



Contents lists available at ScienceDirect

Techniques and Innovations in Gastrointestinal Endoscopy

journal homepage: www.techgiendoscopy.com/locate/tgie

Artificial intelligence for colon polyp detection: Why should we embrace this?



Robin Zachariah, MD^{a,*}, Andrew Ninh^b, William Karnes, MD^{a,b}

^a Department of Gastroenterology, University of California Irvine Medical Center, Orange, California

^b Docbot, Irvine, California

ARTICLE INFO

Article history:

Received 5 June 2019

Accepted 10 July 2019

Keywords:

Machine learning

Convolutional neural networks

Colorectal cancer prevention

Adenoma detection rate

ABSTRACT

Optimal success of colonoscopy for prevention of colorectal cancer is currently measured by adenoma detection rate (ADR), which reflects a colonoscopist's ability to identify colorectal and remove precancerous polyps. Among colonoscopists in the same health care system and shared patient population, ADR varies from 7% to 53%. For every 1% increase in ADR, risk of interval colorectal cancer is reduced by 3%–6%. Beyond attaining excellent exposure of entire mucosal surface during colonoscopy, ADR can be improved with a second observer. Computer-aided detection ("facial recognition" for polyps) has potential to improve ADR as a second observer. Several groups are working to bring this technology into the endoscopy unit. Success will require real-time implementation of an affordable system with very high accuracy and proven benefit to improve ADR and reduce miss rate of precancerous lesions. In just the past year, computer-aided detection systems that run live during colonoscopy have been shown to improve ADR using affordable off-the-shelf computers.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

Colorectal cancer (CRC) is the second leading cause of cancer deaths in the United States [1]. The vast majority of colorectal cancers start as benign precancerous polyps, such as adenomas [2]. These polyps typically have a mean dwell time of around 10 years or more prior to transforming into CRC [3]. Colonoscopy remains the gold standard for finding these precancerous polyps and is the only non-surgical intervention capable of removing them. The National Polyp Study (NPS) showed that up to 90% of colorectal cancers are preventable with polyp removal [4]. The same NPS colonoscopy cohort had a 53% decrease in CRC mortality compared to the SEER population, in whom an unknown percentage had colonoscopy [2].

Adenomas are the most common type of precancerous polyp [5]. The adenoma detection rate (ADR) represents the percent of colonoscopies in which at least one adenoma is found. Ideally, ADR should equal adenoma prevalence, estimated to be greater than 50% among the screening age population [6,7]. Unfortunately, ADRs vary widely, with some colonoscopists having ADRs as low as 7% [8]. Furthermore,

in tandem colonoscopy studies, up to 24% of adenomas were missed [6]. Missed adenoma rate is reflected, in part, by the gap between ADR and adenoma prevalence. Two large studies showed that for each 1% increase in ADR, the interval colorectal cancer rate was decreased by 3%–6% [8,9]. Kaminski et al further showed that colonoscopists can be trained to improve ADR and achieve lower interval colorectal cancer rates. These data emphasize the importance of "leaving no polyp behind" to achieve the potential for 90%+ prevention of colorectal cancer with colonoscopy. Accordingly, ADR has become a key quality measure reportable to the Centers for Medicare and Medicaid and affects reimbursement under the Medicare Access and CHIP Reauthorization Act of 2015 and Merit Based Incentive Payments System since 2017 [10].

A colonoscopist's ADR is positively correlated with level of training, time spent during withdrawal, inspection technique, and bowel preparation quality [11]. Higher ADR can be achieved with special training [12,13], second look [14], Narrow Band Imaging [15], trained second observers [16], and knowledge that the colonoscopy is being recorded [17]. Scope attachments and modifications that expose mucosal surfaces behind folds and stabilize views such as Endocuff and Amplifeye [18] are also associated with increased ADR. Interestingly, multicamera systems designed to expose more of the mucosal surface on multiple viewing screens have failed to consistently improve ADR. These latter results suggest that the small angles of human central/macular vision (5% and 18% of entire visual field, respectively) may neutralize the benefit of technologies that improve

Abbreviations: AI, artificial intelligence; ADR, adenoma detection rate; CADE, computer-aided detection; CNN, convolutional neural network; CRC, colorectal cancer; GPU, graphics processing unit; NPS, National Polyp Study

Conflict of Interest: Dr William Karnes and Andrew Ninh are co-founders of Docbot, Inc., a company responsible for CNN development and implementation.

