	84.8% (67.3–94.3)	89.3% (70.6–97.2)	84.9% (71.9–92.8)	77.3% (54.2–91.3)	89.1% (75.6–95.9)	88.5% (68.7–97.0)	80.0% (63.9–90.4)
Accuracy (95% CI)	83.3% (77.7–88.0)	85.5% (80.0–89.9)	76.3% (68.7–82.8)	79.8% (72.8–85.6)	80.1% (73.3–85.7)	85.8% (80.0–90.4)	80.9% (74.8–86.1)	80.1% (73.6–85.6)

AI, artificial intelligence; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

(95% CI 94.0–98.6), 38.1% (95% CI 28.9–48.1), 83.3% (95% CI 69.2–92.0), 80.1% (95% CI 75.2–84.2), and 80.5% (95% CI 76.2–84.4), respectively (▶ **Table 2**). Results of standalone AI OD subanalysis according to polyp location and endoscopist expertise are reported in ▶ **Table 3**.

AI-assisted OD diagnostic performances

Overall, 376 DPCs were characterized by the endoscopist with the assistance of AI (AI-assisted OD) with high confidence. The AI-assisted sensitivity, specificity, negative and positive predictive value, and overall accuracy were 94.8% (95% CI 91.1–97.1), 58.9% (95% CI 49.7–67.5), 84.9% (95% CI 75.2–91.4), 82.4% (95% CI 77.4–86.5), and 83.0% (95% CI 78.8–86.6), respectively (▶ **Table 2**).

A relevant increase in specificity (58.9%; 95% CI 49.7–67.5 vs.38.1%; 95% CI 28.9–48.1) and a trend toward an increase in other performance variables (▶ **Table 2**) was observed by comparing standalone AI OD and AI-assisted OD diagnostic performances. Results of AI-assisted OD subanalysis according to polyp location and expertise are reported in ▶ **Table 3.**

Discussion

Our data show that in clinical practice, an AI white light-based system provides an OD (i. e., adenoma or non-adenoma) that is potentially clinically useful for further polyp management for more than 80% of diminutive colonic polyps, with good overall accuracy, but low specificity. The rate of DCPs labeled as adenoma or non-adenoma by AI was high, but it was somewhat lower than the one reported in previous studies conducted in a similar setting [32, 33, 39, 40, 41]. This finding relates to a technical peculiarity of GI Genius, namely the presence of three possible outcomes and particularly the "non-predictable" one, which is not available in other AI systems for polyp characterization. While this option reduces by about 20% the portion of diminutive polyps for which the system can provide definite dichotomous adenoma/non-adenoma feedback, it may help to simplify the characterization process and decrease the subjectivity of interpretation by quickly and clearly identifying the subset of diminutive polyps for which an automatic assessment may be less reliable.

We found that standalone AI had good accuracy, even though it resulted in slightly lower accuracy than that reported in previous studies in which the same AI system was used [32, 33, 40, 41]. In the present study we observed low specificity. This finding might have a relevant impact on the cost-effectiveness of the resect-and-discard and leave-in process, because on one hand, low specificity minimizes risk of leaving adenomatous polyps in place, but it also potentially increases the number of hyperplastic polyps unnecessarily removed, reducing possible cost savings, which is one of the major drivers of an OD-based policy. This finding sharply differs from the standalone AI specificity reported by other trials [27, 32, 33, 39, 41], which were explicitly designed and powered to target the leave-in-situ strategy threshold, and in which all rectosigmoid diminutive polyps were included. Conversely, in line with current routine practice in most endoscopy centers, we deliberately decided to exclude very tiny (1–3 mm) rectal polyps, with obvious hyperplastic appearance, which were evaluated by the endoscopist alone and left in situ. This policy led us to select a subset of polyps rich in adenomas, because in the present study, the rate of adenomatous DDCPs was 70.0%, whereas in the study by Hassan et al. [33], this figure was less than 20%. However, this represents the actual real-life scenario and the polyps we selected were exactly those with which endoscopists can benefit the most from the presence of an automatic, independent characterization system to support them in decision making.

It could also be speculated that the performance of standalone AI also may be affected by inclusion of sessile serrated lesions (SSLs) within the category of hyperplastic polyps. Although this is certainly one of the limitations of currently available AI systems and the possibility of differentiating SSLs from hyperplastic polyps will represent a step forward in this technology in the near future, diminutive polyps with SSL features represent a very small proportion of all diminutive colonic polyps (1.7% in the present study); therefore, it does not seem to represent a major issue at the present time.

