ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 March 12; 2(1): 1-11

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 1 March 12, 2021

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 1 March 12, 2021

Contents

ABOUT COVER

Editorial Board Member, Giovanni Di Nardo, MD, PhD, Assistant Professor, Pediatric Unit - Pediatric Gastroenterology Section, NESMOS Department, Sapienza University of Rome, Rome 00189, Italy. giovanni.dinardo@icloud.com

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Fatih Altintoprak, Sahin Coban, Krish Ragunath

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
March 12, 2021

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 March 12; 2(1): 1-11

DOI: 10.37126/aige.v2.i1.1

ISSN 2689-7164 (online)

MINIREVIEWS

5G mobile communication applications for surgery: An overview of the latest literature

Leandra Börner Valdez, Rabi R Datta, Benjamin Babic, Dolores T Müller, Christiane J Bruns, Hans F Fuchs

ORCID number: Leandra Börner Valdez 0000-0002-6998-2255; Rabi R Datta 0000-0002-6829-7379; Benjamin Babic 0000-0003-0313-6862; Dolores T Müller 0000-0002-1833-5664; Christiane J Bruns 0000-0001-6590-8181; Hans F Fuchs 0000-0003-4764-8050.

Author contributions: Börner Valdez L, Datta RR, Babic B, Müller DT, Bruns CJ, Fuchs HF wrote and revised the manuscript.

Conflict-of-interest statement: The authors declare no conflict of interests in relation to this article.

Manuscript source: Invited manuscript

Specialty type: Surgery

Leandra Börner Valdez, Rabi R Datta, Benjamin Babic, Dolores T Müller, Christiane J Bruns, Hans F Fuchs, Department of Surgery, University of Cologne, Cologne 50937, Germany

Corresponding author: Hans F Fuchs, MD, Assistant Professor, Surgeon, Department of Surgery, University of Cologne, Kerpener Str. 62, Cologne 50937, Germany. hans.fuchs@uk-koeln.de

Abstract

Fifth-generation wireless network, 5G, is expected to bring surgery to a next level. Remote surgery and telementoring could be enabled and be brought into routine medical care due to 5G characteristics, such as extreme high bandwidth, ultra-short latency, multiconnectivity, high mobility, high availability, and high reliability. This work explores the benefits, applications and demands of 5G for surgery. Therefore, the development of previous surgical procedures from using older networks to 5G is outlined. The current state of 5G in surgical research studies is discussed, as well as future aspects and requirements of 5G in surgery are presented.

Key Words: 5G; Wireless networks; Remote surgery; Telesurgery; Telementoring; Robotic surgery

Core Tip: Very few research studies have been conducted to prove efficacy and feasibility of 5G in surgery so far, with most of these studies being case studies. All of them reported a stable 5G network proving 5G to be feasible for surgery. However, detailed information about the data rate and latency are missing. More research efforts are demanded to explore questions like the combination with new technologies, *e.g.*, Virtual Reality, political regulations, or cyber-security if 5G becomes the backbone of next-generation surgery.

Citation: Börner Valdez L, Datta RR, Babic B, Müller DT, Bruns CJ, Fuchs HF. 5G mobile communication applications for surgery: An overview of the latest literature. *Artif Intell Gastrointest Endosc* 2021; 2(1): 1-11

URL: https://www.wjgnet.com/2689-7164/full/v2/i1/1.htm

Country/Territory of origin: Germany

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): D
Grade E (Poor): 0

Received: January 23, 2021
Peer-review started: January 23, 2021
First decision: February 28, 2021
Revised: March 10, 2021
Accepted: March 12, 2021
Article in press: March 12, 2021
Published online: March 12, 2021

P-Reviewer: Ielpo B, Jiang T
S-Editor: Wang JL
L-Editor: Filipodia
P-Editor: Wang LL

DOI: https://dx.doi.org/10.37126/aige.v2.i1.1

INTRODUCTION

The use of the Internet in the medical field emerged over time. With the current global situation dealing with coronavirus disease 2019, it has become clear that the healthcare system is dependent on the Internet more than ever. The Internet presents a base to solve existing problems in our health care system. First, medicine being impersonal instead of individual. Second, healthcare being provider-centered instead of invalid-centered. Finally, medical treatment being unevenly available instead of accessible to any ethnicity, income, and geographic location[1].

Looking at the problems in the field of surgery, it has been reported that surgical procedures are not provided worldwide due to a lack of trained professionals[2,3], and besides the invention of laparoscopic surgery most surgeries are technically still performed as they were a hundred years ago[4]. In 1994, Simon[5] already forecasted that the future of surgery in the 21st century will be characterized by an increased distance between the patient and the surgeon.

One step in this trend was the introduction of robotic surgery. Though, it has not passed the threshold to standard use in clinic yet. To play a role in surgery where high precision and reliability are an unquestionable requirement, the evolution of the Internet to 5G - 5th generation mobile network - seems to be the promising and necessary surgical tool in the operating room that will ultimately avail the existing innovations of surgical technologies.

This review aims to clarify the potential benefits and applications of 5G in surgery. Therefore, a brief historical overview and the development of previous surgical procedures using the Internet are provided. Next, 5G and its characteristics are described, followed by a current state summarization of 5G in surgical research studies. Finally, future aspects and requirements of 5G in surgery are outlined.

LITERATURE REVIEW

A literature review was conducted by two researchers (Börner Valdez L, Fuchs HF) individually. The search terms of "5G", "5G network" or "wireless network" paired with "surgery", "remote surgery" or "telesurgery" applied to PubMed yielded a total of 6538 search results. An additional Internet search and reviewed references yielded an additional 20 records. After screening the abstracts for eligibility and removing duplicates, 137 articles were left. Research studies, case reports, reviews or book chapters in English were included for the purpose of this study. After full-text articles were read, 53 articles remained for this study.