*Corresponding author.

E-mail address: rzachar1@uci.edu (R. Zachariah).

<https://doi.org/10.1016/j.tgie.2019.150631>

1096-2883/© 2019 Elsevier Inc. All rights reserved.

mucosal surface area exposure if displayed on multiple or large screens during colonoscopy. Indeed, gaze pattern has been shown to affect ADR [19]. Hawthorne effect as well as human variability in gaze pattern and limitations in central/macular vision may explain improved ADR with second observers.

2. Computer-aided detection

Computer-aided detection (CADE) of polyps has the potential to function as a second observer and reduce the miss-rate of polyps. When utilized in conjunction with quality colonoscopy practices, CADE has the potential to close the gap between ADR and adenoma prevalence and therefore reduce interval colorectal cancer rates. To achieve these lofty goals, a CADE technology must be real time (<10 ms latency), easy to implement, reliable, provide near 100% sensitivity, and a nondistracting low false positive rate. Widespread use and demand for this technology will depend on proven ability to increase ADR, clearance by regulatory agencies, and financial benefit for users. Beyond the potential for increased reimbursement through merit-based incentive systems, a reimbursement code would seal the deal for many potential users.

3. The past, present, and future

CADE of polyps has a surprisingly long history. In 2003, Karkanis et al described computer-assisted polyp detection software that used color and texture analysis to identify polyps [20]. Their system, Colorectal Lesion Detector, had an accuracy greater than 95% for polyp identification but was applicable only on static images due to high latency. Later studies evaluated shape [21,22], spatiotemporal features [23], and edge features for polyp detection [24]. For example, based on just 25 unique polyp images, Tajbakhsh et al created a CADE system using a hybrid shape analysis and achieved 88% sensitivity and 300 ms latency [21], still much too long for real-time feedback on a video stream.

These early CADE systems relied on brute force programming and early machine learning techniques, in which human programmers inform the computer about unique polyp features. Inherently, these techniques are resource intensive, time consuming, and impose human bias and error onto the software. A viable future of CADE for polyps had to wait until this decade when deep learning models using convolutional neural networks (CNNs) became available. CNNs take on the task of discovering polyp-specific features, independently of human input, learning much like a human learns to recognize a face. Now, with increasingly powerful and affordable graphics processing units (GPUs), we are at the cusp of developing and running highly accurate CNNs for polyp detection with latencies less than 10 ms, easily capable of real-time analysis during colonoscopies running at 60 frames per second.

In 2016, Li et al were one of the first to describe a deep learning system used for polyp detection based on images. However, their system was only able to achieve an accuracy of 86%, sensitivity of 73% [25]. In 2017, Wang et al presented their CNN, developed using SegNet Architecture, trained on greater than 5000 annotated images and validated on >27,000 images [26,27]; however, their described system had a latency of 77 ms operating at 25 frames per second. In 2018, Misawa et al described their version of a deep learning algorithm with a false positive rate of 60% [28].

In 2018, Urban et al described the development of a CNN algorithm theoretically capable of operating in real time with a latency of 10.2 ms (98 frames per second), achieving an accuracy of 96%, sensitivity of either 96.9% (assuming a 5% false positive rate) or 88.1% (assuming a 1% false positive rate) [29]. Their algorithm was further validated on over 5 hours of colonoscopy videos. The algorithm missed no polyps found by expert reviewers and had a false positive rate of 7%. When videos were reviewed with overlaid CNN, expert reviewers found 20% additional polyps. While flawed as a retrospective single-center video

validation study, the high performance achieved by their CNN demonstrated the potential of this new technology.

