Artificial intelligence in the detection of colonic polyps: what the studies say

The impact that colonoscopy and polypectomy of adenomatous polyps have on reducing the incidence and mortality of colorectal cancer (CRC) is widely known; however, it appears that endoscopic advances have not translated into reduced interval cancer rates (1, 2).

Adenoma detection rate (ADR) is a colonoscopy performance quality indicator defined as the proportion of patients in which one or more adenomas are detected during the procedure, and where a >25 % overall detection rate (≥ 30 % in men and ≥ 20 % in women) during screening colonoscopies is considered acceptable (3). This indicator allows assessing the efficacy of an endoscopist in detecting adenomatous lesions and it is of utmost importance, since for each 1 % increase in the endoscopist’s ADR there will be a 3 % decrease in the risk of CRC (4). Unfortunately, several studies have shown that ADR greatly varies among endoscopists, with ranges between 7 % and 53 % (4), and 20 % to 30 % adenoma miss rate during screening colonoscopies (5), being the cause of up to 9 % cases of interval CRC (6).

There are multiple causal factors: some inherent to the patient, such as inadequate bowel preparation for colonoscopy or lesions located in areas where evaluation is difficult, and other inherent to the endoscopist, whose inter-operator variability may be related to lack of experience or recognition errors, performing procedures with suboptimal quality indicators, and other factors such as lack of motivation, distraction, or decreased alertness during the colonoscopy (7). Therefore, strategies to minimize the rate of undetected lesions during colonoscopy are of great importance, including medical training for endoscopists and nurses, the performance of high-quality colonoscopies, the use of high-definition equipment, the use of advanced endoscopic imaging technologies such as narrowband imaging (NBI) or linked color imaging (LCI), as well as auxiliary devices such as cap or endocoff, which in some studies have been shown to improve the detection rate of polyps and adenomas (8-12).

Similarly, the fact that a large part of the diagnosis of gastrointestinal conditions is based on endoscopic procedures that use digital images makes computer-aided diagnosis (CAD) and artificial intelligence (AI) tools of interest for this purpose, as they can reduce unintentional errors taking place during colonoscopy, such as polyp detection (computer-aided polyp detection - CADe) and/or diagnosis (computer-aided detection diagnosis - CADx) (13). On the one hand, CADe has the potential to reduce the undetected polyps rate, which contributes to improving the detection of adenomas. On the other, CADx focuses on improving the accuracy of the optical diagnosis of colorec-
tal polyps, which could translate into a reduced number of unnecessary polypectomies and unnecessary resection of distal non-neoplastic polyps (13).

These AI systems have been developed using an automated and supervised learning model that extracts covariates from training data using images and manages to recognize or classify patterns that are then used for predicting results (14). Among the several existing models, deep learning (DL), which is based on convolutional neural networks (CNN), has shown an outstanding performance in terms of the analysis of these images. These networks are made up of multiple layers with artificial neurons that act as a filter for the extraction of important characteristics of the image (14). There are connection layers that can speed up the underlying calculation and, once all layers are connected, combine the characteristics of the image to create a model that allows classifying different results with the least degree of error (15) without requiring human intervention or indication.

Initial studies were either exploratory or retrospective and were performed using stored images and probably overestimated the results obtained with such data. In the Urban study, and based on a collection of 20 videos lasting 5 hours, a CNN was designed and trained for the detection of polyps with a diagnostic accuracy of 96.4% and an area under the curve (AUC) of 0.991. Furthermore, the total number of polyps detected by the AI system was significantly higher than the number reported by expert endoscopists (by 37%), although this difference was caused by small 1-6 mm lesions (16). Wang conducted another prospective study in more than 1000 patients (17) and randomized them into two groups (colonoscopy with or without AI assistance) in order to assess the ADR and the mean number of adenomas per patient, increasing from 20.3% to 29.1% and from 0.31 to 0.53, respectively. Similar to the Urban study, in the study by Wang, the increase was caused by very small polyps, and there was no a statistical difference in the detection of advanced adenomas.