INTERNET – BACKBONE FOR MODERN SURGERY

History of the Internet in surgery

Two studies are frequently cited in the related research and constitute noteworthy enablers for a next step in modern surgery. One is the "Lindbergh Operation" conducted by Marescaux *et al*[6] in 2001. Marescaux *et al*[6] performed a safe and uneventful cholecystectomy on a woman in Strasbourg, France, using the ZEUS robotic system. The extraordinary feature of this case was that the procedure was performed from New York City, United States, 6.000 km away from the patient's bedside. To do so, a transatlantic fiberoptic connection (referred to as ATM) was established. This connection enabled the use of a guaranteed bandwidth of 10 Mbit/s for both the robot motion and video data. A latency of 155 ms was measured, thereby ranging under the previously stated safe latency threshold of 300 ms[7,8]. The costs for the medical and technical staff, the robot and the Internet connection exceeded 1 million US dollars[6].

In 2003, Anvari[9] aimed to show that remote surgery is advantageous, especially for countries with a greater portion of rural areas. Therefore, 22 abdominal surgeries were performed remotely between two locations in Canada, 400 km away from each other, using the ZEUS robotic system without any intra- or postoperative complications. Compared to the Marescaux's operation, a commercial Internet Protocol/Virtual

Private Network (*i.e.* IP/VPN) was used at a bandwidth of 15 Mbit/s. The latency was comparable, between 135 and 140 ms[10]. The costs reached 2.5 million US dollars, including the costs for exploring the Internet requirements and solutions for this operation[11].

Both studies were conducted successfully and uneventfully, proving that the technical requirements to separate the surgeon from the patient side already existed 20 years ago. These studies also show that, in remote surgeries, the Internet plays a fundamental role in the operating room, even if some fine touch adjustments were needed in matters of bandwidth, latency, and costs at that time.

Advantages of the Internet in surgery

Taking the idea of using the Internet in surgery one step further, possibilities open up which are not new as an idea but have not been fully reached nor transferred into daily practice yet. As mentioned before, remote surgery could serve areas which are geographically difficult accessible, or by simply saving long-distance travel and costs. Hence, the expertise of specialized surgeons could still be experienced from far away[10,11].

Minimizing the transmission of infectious diseases or avoiding dangerous environments in general is more relevant than ever[12,13]. Both the patients and the medical staff could minimize these risks if remote surgery would become more applicable[14,15]. To provide surgical care in risky surroundings or battlefields has always been of interest to the military and naval sectors. Therefore, it is not surprising that a lot of research in telesurgery has been initiated by the military[16,17].

However, not only extreme situations have to be thought of in regards to how the Internet could improve surgical care. Specialized surgeons could either collaborate on a national or international level, or unexperienced surgeons could be trained and mentored through remote surgery[18,19]. Surgical education could therefore expand in both directions — in width and depth.

Problems and requirements of the Internet in surgery

Already in 1996, Smithwick[20] defined the network requirements for surgery, as follows: "reliability; an acceptable end-to-end delay; the ability to transfer data from sources with widely different data rates; low data error rate." Furthermore, availability needs to be added to this listing, especially if remote surgery or telementoring shall be utilized at any time, from any place[14,16].

Previous studies exposed the flaws and examined the necessary requirements for networks in surgery, mainly bandwidth and latency. The bandwidth describes the possible data volume transmitted per time unit[21]. Rayman *et al*[22] explored the minimum bandwidth for safe remote surgery and found that a bandwidth above 5 Mbit/s should be achieved, whereas Anvari[11] stated that 7 Mbit/s would be needed for remote surgery. Bandwidths used in previous studies for remote surgery - mainly for the video and the robotic signals – have ranged from 10[6] and 15[10] to 23[23,24] and 40[25] Mbit/s.

Latency describes the time amount for data to be transmitted from the sending source to the receiver[21]. For remote surgery, latency must be kept as low as possible. The higher the latency, the more surgical performance deteriorates[8,25-27], even though adaption occurs[28,29]. Acceptable latency in remote surgery has been determined as below 300 ms, in various studies[6,9,30-32]. Both ideal bandwidth and latency are crucial determinants for the Internet to be sufficient in modern surgery.

5G - FIFTH-GENERATION WIRELESS NETWORKS

Properties of 5G

Networks used for surgery evolved over time from satellite[22,24,28,29], Integrated Services Digital Network[8,33-35], ATM[6,7] and IP/VPN[10,27] to the current wireless networks[36]. 5G could meet the previously outlined demands for surgery which were not reached by its predecessor, the 4G/Long-term evolution (LTE) mobile communication standard[1,37,38]. To have a better understanding of the improvements compared to LTE, the properties of 5G are described herein (Figure 1, Table 1).

5G is characterized by its extremely high data rate, up to 10 Gbit/s, thus being 100 times faster than LTE[39-42]. The high data transmission is explained by high frequencies, up to 30 GHz[40,43]. However, the high frequencies of 5G explain two disadvantages compared to LTE. With higher frequencies, wavelengths become smaller and therefore have worse penetration of objects. Consequently, LTE is less susceptible to blockage by

Table 1 Comparison between long-term evolution and 5G[40]

Characteristic	4G/LTE	5G
Data rate	0.01-1 Gbit/s	0.1-20 Gbit/s
Latency	10 ms	1 ms
Mobility	~ 360 km/h	~ 500 km/h
Energy efficiency	0.1 mJ per 100 bits	0.1 μJ per 100 bits

LTE: Long-term evolution.