In 2019, Wang et al published results of the first randomized trial of CADE for polyp detection in a single center of 1058 colonoscopies in China [30]. Their CNN achieves a per-image sensitivity of 94.4% and per-image specificity of 95.9% with a latency of 77 ms (25 frames per second). Suboptimal latency necessitated use of 2 viewing screens, one showing the native colonoscopy video stream viewed by the colonoscopist, and the other showing the AI-processed video stream. An alarm system was developed to notify the colonoscopist when the AI predicted the presence of a polyp. The artificial intelligence (AI) system significantly increased ADR (29.1% vs 20.3%, $P < 0.001$) and the mean number of adenomas per patient (0.53 vs 0.31, $P < 0.001$). This study is the first to demonstrate the utility of CADE live during colonoscopy, and the first to show that its use improves ADR.

Progress is rapid toward a viable CADE system that achieves accurate polyp detection in real time on one screen. This has likely already been achieved but unpublished. Current CADE CNNs for polyp detection already demonstrate performance characteristics that are at the cusp of live implementation after regulatory clearance. But what of the other AIs being developed for colonoscopy and other endoscopic procedures. How will these AI “apps” play together in a single hardware/software solution? Can they even be run simultaneously with current hardware/GPU limitations in an affordable and compact form-factor? There is no doubt that they will, perhaps sooner than we expect!

4. Conclusion: Embracing AI

New technologies designed to reduce colorectal cancer deaths often strike fear and skepticism in colonoscopists. Will novel stool and blood tests, virtual colonoscopy, colon pill cameras and x-ray devices, or polyp chemopreventive drugs put my career at risk? None of these technologies have reduced the demand for expert colonoscopy for definitive screening, diagnosis, and intervention. In many cases, these technologies have increased demand by engaging more of the population in screening.

We are at the cusp of reaping benefit from AI for colonoscopy. A CADE that behaves as an expert second observer and improves our detection of polyps will reduce the risk of interval colorectal cancers. However, even the most accurate CADE system cannot assist detection of polyps that are not exposed, it cannot compensate for inadequate cleaning, low cecal intubation rate, fast withdrawal time, and poor inspection technique. Colonoscopists remain responsible for the care of his or her patient, for deliberately exposing all mucosal surfaces, interpreting potential pathology, and determining appropriate intervention. It will be a long time before AI and robotics can replace a skilled colonoscopist.

But what if the use of CADE desensitizes our acuity to recognize polyps? What if we ignore important lesions unrecognized by CADE? What if we become dependent on and overly trusting of CADE and remove all identified lesions, even those that require no intervention such as a suction pseudopolyp or inverted diverticulum? Future studies will be needed to determine if use of CADE systems sensitize or desensitize acuity for recognition, interpretation and appropriate intervention of suspected abnormalities, especially among those in fellowship training. In the meantime, we should embrace CADE systems only as a “second observer,” one that questions us: “what is this; is it important?”

The near future is likely to bring other benefits of AI to colonoscopy, including automated cecal intubation rate [31], withdrawal time [31], Boston bowel prep score [32], polyp size [33], polyp pathology [34], Mayo Endoscopic Score [35], etc., enabling automated endoscopic reports, standardized scoring, and quality measure reporting. For example, CADE systems are under development by several groups