Notably, we observed that when the AI output was coupled with endoscopist evaluation, all performance measures tended to improve, with an increase in both colorectal and diminutive distal polyp OD specificity. This finding reinforces the key role of the endoscopist in the OD process, and the superiority of so-called hybrid intelligence over standalone AI output. In fact, endoscopists do not merely record output provided by the AI system, but they integrate it within a complex cognitive process that takes into account their assessment, the AI output, and also other technical and clinical data (e. g., the location of the polyp) in order to provide the final OD. This seems to be confirmed by the subanalysis of AI-assisted OD we performed according to endoscopist expertise. We observed that the performance of experts tended to be superior to that of non-experts. In particular, when we looked at DDCPs, which are the target of the leave-in situ strategy, the results of the experts in our study were very close to the values observed in the study by Hassan et al. [33,] which mostly involved expert endoscopists, and to the thresholds required for implementation of a leave-in situ strategy in clinical practice.

Our study had some limitations. First, as already mentioned, we designed a feasibility pilot study, with an arbitrary sample size of 500 patients, because when the study was planned clinical data were not available. Second, as anticipated, we excluded some very tiny (1–3 mm) rectal polyps with obvious hyperplastic appearance and we included among distal polyps those located in the descending colon, so we cannot draw definitive conclusions regarding the use of the AI system in the framework of resect-and-discard or leave-in policies. However, in the present feasibility study, we took a pragmatic approach from clinical practice, where anatomical landmarks for distinguishing between the descending and sigmoid colon are often lacking (while it is easier to use the splenic flexure as a reference) and very tiny polyps in the distal rectum do not need any additional tool to be correctly recognized as hyperplastic. We also focused mostly on feasibility and accuracy, instead of negative predictive value or sensitivity and specificity as suggested by PIVI [21, 24] or Simple Optical Diagnosis Accuracy (SODA) [20] thresholds. In addition, although inclusion of a high number of heterogenous endoscopists might represent a reliable picture of everyday clinical practice, differences in colonoscopy tech-

nique such as endoscope handling, polyp framing, and colonic lumen cleansing and the relatively low number of DCPs evaluated by each endoscopist could have affected the final results. To minimize this issue, we decided to analyze aggregate data by comparing experts and non-experts, instead of providing per-endoscopist data. Furthermore, in the present study, we did not collect the characterization of DCPs by an endoscopist alone (i.e., without the support of AI). Third, although in our study some patients presented with more than one polyp, we did not apply a correction for possible statistical dependency. From a theoretical standpoint, characterizing multiple polyps might be more straightforward when they are obtained from the same patient, rather than from different patients. Nevertheless, we opted to exclude patients with predisposing conditions (such as a history of colorectal cancer, hereditary polyposis syndrome, or hereditary non-polyposis colorectal cancer) that might lead to the presence of polyp clusters with a specific histotype. Our approach closely aligns with methodologies employed in other studies addressing the same topic. In addition, considering the median number of polyps per patient (2, IQR 1–4) and the relatively low occurrence of patients with multiple polyps (**Supplementary material Table 2**), we anticipate a negligible impact on the analysis results by not adjusting for multiplicity. Lastly, due to the intrinsic characteristics of the AI system we evaluated (GI Genius, Medtronic Co), which provides immediate output of characterization once a polyp is identified, we were unable to evaluate standalone endoscopist performance, which would be the ideal comparator to better understand system performance and, most importantly, the contribution of AI to the OD process. To address this issue, an additional screen with the native endoscopic image (without the AI output overlay) would be needed. Unfortunately this was not available at our center during the study period.

Conclusions

In conclusion, white light standalone AI OD showed good feasibility and acceptable overall accuracy in everyday clinical practice, with a low specificity that might hamper the cost-effectiveness of OD-based strategies, such as leave-in-situ and resect and discard. We reported an increase in some relevant performance variables when humans interacted with the machine, thus supporting the superiority of the so-called hybrid intelligence over the AI-standalone performance, which is consistent with other studies [27, 30, 40], especially when there is a high level of expertise in OD and when polyps difficult to characterize are included. If one of the possible advantages of AI is standardization of the OD process in small and inexperienced centers, the results of our study do not fully support that expectation. However, we used the first available GI Genius release and further improvements are expected with the second-generation software, which is going to be launched soon.