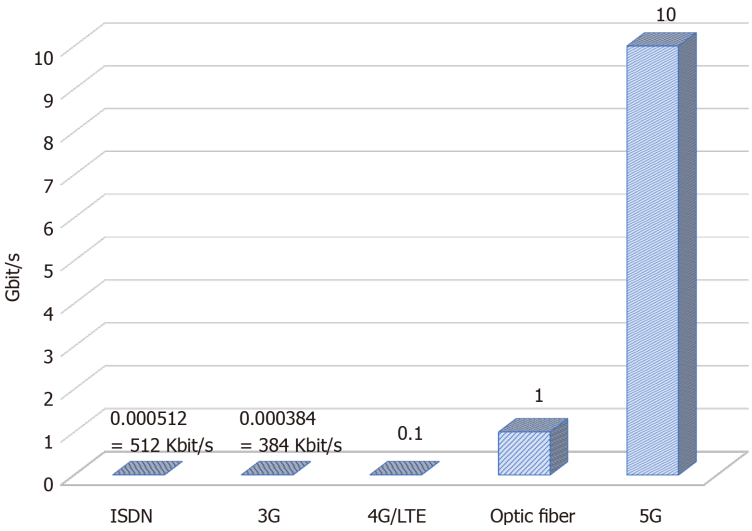

Figure 1 Data rate of distinct networks in Gbit/s[59]. ISDN: Integrated services digital network; LTE: Long-term evolution.

objects in a room than 5G[44,45]. Second, 60 times less distance can be overcome by the high bands of 5G[44].

The latency of 5G is about 1 -2 ms, which represents a 10-fold decrease in latency of LTE[38-42]. 5G is less energy-consuming than LTE (0.1 μJ per 100 bits *vs* 0.1 mJ per 100 bits)[39,41]. With 5G, simultaneous mobile use and device connectivity (massive multiple-input multiple-output; *i.e.* MMIMO) increase by a 100-fold compared to 4G/LTE. Further key features of 5G are high availability and stable reliability[41,42].

Benefits and applications of 5G in surgery

Knowledge of the 5G characteristics makes its use as a platform for new-generation surgery seem possible.

As outlined before, key requirements for remote surgery are a high bandwidth with a fast data transfer and without delays. 5G mobile networks meet these requirements. With a data rate in the gigabit range, previously stated necessary bandwidths of 7 Mbit/s[11] are easily met. The same applies to the necessary latency for remote surgery, which has been defined as below 300 ms[6,9,30-32]. 5G offers a 1 ms latency and therefore represents a huge improvement for this crucial aspect in surgery.

With its high speed, low latency and wireless transfer, multi-connectivity between multiple devices or users is possible[41,42]. This opens the door for real-time telementoring with the option to participate and interfere from a distant location. This does not only mean a gain in surgical quality. It can mean a cost and time reduction in the microcosmos of a hospital under increasing economic pressure of the health system, *e.g.*, a surgeon can operate from their office, with less staff in the operating theater and saving of materials. Globally, long-distance travel can be decreased, even though real-time exchange would still be possible.

Not only videos or pictures could be transferred in high definition (*e.g.*, 4K and 8K)[38]. The sensual experience in robotic surgery could be extended in matters of tactile sensation. Not only can the executive device adapt the movements of the surgeon but also the surgeon can experience haptic feedback when the patient is connected to a sensing device[46,47]. This equates to a great data load, which could be transmitted

through 5G. Especially in surgery, the sense for tissue is of great importance. The transfer of this information represents an innovative possibility.

Robotic surgery could furthermore be combined with virtual reality (referred to as VR) and augmented reality (referred to as AR) technology. With 5G, this technology can expand and be refined for surgery. It could be used for surgical education as well as for safer and more accurate surgery[38,41]. Collected data during robotic surgery could be processed for either machine learning (referred to as ML) or artificial intelligence (referred to as AI)[1]. Furthermore, it could be shared and transferred at another speed and size for research purposes[41].

With 5G, it is expected that the Internet of Things (referred to as IoT) technology will become possible. IoT means that any physical object and its use could be connected to each other virtually, creating a whole new data cloud[45]. At best, these data will facilitate processes in daily life. An already existing IoT-technology in the medical sector is the monitoring of blood sugar levels of patients with diabetes *via* an integrated sensing device and smartphones. This technology could be extended, meaning that other human measurable parameters could be directly transmitted from an integrated sensing device to a monitor wirelessly and remotely[1].

In surgery, these health devices could monitor patients before and after the operation to filter and prevent complications early or simply to come to an early diagnosis. Monitoring and extracting medical information through these devices could offer a new perspective about diseases on an individual basis and therefore bring about new aspects of treatment[1,47].

CURRENT STATUS OF RESEARCH: 5G IN SURGERY

Some of these possible features and applications still sound far away. However, the global implementation of 5G has already begun. Obviously, experiments with only machines or devices are easier to conduct and have less risks. Therefore, it is not surprising that medicine lags with the incorporation of new technologies. Very few clinical studies have been conducted so far to prove the benefit of 5G for surgery, most likely due to the lack of a standard 5G network and the expense to establish a technical setup.

It is of note that the first telesurgical procedures with 5G network were reported from China. One was a porcine liver resection and the other was a human brain surgery in 2019. Both procedures were supported by Huawei, the biggest network provider of 5G[48,49]. Although, it is hard to find any detailed information about these two procedures outside of the Huawei online page and Chinese media. No related scientific research papers have been published for these two cases, to the best of our knowledge. Therefore, only published research papers on 5G in surgery will be considered herein (Table 2).

Jell *et al*[37] evaluated whether 5G technology is suitable for surgery, in 2019. The study design contained a theoretical and a technical part. First, a structured questionnaire (Delphi study) was conducted to explore the benefits and demands for surgery by 12 professionals; nine of the participants were from the industry. Second, two case studies examined the technical feasibility and parameters of 5G in the operation room. Therefore, tracking and tracing of static (n = 4) and moving (n = 4) objects was investigated in the operating theater by 5G locators. The second case study simulated the remote robotic camera control (SoloAssist; AKTORmed GmbH) in a phantom, by both inexperienced and advanced medical staff (n = 15). Only camera positioning was performed without any surgical tasks. The Delphi study showed that the majority agreed in the general, useful capability of 5G for the health care sector but more research efforts, especially in daily clinical routine, and global standards are requested. More than a half of the respondents expected benefits for rural areas through telemedicine and advocated for state-funded support. The participants were neutral about industrial funding but recommended early participation of the industry for the realization of 5G in medicine. No statement about the additional effort and justified costs to establish 5G in hospitals was possible. The tracking and tracing of static and moving objects in the operating room was possible; though, the tracking of objects revealed some inaccuracies. The robotic test case showed data rates of the video and robotic signals together around 8 Mbit/s. The latency ranged between 2-60 ms, with most data (75%) being transferred after 30 ms. No disruption of the network was noted.