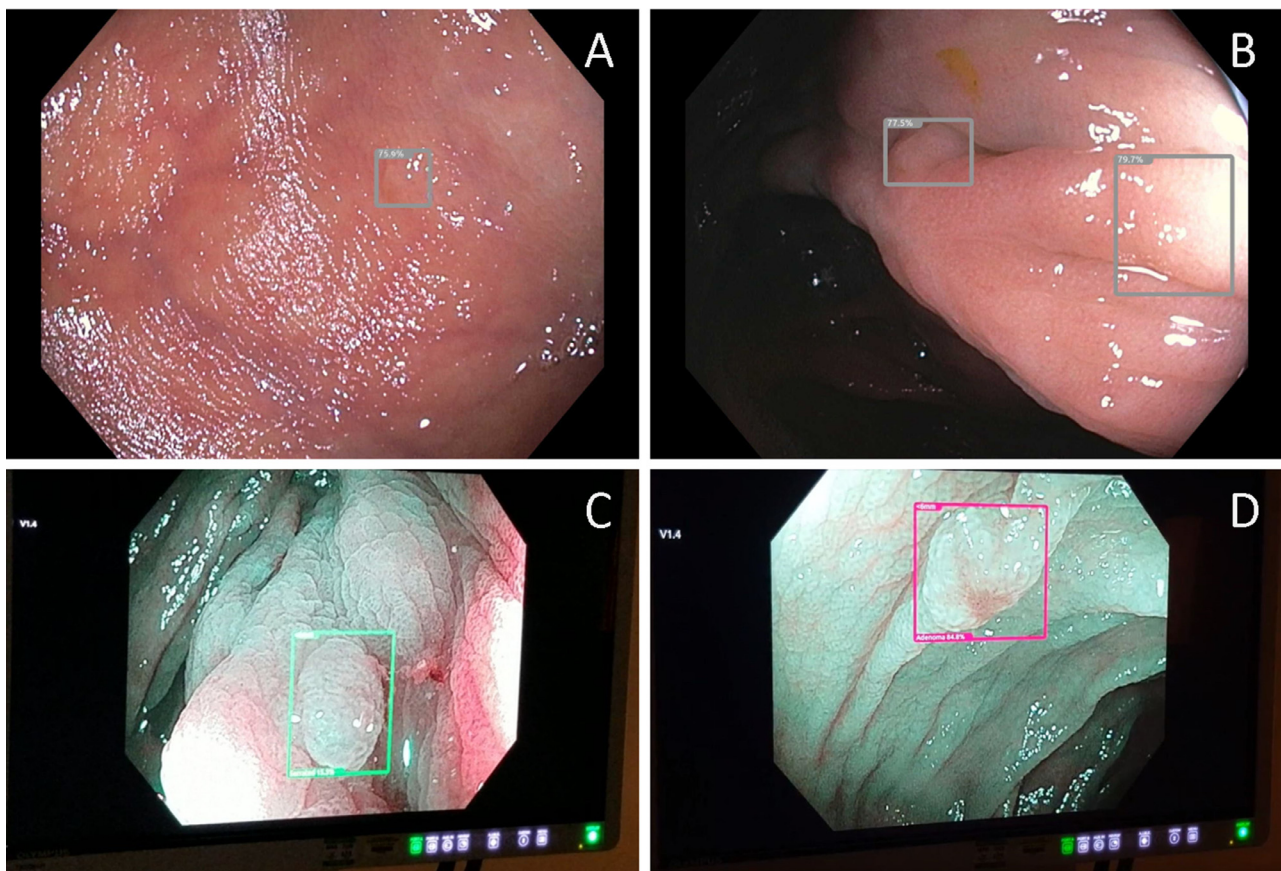


Fig. 1. Images of video frames from live colonoscopies overlaid with AIs for polyp detection (panels A and B) and optical pathology predictions (panels C and D). Polyp detection AI runs continuously while optical pathology AI runs only on “freeze frames.” Panel A shows a single predicted polyp. Panel B shows multiple predicted polyps. Panel C shows a polyp predicted to be serrated (green box). Panel D shows a polyp predicted to be adenomatous (red box). (Figure provide courtesy of Docbot, Inc.)

capable of both polyp detection and polyp optical pathology prediction (Figure 1). The future appears bright for AI-assisted improvements in the delivery of efficient, accurate, and cost-effective patient care [36].