Gómez et al., in their paper entitled “Artificial intelligence techniques for the automatic detection of colorectal polyps”, published in this issue of the Revista Colombiana de Gastroenterología, illustrate extensively the development of the architecture of a deep-learning CNN. Based on 3 validated CNNs (InceptionV3, Vgg16 and ResNet50), and using data collected from 6 image databases with a total of 1875 colorectal polyps, they obtained a diagnostic accuracy index for the detection of polyps of 0.7, with a sensitivity, specificity, and AUC of 0.89, 0.71, and 0.87, respectively. When compared with experienced endoscopists, the AI system showed a better sensitivity, and the possibility of being used as a second reader in a colonoscopy service was proposed, as it could correct human errors during a colonoscopy in real time. Also, Gómez et al. highlight the role of AI in bridging the gap between expert endoscopists and those still in training in terms of both detection rate and adenoma characterization.

Although the designs and algorithms of these studies are not uniform, AI has shown its ability to perform a histological differentiation, obtaining better results in studies where a deep learning model was used, as reported by Byrne’s group, who, using video endoscopic images of colorectal polyps, managed to differentiate a hyperplastic polyp from an adenomatous polyp with a sensitivity of 98% and a specificity of 83% (18). In the study by Chen, increased colonoscopy and NBI were used and a sensitivity of 96% and a specificity of 78.2% were reproduced in terms of histological differentiation (19).

Several prospective studies evaluating the role of the Cade system with CNN-based deep learning and with visual alerts allowing the real-time identification of polypoid lesions on the endoscopy monitor have been published in recent years. Wang, in a study conducted in 2020 in 369 patients alternately that were assigned to CADe colonoscopy and conventional colonoscopy groups, found that the adenoma miss rate was significantly lower in the AI assisted colonoscopy group compared to the conventional colonoscopy group (13.89 vs. 40%). Similarly, the polyp miss rate was significantly lower in the CADe colonoscopy group (12.98% vs. 45.90%) (20). Also in 2020, Repici conducted a multicenter randomized controlled trial (RCT) to evaluate the safety and efficacy of a real-time CADE system, finding that the ADR was significantly higher in the CADe group compared to the conventional colonoscopy group (54.8% vs. 40.4%). Adenomas ≤5 mm were detected in 33.7% of the cases in the CADe group compared to 26.5% in the control group, and a similar situation was observed with 6 to 9 mm adenomas (10.6% vs. 5.8%) (21).

Gong evaluated 704 patients with a deep CNN system (ENDOANGEL) to monitor the ADR, which was significantly higher in the AI group in comparison with the control group (16% vs. 8%) (22). Another meta-analysis that included 5 RCTs with a total of 4354 patients found that the pooled ADR was significantly higher in the CADE group (36.6%) than in the control group (25.2%) (23). Aziz and Cruda, in a systematic review that retrieved 3 RCTs (2815 patients in total), reported that the ADR was significantly better in the AI assisted colonoscopy group (32.9% vs. 20.8% in the control group), with a polyp detection rate (PDR) of 43.0% vs. 27.8%, respectively (24).

The largest meta-analysis on histological prediction of polyps was recently published by Lui et al. (25). Said study included 18 studies and a total of 7680 colonic polyp images for analysis, and found that, in terms of histological prediction, AI has a pooled sensitivity of 92.3%, a pooled specificity of 89.8%, and an area under the curve (AUC) of 0.92.
In conclusion, AI systems have an exponential development in digestive endoscopy and subcategories such as deep learning and neural networks will undoubtedly make large contributions to the diagnosis of gastrointestinal diseases in such a way that eventually these technologies will be considered real-time virtual assistants during colonoscopy and will help improve both ADR and adenoma miss rate, as well as reduce the incidence of colorectal carcinoma and, probably, decrease the frequency of interval neoplasms.

REFERENCES

Artificial intelligence in the detection of colonic polyps: what the studies say