Lacy *et al*[50] performed the first telementored surgery using 5G network to prove feasibility and benefits. Two procedures (laparoscopic high or low anterior resection)

Table 2 Research studies using 5G for surgery

Ref.	Yr	Study type	Robot system	Location - remote distance	Data rate	Latency
Jell *et al*[37]	2019	(1) Delphi study; and (2) 2 use cases: (a) Track and tracing; and (b) Remote robotic camera control on phantom	SoloAssist (AKTORmed)	Munich (Germany): Not indicated	~ 8 Mbit/s	2-60 ms
Lacy *et al*[50]	2019	Telementoring with telestration: (1) Laparoscopic ant. high resection; and (2) Laparoscopic ant. low resection	-	(1) Barcelona (Spain): 4 km; and (2) Shanghai (China): 6 km	(1) 95-102 Mbit/s; and (2) 99-106 Mbit/s	(1) 202 ms; and (2) 146 ms
Acemoglu *et al*[51]	2010	Brief research report: Remote microsurgical cordectomy on a cadaver	Panda robot (Franka Emika); 3D-microscop (Karl Storz)	Milan (Italy): 15 km	-	One-way video latency: 102 ± 2 ms (max. ≤ 140 ms); Round-trip latency: 280 ms
Tian *et al*[52]	2020	Remote spinal surgery on human (n = 12)	TiRobot system (Tinavi)	Beijing (China): 120 km; 280 km; 750 km; 1200 km; 3000 km	-	28 ms
Zheng *et al*[53]	2020	Remote laparoscopic surgery on porcine model (n = 4)	MicroHand (WEGO Group)	Quingdolo (China): 3000 km	-	Mean round trip latency: 264 ms (258-278 ms)

were safely performed in Spain and China. The distance between the mentoring surgeon and the operation side was 4 and 6.2 km respectively. The remote surgeon carried out his mentoring function through verbal communication and telestration, so that the surgeon to drew on the laparoscopic image which was received by the operating team at the same time. Both surgeries were performed without signal interruption. Image and transmission quality were highly rated by both surgeons. A data rate of 98 and 101 Mbit/s and a latency of 202 and 146 ms were recorded respectively.

In 2020, Acemoglu *et al*[51] presented a case report where a remote microsurgical cordectomy was performed effectively on a human cadaver from a distance of 15 km using 5G network. The robot used for surgery was the Panda robot (Franka Emika) and visualization was established by 3D-microscopy (Vitom 3D; Karl Storz). Neither information about the data rate nor about the latency of the robotic effector movement solely was indicated by the authors. The maximum video-transmission latency was below 140 ms. The maximum round-trip latency was 280 ms.

Tian *et al*[52] performed 12 cases of remote spinal surgeries on humans with lumbar spinal disorders using a surgical robot (TiRobot system) and 5G network. The aim was also to test the reliability and therefore applicability of 5G for surgery. The remote surgeon, located in Beijing, conducted the pedicle screw planning and robot arm positioning for five different patient sides, with distances ranging from 120 to 3000 km. The surgeons on the patient side performed the screw placement. No detailed data about the number of surgeries or technical parameters of the network at each location was provided. Additionally, the remote surgeon guided the following two operations at different locations at one time, called "one-to-many-remote-surgery": A one-to-two operation sides (Beijing to Shangdong and Zhejiang) and a one-to-three operation sides (Beijing to Xianjiang, Hebei and Tianjin). There was no network disruption in any case. The mean network latency was reported to be 28 ms but was not explained further. No data was provided about the data rate. The pedicle screw implantation was rated acceptable in all cases and there were no intraoperative complications.

Zheng *et al*[53] evaluated the efficacy, availability, reliability, and safety of 5G in four remote laparoscopic surgeries on a porcine model using the MicroHand robot system (n = 4; nephrectomy, hepatectomy, cholecystectomy and cystectomy). The remote surgeon was located 3000 km away from the patient side in China, connected *via* 5G network with a bandwidth of 1 Gb/s. A wired Internet connection served as reference with a bandwidth of 100 Mbit/s. The procedures were conducted safely without any adverse effects. No network errors were noted. The mean total latency of 5G was 264 ms (258-278 ms), whereas the mean latency of the wired Internet connection was shorter, with 206 ms (204-210 ms). The parts of the total latency consisted of the mean round-trip delay, the servo period of the surgical robot (< 1 ms), the mechanical response delay of the robot (40 ms), the endoscope imaging and image processing delay (50 ms), and the video codec delay (60 ms). Mentioned times were the same in both network setups but the mean round-trip delay was 114 ms (108-124 ms) for 5G and 56 ms (54-60 ms) for the wired network setup. The data rate was not mentioned in this study.

DISCUSSION, FUTURE ASPECTS AND DEMANDS FOR 5G IN SURGERY

The described studies evaluating 5G for surgery comprise a technical report using a use case on a phantom[37], one remote telementoring study with two use cases on humans[50], one brief research report about remote microsurgery on a cadaver[51] and a remote laparoscopic surgery on four pigs[53] but only one interventional study with 12 patients undergoing remote robotic neurosurgery[52]. This *status quo* demonstrates that surgery using 5G has not passed the threshold of a first-line approach. As a matter of fact, only very few randomized studies on telesurgery and telementoring using older networks have been published so far[54,55]. Though, randomized studies must follow.