5. Epilogue: Stardate 41254.7

One day, the skills of expert colonoscopists will be obsolete and replaced by a new technology. Imagine, for example, an ingested self-driving nanodevice, available by next day delivery from Amazon or Alibaba and taken before bedtime. As you sleep, the device accurately identifies, diagnoses, and removes all precancerous polyps—no need for colonoscopy prep, sedation, imposition on someone to drive you home, or time off work. That day may be generations ahead and may be supplanted by another technology that prevents polyps and cancers from ever forming. In the meantime, colonoscopist can and should be the best they can be and embrace new technologies that assist in our purpose: reduce the burden of colorectal cancer.

References

- [1] Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat database: Mortality-all COD, aggregated with state, total US (1969-2014) <Katrina/Rita Population Adjustment>. Bethesda, MD: National Cancer Institute; 2016 Updated September 10, 2018.
- [2] Zauber AG, Winawer SJ, O'Brien MJ, Lansdorp-Vogelaar I, Ballegooijen MV, Hankey BF, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med* 2012;366:687–96.
- [3] Kuntz KM, Lansdorp-Vogelaar I, Rutter CM, Knudsen AB, van Ballegooijen M, Savarino JE, et al. A systematic comparison of microsimulation models of colorectal cancer. *Med Decis Making* 2011;31(4):530–9.
- [4] Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med* 1993;329:1977–81.
- [5] Shinya H, Wolff W. Morphology, anatomic distribution and cancer potential of colonic polyps. *Ann Surg* 1979;190(6):679–83.
- [6] Leufkens A, van Oijen M, Vlegga F, Siersema PD. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. *Endoscopy* 2012;44(5):470–5.
- [7] El-Halabi M, Rex D, Saito A, Eckert GJ, Kahi CJ, et al. Defining adenoma detection rate benchmarks in average-risk male veterans. *Gastrointest Endosc* 2019;89(1):137–43.
- [8] Corley D, Jensen C, Marks A. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014;370(26):2539–41.
- [9] Kaminski M, Wieszczyni P, Rupinski M, Wojciechowska U, Didkowska J, Kraszewska E, et al. Increased rate of adenoma detection associates with reduced risk of colorectal cancer and death. *Gastroenterology* 2017;153(1):98–105.
- [10] GI Quality Measures for 2017 Released in MACRA Final Rule. Available at: <https://www.gastro.org/news/gi-quality-measures-for-2017-are-released-in-macra-final-rule>. [Accessed 3 May 2019].
- [11] Anderson JC, Butterly LF. Colonoscopy: quality indicators. *Clin Transl Gastroenterol* 2015;6(2):e77.
- [12] Eckardt A. Optimizing adenoma detection rates: equipment, experience or education? *Ann Gastroenterol* 2013;26(3):272–4.
- [13] Peters S, Hasan A, Jacobson N, Austin GL. Level of fellowship training increases adenoma detection rates. *Clin Gastroenterol Hepatol* 2010;8(5):439–42.
- [14] Clark BT, Parikh ND, Laine L. Yield of repeat forward-view examination of the right side of the colon in screening and surveillance colonoscopy. *Gastrointest Endosc* 2016;84(1):126–32.
- [15] Yoshida N, Inoue K, Yasuda R, Hirose R, Dohi O, Naito Y, et al. An additional 30-s observation of the right-sided colon with narrow band imaging decreases missed polyps: a pilot study. *Dig Dis Sci* 2018;63(12):3457–64.
- [16] Xu L, Zhang Y, Song H, Wang W, Zhang S, Ding X. Nurse participation in colonoscopy observation versus the colonoscopist alone for polyp and adenoma detection: a meta-analysis of randomized, controlled trials. *Gastroenterol Res Pract* 2016;2016:1–7.
- [17] Madhoun M, Tierney W. The impact of video recording colonoscopy on adenoma detection rates. *Gastrointest Endosc* 2012;75(1):127–33.
- [18] Chin M, Karnes W, Jamal M, Lee J, Lee R, Samarasena J, et al. Use of the Endocuff during routine colonoscopy examination improves adenoma detection: a meta-analysis. *World J Gastroenterol* 2016;22(43):9642–9.