Acknowledgement

The authors sincerely thank Costantini A and Facciorusso A for support in the editing process.

Conflict of Interest

Paggi S and Amato A: speaker honoraria from Fujifilm Co. Rondonotti E: speaker honoraria from Fujifilm Co., Medtronic Co. consultancy. Radaelli F: Speaker honoraria from Fujifilm Co; research grant from Fujifilm Co; endoscopy equipment loan from Medtronic Co. Hassan C: Medtronic Co, Fujifilm Co and Odin Co. consultancy. All the other authors declare no conflict of interest.

Clinical trial

ClinicalTrials.gov (http://www.clinicaltrials.gov/)
Registration number (trial ID): NCT05492656
Type of Study: Prospective single centre

References

[1] Hill MJ, Morson BC, Bussey HJR. Aetiology of adenoma- carcinoma sequence in large bowel. Lancet 1978; 311: 245–247 doi:10.1016/s0140-6736(78)90487-7

[2] Winawer SJ, Zauber AG, Ho MN et al. Prevention of colorectal cancer by colonoscopic polypectomy. N Engl J Med 1993; 329: 1977–1981 doi:10.1056/NEJMoa1100370

[3] Kaminski MF, Regula J, Kraszewska E et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010; 362: 1795–1803 doi:10.1056/NEJMoa0907667

[4] Kaminski MF, Wieszczy P, Rupinski M et al. Increased rate of adenoma detection associated with reduced risk of colorectal cancer and death. Gastroenterology 2017; 153: 98–105

[5] Butterly LF, Chase MP, Pohl H et al. Prevalence of clinically important histology in small adenomas. Clin Gastroenterol Hepatol 2006; 4: 343–348 doi:10.1016/j.cgh.2005.12.021

[6] Yoo TW, Park DI, Kim YH et al. Clinical significance of small colorectal adenoma less than 10 mm: The KASID study. Hepatogastroenterology 2007; 54: 418–421

[7] Ponugoti PL, Cummings OW, Rex DK. Risk of cancer in small and diminutive colorectal polyps. Dig Liver Dis 2017; 49: 34–37 doi:10.1016/j.dld.2016.06.025

[8] Gupta N, Bansal A, Rao D et al. Prevalence of advanced histological features in diminutive and small colon polyps. Gastrointest Endosc 2012; 75: 1022–1030 doi:10.1016/j.gie.2012.01.020

[9] Denis B, Bottlaender J, Weiss AM et al. Some diminutive colorectal polyps can be removed and discarded without pathologic examination. Endoscopy 2011; 43: 81–86

[10] Tsai FC, Strum WB. Prevalence of advanced adenomas in small and diminutive colon polyps using direct measurement of size. Dig Dis Sci 2011; 56: 2384–2388 doi:10.1007/s10620-011-1598-x

[11] Chaput U, Alberto SF, Terris B et al. Risk factors for advanced adenomas amongst small and diminutive colorectal polyps: A prospective monocenter study. Dig Liver Dis 2011; 43: 609–612 doi:10.1016/j.dld.2011.02.002

[12] Lieberman D, Moravec M, Holub J et al. Polyp size and advanced histology in patients undergoing colonoscopy screening: implications for CT colonography. Gastroenterol 2008; 135: 1100–1105 doi:10.1053/j.gastro.2008.06.083

[13] Rex DK, Overhiser AJ, Chen SC et al. Estimation of impact of American College of Radiology recommendations on CT colonography reporting for resection of high-risk adenoma findings. Am J Gastroenterol 2009; 104: 149–153 doi:10.1038/ajg.2008.35

[14] Hewett DG, Huffman ME, Rex DK. Leaving distal colorectal hyperplastic polyps in place can be achieved with high accuracy by using narrow-band imaging: an observational study. Gastrointest Endosc 2012; 76: 374–380 doi:10.1016/j.gie.2012.04.446

[15] Wilson AI, Saunders BP. New paradigms in polypectomy: resect and discard, diagnose and disregard. Gastrointest Endosc Clin N Am 2015; 25: 287–302 doi:10.1016/j.giec.2014.12.001