However, case studies are a good start to develop an understanding about the technical details of the network and to detect first obstacles. Jell *et al*[37] and Lacy *et al*[50] provided the first data about the particular data rate and latency of the video and robotic camera movement signals. Both studies, though, did not use 5G technology to directly perform surgical tasks. Only the work of Tian *et al*[52] included robotic movement control for spinal surgery on humans but they did not provide any data about the particular data rate. However, it would be of great interest to know how a higher data size of robotic control would be transmitted through 5G.

Only Zhen *et al*[53] provided detailed information about their results regarding latency. The measured total latency using the 5G network (264 ms), though, was longer than the reference setup using a wired connection (206 ms). When looking closer at the parts causing the latency, it is noticeable that the 5G and wired connection setup only differ in the time of the mean round trip delay. A detailed explanation for the longer mean round trip delay of the 5G network was not provided by the authors. A lack of enough 5G antennas between the patient and surgeon side might be an explanation. The authors pleaded for 5G, due to such benefits as its multimodal data transmission, higher bandwidth and being wireless, which therefore provide great mobility, even in rural areas which are difficult to access.

Lacy *et al*[50] argued that recorded latencies could have been faster in their setup if 5G would have been the connecting network of any device in the whole study setup, and not only two 5G antennas between the patient and mentor side. Furthermore, adequate image processing software for coding and decoding video signals would lead to shortened latencies. Jell *et al*[37] explained measured inaccuracies in their tracking setup due to other, intermediate objects interrupting the signals to the 5G receiver. Both examples show that further software innovations for the use of 5G is demanded. Overall, published data using 5G in surgery lacks technical details and needs to provide more information about the composition of data rate and latency for further improvements.

None of the cited studies used the Da Vinci robot system (Intuitive Surgical Inc.). However, the Da Vinci surgical system is the most used surgical robot worldwide[56]. To evaluate whether 5G works for remote robotic surgery in a broad spectrum of surgical fields, research efforts need assess the technical feasibility with the Da Vinci system. Furthermore, new technologies such as AI, ML, VR or AR that are being claimed to contribute to surgery have not been studied with 5G and surgery scientifically, to our knowledge.

All authors of the studies using 5G for surgery agreed in the benefit of telesurgery for remote areas. This is in accordance with previous literature as outlined before. However, none of the described studies mentioned the costs to establish 5G-based surgery. Van Wynsberghe *et al*[15] justifiably raised the objection that aforementioned, remote areas - usually rural and low-income areas – will not have the means to afford highly advanced technology.

It is estimated that the network expansion of 5G in a country of size and economic status similar to Germany's would cost tens of billions of Euros[57]. Therefore, a depiction of the costs establishing such a technical system like robotic surgery with 5G is important. It could elucidate to what extent financial support or funding by the government or a Union of States is needed. This has also been demanded in the Delphi study of Jell *et al*[37].

Three of the five described studies used a 5G network provided by Huawei[37,52,53]. Huawei – a Chinese company - is the biggest 5G technology provider worldwide, far beyond other providers. In times of the predominance of virtual technologies, Huawei has been in the spotlight of geopolitical interests between China and the United States[58]. The matters of security and reliability are crucial in surgery. Not only does the network itself have to be stable enough to perform safe surgery, which has been proven by the outlined studies[37,50-53], but also cybersecurity now needs to be an issue of interest if the 5G wireless network will be the basis of modern surgery. Patient data and data transmission must be protected against cyber-attack.

Although the 5G technology already exists, the biggest technical issue is simply its pending implementation in daily life. In Germany, the expansion of the 5G network was subjected to a politically regulated and costly tender. Now, four providers have started to offer 5G in Germany, but it is far from being universally available[57]. Due to its high frequencies, more antennas are needed to create a stable 5G network. Therefore, more radio masts need to be constructed in urban as well as in natural areas. However, this might interfere with citizen movement and nature conservation organization interests.

From an ethical point of view, the topic of dehumanizing can become a greater issue when remote surgery becomes reality with 5G. The surgeon will not physically interact with his patient and therefore may be more likely to comprehend his patient just as a data set[15]. The sense of responsibility can be lost, and the criticized mechanization of medicine could be enhanced instead of resolved by 5G.

Consequently, 5G in surgery is not just a topic of medicine. It is a global, political issue which needs to be discussed from a political point of view as well. Recommendations and laws regarding how to handle this new technology in a practical but safe manner in medicine need to be addressed. These laws cannot just follow the geographical boundaries of states. One of the major advancements of the Internet is global networking. Practicing the advancement that surgery could be possible with 5G regardless of boundaries, legal and ethical agreements must be established on an international level.

CONCLUSION

Introduction of the fifth-generation wireless network, 5G, is forecasted to finally enable long-awaited establishment of telesurgery and telementoring in routine medical care. The first trials to prove the feasibility of remote surgery using the Internet were conducted decades ago and identified important network parameters for safe surgery, such as bandwidth, data rate and latency. 5G is supposed to meet these requirements with its enormous bandwidth, very short latency, multi-connectivity, high mobility, high availability, and high reliability.

Very few research studies are present in the literature to prove efficacy and feasibility of 5G in surgery so far and most of these studies are case studies. Nevertheless, all of them have reported safe surgery without connection disruption of the 5G network. However, these studies lack detailed information about the data rate and latency. More in-depth studies as well as finally randomized studies need to follow.

Combination of surgery with new technologies such as AI, ML, VR and AR using 5G as the providing network remains an issue of interest. Furthermore, questions like costs, political regulations on a national as well as international level, and data security need to be taken in consideration if 5G becomes an integral part in next-generation surgery.