- [19] Almansa C, Shahid MW, Heckman MG, Preissler S, Wallace MB. Association between visual gaze patterns and adenoma detection rate during colonoscopy: a preliminary investigation. *Am J Gastroenterol* 2011;106(6):1070–4.
- [20] Maroulis DE, Iakovidis DK, Karkanis SA, Karras DA. CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames. *Comput Methods Programs Biomed* 2003;70(2):151–66.
- [21] Tajbakhsh N, Gurudu SR, Liang J. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Trans Med Imaging* 2016;35:630–44.
- [22] Hwang S, Oh J, Tavanapong W, Wong J, de Groen P. Polyp detection in colonoscopy video using elliptical shape feature. *Proc ICIP* 2007;2:465–8.
- [23] Bernal J, Sánchez FJ, Fernández-Esparrach G, Gil D, Rodríguez C, Vilarinho F, et al. WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians. *Comput Med Imaging Graph* 2015;43:99–111.
- [24] Wang Y, Tavanapong W, Wong J, Oh JH, de Groen PC. Polyp-alert: near real-time feedback during colonoscopy. *Comput Methods Programs Biomed* 2015;120(3):164–79.
- [25] Li T, Cohen J, Craig M, Tsourides K, Mahmud N, Berzin T. The next endoscopic frontier: a novel computer vision program accurately identifies colonoscopic colorectal adenomas. *Gastrointest Endosc* 2016;83:AB482.
- [26] Wang P, Xiao X, Brown JR, Berzin T, Tu M, Xiong F, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2018;2(10):741–8.
- [27] Wang P, Xiao X, Liu J, Li L, Tu M, He J, et al. A prospective validation of deep learning for polyp auto-detection during colonoscopy. *Am J Gastroenterol* 2017;112:S106–10.
- [28] Misawa M, Kudo SE, Mori Y, Cho T, Kataoka S, Yamauchi A, et al. Artificial intelligence-assisted polyp detection for colonoscopy: initial experience. *Gastroenterology* 2018;154(8):2027–9.e3.
- [29] Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 2018;155(4):1069–78.e8.
- [30] Wang P, Berzin T, Glissen Brown J, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019. <https://doi.org/10.1136/gutjnl-2018-317500>.
- [31] Karnes WE, Ninh A, Dao T, Requa J, Samarasena J. Real-time identification of anatomic landmarks during colonoscopy using deep learning. *Gastrointest Endosc* 2018;87(6). <https://doi.org/10.1016/j.gie.2018.04.447>. (abstr Sa1925).
- [32] Karnes WE, Ninh A, Dao T, Requa J, Samarasena J. Unambiguous real-time scoring of bowel preparation using artificial intelligence. *Gastrointest Endosc* 2018;87(6). <https://doi.org/10.1016/j.gie.2018.04.461>. (abstr Sa1940).
- [33] Requa J, Dao T, Ninh A, Karnes W. Can a convolutional neural network solve the polyp size dilemma? *Am J Gastroenterol* 2018;113:S158.
- [34] Zachariah R, Ninh A, Dao T, Requa J, Karnes W. Can artificial intelligence (AI) achieve real-time “resect and discard” thresholds independently of device or operator? In: Presented at the American College of Gastroenterology Annual Scientific Meeting, Wolters Kluwer, Philadelphia; 2018.
- [35] Abadir A, Requa J, Ninh A, Karnes W, Mattar M. Unambiguous real-time endoscopic scoring of ulcerative colitis using a convolutional network. *Am J Gastroenterol* 2018;113:S349.
- [36] Abu Dayyeh BK, Thosani N, Konda V, Wallace MB, Rex DK, Chauhan SS, et al. ASGE Technology Committee. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE PIV1 thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc* 2015;81:502.e1-16.