[16] Hassan C, Pickhardt PJ, Rex DK. A resect and discard strategy would improve cost effectiveness of colorectal cancer screening. Clin Gastroenterol Hepatol 2010; 10: 865–869 doi:10.1016/j.cgh.2010.05.018

[17] Abu Dayyeh BK, Thosani N, Konda V 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: 501–516 doi:10.1016/j.gie.2014.12.022

[18] 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 doi:10.1016/j.gie.2011.01.023

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

[20] Houwen BBSL, Hassan C, Coupé VMH et al. Definition of competence standards for optical diagnosis of diminutive colorectal polyps: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy 2022; 54: 88–99

[21] Rees C, Rajashekar P, 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

[22] Neumann H, Vieth M, Bisscops R et al. Leaving colorectal polyps in place can be achieved with high accuracy using blue light imaging (BLI). United European Gastroenterol J 2018; 6: 1099–1105

[23] Rondonotti E, Paggi S, Amato A et al. Blue-light imaging com- pared with high-definition white light for real-time histology pre- diction of colorectal polyps less than 1 centimeter: a prospective randomized study. Gastrointest Endosc 2019; 89: 554–564

[24] Neumann H, Fujishiro M, Wilcox C et al. Present and future perspectives of virtual chromoendoscopy with i-scan and optical enhancement technology. Dig Endosc 2013; 26: 43–51

[25] Vleugels JLA, Hazewinkel Y, Dijkgraa MGW. Optical diagnosis expanded to small polyps: Post-hoc analysis of diagnostic performance in a prospective multicenter study. Endoscopy 2019; 51: 244–252

[26] 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 doi:10.1055/a-1132-5371

[27] 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

[28] Gross S, TrautweinCBehrens A et al. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. Gastrointest Endosc 2011; 74: 1354–1359

[29] Kominami Y, Yoshida S, Tanaka S et al. Computer-aided diagno-sis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. Gastrointest Endosc 2016; 83: 643–649

[30] 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 doi:10.1055/a-1372-0419

[31] Biffi C, Salvagnini P, Dinh NN et al. A novel AI device for real-time optical characterization of colorectal polyps. NPJ Dig Med 2022; 5: 84 doi:10.1038/s41746-022-00633-6

[32] Hassan C, Balsamo G, Lorenzetti R et al. Artificial intelligence allows leaving-in-situ colorectal polyps. Clin Gastroenterol Hepatol 2022; 20: 2505–2513

[33] Hassan C, Sharma P, Mory Y et al. Comparative performance of artificial intelligence optical diagnosis systems for leaving in situ colorectal polyps. Gastroenterology 2022; 164: 467–469.e4

[34] Senore C, Bellisario C, Segnan N. Distribution of colorectal polyps: implication for screening. Best Pract Res Clin Gastroenterol 2017; 31: 481–488 doi:10.1016/j.bpg.2017.04.008

[35] Endoscopic Classification Review Group. Update on the Paris classification of superficial neoplastic lesions in the digestive tract. Endoscopy 2005; 37: 570–578 doi:10.1055/s-2005-861352

[36] Schlemper MJ, Riddell RH, Kato Y et al. The Vienna classification of gastrointestinal epithelial neoplasia. Gut 2000; 47: 251–255 doi:10.1136/gut.47.2.251

[37] Pai RK, Makinen MJ, Rosty C. WHO Classification of Tumours. Digestive System Tumours. 5th Edition. IARC; 2019

[38] Payton ME, Greenstone MH, Schenker N. Overlapping confidence intervals or standard error intervals: what do they mean in terms of statistical significance? J Insect Sci 2003; 3: 34 doi:10.1093/jis/3.1.34

[39] Rondonotti E, Hassan C, Tamanini G et al. Artificial intelligence-assisted optical diagnosis for the resect and discard strategy in clinical practice: the Artificial intelligence BLI Characterization (ABC) study. Endoscopy 2023; 55: 14–22 doi:10.1055/a-1852-0330

[40] Reverberi C, Rogon T, Solari A et al. Experimental evidence of effective human-AI collaboration in medical decision making. Sci Rep 2022; 12: 14952

[41] Baumer S, Kilian S, Alqahtani SA et al. Accuracy of polyp characterization by artificial intelligence and endoscopists: a prospective, non-randomized study in a tertiary endoscopy center. Endosc Int Open 2023; 11: E818–E828