REFERENCES

1 **Latif S**, Qadir J, Farooq S, Imran MA. How 5g wireless (and concomitant technologies) will revolutionize healthcare? *Future Int* 2017; **9**: 93 [DOI: 10.3390/fi9040093]

2 **Meara JG**, Leather AJ, Hagander L, Alkire BC, Alonso N, Ameh EA, Bickler SW, Conteh L, Dare AJ, Davies J, Mérisier ED, El-Halabi S, Farmer PE, Gawande A, Gillies R, Greenberg SL, Grimes CE, Gruen RL, Ismail EA, Kamara TB, Lavy C, Lundeg G, Mkandawire NC, Raykar NP, Riesel JN, Rodas E, Rose J, Roy N, Shrime MG, Sullivan R, Verguet S, Watters D, Weiser TG, Wilson IH, Yamey G, Yip W. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Int J Obstet Anesth* 2016; **25**: 75-78 [PMID: 26597405 DOI: 10.1016/j.ijoa.2015.09.006]

3 **World Health Organzation**. The world health report: 2006: working together for health. Geneva: World Health Organization, 2006

4 **Ballantyne GH**. Robotic surgery, telerobotic surgery, telepresence, and telementoring. Review of early clinical results. *Surg Endosc* 2002; **16**: 1389-1402 [PMID: 12140630 DOI: 10.1007/s00464-001-8283-7]

5 **Simon IB**. Telepresence surgery: remote surgery in the near future. *Bildgebung* 1994; **61** Suppl 1: 9-10 [PMID: 7919899]

6 **Marescaux J**, Leroy J, Rubino F, Smith M, Vix M, Simone M, Mutter D. Transcontinental robot-assisted remote telesurgery: feasibility and potential applications. *Ann Surg* 2002; **235**: 487-492 [PMID: 11923603 DOI: 10.1097/00000658-200204000-00005]

7 **Marescaux J**, Leroy J, Gagner M, Rubino F, Mutter D, Vix M, Butner SE, Smith MK. Transatlantic robot-assisted telesurgery. *Nature* 2001; **413**: 379-380 [PMID: 11574874 DOI: 10.1038/35096636]

8 **Fabrizio MD**, Lee BR, Chan DY, Stoianovici D, Jarrett TW, Yang C, Kavoussi LR. Effect of time delay on surgical performance during telesurgical manipulation. *J Endourol* 2000; **14**: 133-138 [PMID: 10772504 DOI: 10.1089/end.2000.14.133]

9 **Anvari M**. Telesurgery: remote knowledge translation in clinical surgery. *World J Surg* 2007; **31**: 1545-1550 [PMID: 17534550 DOI: 10.1007/s00268-007-9076-5]

10 **Anvari M**, McKinley C, Stein H. Establishment of the world's first telerobotic remote surgical service: for provision of advanced laparoscopic surgery in a rural community. *Ann Surg* 2005; **241**: 460-464 [PMID: 15729068 DOI: 10.1097/01.sla.0000154456.69815.ee]

11 **Anvari M**. Remote telepresence surgery: the Canadian experience. *Surg Endosc* 2007; **21**: 537-541 [PMID: 17279304 DOI: 10.1007/s00464-006-9040-8]

12 **Stanberry B**. Telemedicine: barriers and opportunities in the 21st century. *J Intern Med* 2000; **247**: 615-628 [PMID: 10886483 DOI: 10.1046/j.1365-2796.2000.00699.x]

13 **Choi PJ**, Oskouian RJ, Tubbs RS. Telesurgery: Past, Present, and Future. *Cureus* 2018; **10**: e2716 [PMID: 30079282 DOI: 10.7759/cureus.2716]

14 **Shahzad N**, Chawla T, Gala T. Telesurgery prospects in delivering healthcare in remote areas. *J Pak Med Assoc* 2019; **69** (Suppl 1): S69-S71 [PMID: 30697023]

15 **van Wynsberghe A**, Gastmans C. Telesurgery: an ethical appraisal. *J Med Ethics* 2008; **34**: e22 [PMID: 18827095 DOI: 10.1136/jme.2007.023952]

16 **Evans CR**, Medina MG, Dwyer AM. Telemedicine and telerobotics: from science fiction to reality. *Updates Surg* 2018; **70**: 357-362 [PMID: 30056519 DOI: 10.1007/s13304-018-0574-9]

17 **Haidegger T**, Sándor J, Benyó Z. Surgery in space: the future of robotic telesurgery. *Surg Endosc* 2011; **25**: 681-690 [PMID: 20652320 DOI: 10.1007/s00464-010-1243-3]

18 **Pande RU**, Patel Y, Powers CJ, D'Ancona G, Karamanoukian HL. The telecommunication revolution in the medical field: present applications and future perspective. *Curr Surg* 2003; **60**: 636-640 [PMID: 14972207 DOI: 10.1016/j.cursur.2003.07.009]

19 **Raison N**, Khan MS, Challacombe B. Telemedicine in Surgery: What are the Opportunities and Hurdles to Realising the Potential? *Curr Urol Rep* 2015; **16**: 43 [PMID: 26025497 DOI: 10.1007/s11934-015-0522-x]

20 **Smithwick M**. Network options for wide-area telesurgery. *J Telemed Telecare* 1995; **1**: 131-138 [PMID: 9375133 DOI: 10.1177/1357633X9500100302]

21 **Hall JC**. The internet: from basics to telesurgery. *ANZ J Surg* 2002; **72**: 35-39 [PMID: 11906422 DOI: 10.1046/j.1445-2197.2002.02288.x]

22 **Rayman R**, Croome K, Galbraith N, McClure R, Morady R, Peterson S, Smith S, Subotic V, Van Wynsberghe A, Patel R, Primak S. Robotic telesurgery: a real-world comparison of ground- and satellite-based internet performance. *Int J Med Robot* 2007; **3**: 111-116 [PMID: 17554810 DOI: 10.1002/rcs.133]

23 **Nguan C**, Miller B, Patel R, Luke PP, Schlachta CM. Pre-clinical remote telesurgery trial of a da Vinci telesurgery prototype. *Int J Med Robot* 2008; **4**: 304-309 [PMID: 18803341 DOI: 10.1002/rcs.210]

24 **Nguan CY**, Morady R, Wang C, Harrison D, Browning D, Rayman R, Luke PP. Robotic pyeloplasty using internet protocol and satellite network-based telesurgery. *Int J Med Robot* 2008; **4**: 10-14 [PMID: 18265415 DOI: 10.1002/rcs.173]

25 **Sterbis JR**, Hanly EJ, Herman BC, Marohn MR, Broderick TJ, Shih SP, Harnett B, Doarn C, Schenkman NS. Transcontinental telesurgical nephrectomy using the da Vinci robot in a porcine model. *Urology* 2008; **71**: 971-973 [PMID: 18295861 DOI: 10.1016/j.urology.2007.11.027]

26 **Xu S**, Perez M, Yang K, Perrenot C, Felblinger J, Hubert J. Determination of the latency effects on surgical performance and the acceptable latency levels in telesurgery using the dV-Trainer(®) simulator. *Surg Endosc* 2014; **28**: 2569-2576 [PMID: 24671353 DOI: 10.1007/s00464-014-3504-z]

27 **Anvari M**, Broderick T, Stein H, Chapman T, Ghodoussi M, Birch DW, McKinley C, Trudeau P, Dutta S, Goldsmith CH. The impact of latency on surgical precision and task completion during robotic-assisted remote telepresence surgery. *Comput Aided Surg* 2005; **10**: 93-99 [PMID: 16298920 DOI: 10.3109/10929080500228654]

28 **Rayman R**, Croome K, Galbraith N, McClure R, Morady R, Peterson S, Smith S, Subotic V, Van Wynsberghe A, Primak S. Long-distance robotic telesurgery: a feasibility study for care in remote environments. *Int J Med Robot* 2006; **2**: 216-224 [PMID: 17520635 DOI: 10.1002/rcs.99]

29 **Rayman R**, Primak S, Patel R, Moallem M, Morady R, Tavakoli M, Subotic V, Galbraith N, van Wynsberghe A, Croome K. Effects of latency on telesurgery: an experimental study. *Med Image Comput Comput Assist Interv* 2005; **8**: 57-64 [PMID: 16685943 DOI: 10.1007/11566489_8]

30 **Korte C**, Nair SS, Nistor V, Low TP, Doarn CR, Schaffner G. Determining the threshold of time-delay for teleoperation accuracy and efficiency in relation to telesurgery. *Telemed J E Health* 2014; **20**: 1078-1086 [PMID: 25290465 DOI: 10.1089/tmj.2013.0367]

31 **Perez M**, Xu S, Chauhan S, Tanaka A, Simpson K, Abdul-Muhsin H, Smith R. Impact of delay on telesurgical performance: study on the robotic simulator dV-Trainer. *Int J Comput Assist Radiol Surg* 2016; **11**: 581-587 [PMID: 26450105 DOI: 10.1007/s11548-015-1306-y]

32 **Marescaux J**, Rubino F. Robot-assisted remote surgery: technological advances, potential complications, and solutions. *Surg Technol Int* 2004; **12**: 23-26 [PMID: 15455307]

33 **Bove P**, Stoianovici D, Micali S, Patriciu A, Grassi N, Jarrett TW, Vespasiani G, Kavoussi LR. Is telesurgery a new reality? *J Endourol* 2003; **17**: 137-142 [PMID: 12803985 DOI: 10.1089/089277903321618699]

34 **Arata J**, Takahashi H, Pitakwatchara P, Warisawa S, Konishi K, Tanoue K, Ieiri S, Shimizu S, Nakashima N, Okamura K, Kim YS, Kim SM, Hahm JS, Hashizume M, Mitsuishi M. A remote surgery experiment between Japan-Korea using the minimally invasive surgical system. In: Proceedings 2006 IEEE International Conference on Robotics and Automation; 2006 May 15-19; Orlando, USA. IEEE, 2006 [DOI: 10.1109/ROBOT.2006.1641193]

35 **Lee BR**, Png DJ, Liew L, Fabrizio M, Li MK, Jarrett JW, Kavoussi LR. Laparoscopic telesurgery between the United States and Singapore. *Ann Acad Med Singap* 2000; **29**: 665-668 [PMID: 11126706]

36 **Martini MG**, Hewage CT, Nasralla MM, Smith R, Jourdan I, Rockall T. 3-D robotic tele-surgery and training over next generation wireless networks. *Annu Int Conf IEEE Eng Med Biol Soc* 2013; **2013**: 6244-6247 [PMID: 24111167 DOI: 10.1109/EMBC.2013.6610980]

37 **Jell A**, Vogel T, Ostler D, Marahrens N, Wilhelm D, Samm N, Eichinger J, Weigel W, Feussner H, Friess H, Kranzfelder M. 5th-Generation Mobile Communication: Data Highway for Surgery 4.0. *Surg Technol Int* 2019; **35**: 36-42 [PMID: 31694061]

38 **Li D**. 5G and intelligence medicine-how the next generation of wireless technology will reconstruct healthcare? *Precis Clin Med* 2019; **2**: 205-208 [PMID: 31886033 DOI: 10.1093/pcmedi/pbz020]

39 **Ahad A**, Tahir M, Aman Sheikh M, Ahmed KI, Mughees A, Numani A. Technologies Trend towards 5G Network for Smart Health-Care Using IoT: A Review. *Sensors (Basel)* 2020; **20** [PMID: 32708139 DOI: 10.3390/s20144047]

40 **Ahad A**, Tahir M, Yau KA. 5G-Based Smart Healthcare Network: Architecture, Taxonomy, Challenges and Future Research Directions. *IEEE Access* 2019; **7**: 100747-100762 [DOI: 10.1109/ACCESS.2019.2930628]

41 **Dananjayan S**, Raj GM. 5G in healthcare: how fast will be the transformation? *Ir J Med Sci* 2020 [PMID: 32737688 DOI: 10.1007/s11845-020-02329-w]

42 **Panwar N**, Sharma S, Singh AK. A survey on 5G: The next generation of mobile communication. *Physical Communication* 2016; **18**: 64-84

43 **Dohler M**. 5G Networks, Haptic Codecs, and the Operating Theatre. In: Atallah S, editor. Digital Surgery. Cham: Springer International Publishing, 2021: 71-86 [DOI: 10.1007/978-3-030-49100-0_6]

44 **O'Connell E**, Moore D, Newe T. Challenges Associated with Implementing 5G in Manufacturing. *Telecom* 2020; **1**: 48-67

45 **Taheribakhsh M**, Jafari A, Peiro MM, Kazemifard N. 5G Implementation: Major Issues and Challenges. In: 2020 25th International Computer Conference, Computer Society of Iran (CSICC); 2020 Jan 1-2; Teheran, Iran. IEEE, 2020

46 **Chen M**, Yang J, Hao Y, Mao S, Hwang K. A 5G cognitive system for healthcare. *Big Data Cogn Comput* 2017; **1**: 2 [DOI: 10.3390/bdcc1010002]

47 **Ullah H**, Nair NG, Moore A, Nugent C, Muschamp P, Cuevas M. 5G Communication: An Overview of Vehicle-to-Everything, Drones, and Healthcare Use-Cases. *IEEE Access* 2019; **7**: 37251-37268 [DOI: 10.1109/ACCESS.2019.2905347]

48 **China Daily**. China performs first 5G-based remote surgery on human brain: China Daily; 2019. Available from: http://www.chinadaily.com.cn/a/201903/18/WS5c8f0528a3106c65c34ef2b6.html

49 **Huawei**. World's First Remote Operation Using 5G Surgery: Huawei. Available from: https://www.huawei.com/en/industry-insights/outlook/mobile-broadband/wireless-for-sustainability/cases/worlds-first-remote-operation-using-5g-surgery

50 **Lacy AM**, Bravo R, Otero-Piñeiro AM, Pena R, De Lacy FB, Menchaca R, Balibrea JM. 5G-assisted telementored surgery. *Br J Surg* 2019; **106**: 1576-1579 [PMID: 31483054 DOI: 10.1002/bjs.11364]

51 **Acemoglu A**, Peretti G, Trimarchi M, Hysenbelli J, Krieglstein J, Geraldes A, Deshpande N, Ceysens PMV, Caldwell DG, Delsanto M, Barboni O, Vio T, Baggioni S, Vinciguerra A, Sanna A, Oleari E, Camillo Carobbio AL, Guastini L, Mora F, Mattos LS. Operating From a Distance: Robotic Vocal Cord 5G Telesurgery on a Cadaver. *Ann Intern Med* 2020; **173**: 940-941 [PMID: 32658568 DOI: 10.7326/M20-0418]

52 **Tian W**, Fan M, Zeng C, Liu Y, He D, Zhang Q. Telerobotic Spinal Surgery Based on 5G Network: The First 12 Cases. *Neurospine* 2020; **17**: 114-120 [PMID: 32252160 DOI: 10.14245/ns.1938454.227]

53 **Zheng J**, Wang Y, Zhang J, Guo W, Yang X, Luo L, Jiao W, Hu X, Yu Z, Wang C, Zhu L, Yang Z, Zhang M, Xie F, Jia Y, Li B, Li Z, Dong Q, Niu H. 5G ultra-remote robot-assisted laparoscopic surgery in China. *Surg Endosc* 2020; **34**: 5172-5180 [PMID: 32700149 DOI: 10.1007/s00464-020-07823-x]

54 **Challacombe B**, Patriciu A, Glass J, Aron M, Jarrett T, Kim F, Pinto P, Stoianovici D, Smeeton N, Tiptaft R, Kavoussi L, Dasgupta P. A randomized controlled trial of human *vs* robotic and telerobotic access to the kidney as the first step in percutaneous nephrolithotomy. *Comput Aided Surg* 2005; **10**: 165-171 [PMID: 16321914 DOI: 10.3109/10929080500229561]

55 **Augestad KM**, Bellika JG, Budrionis A, Chomutare T, Lindsetmo RO, Patel H, Delaney C; Mobile Medical Mentor (M3) Project. Surgical telementoring in knowledge translation--clinical outcomes and educational benefits: a comprehensive review. *Surg Innov* 2013; **20**: 273-281 [PMID: 23117447 DOI: 10.1177/1553350612465793]

56 **Boys JA**, Alicuben ET, DeMeester MJ, Worrell SG, Oh DS, Hagen JA, DeMeester SR. Public

perceptions on robotic surgery, hospitals with robots, and surgeons that use them. *Surg Endosc* 2016; **30**: 1310-1316 [PMID: 26173543 DOI: 10.1007/s00464-015-4368-6]

57 **Bünder H**. Milliarden Einnahmen für Deutscland - "Das Geld fehlt jetzt für den Netzausbau". Frankfurter Allgemeine Zeitung, 12. Jun 2019. Available from: https://www.faz.net/aktuell/wirtschaft/digitec/5g-mobilfunk-auktion-beendet-6-6-milliarden-euro-einnahmen-fuer-deutschland-16233785.html

58 **Tekir G**. Huawei, 5G Network and Digital Geopolitics. *Inter J Politic Security* 2020; **2**: 113-135

59 **Veneziano D**, Tafuri A, Rivas JG, Dourado A, Okhunov Z, Somani BK, Marino N, Fuchs G, Cacciamani G; ESUT-YAUWP Group. Is remote live urologic surgery a reality? *World J Urol* 2020; **38**: 2367-2376 [PMID: 31701210 DOI: 10.1007/s00345-019-02996-0]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com

ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 12-49

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 2 April 28, 2021

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 2 April 28, 2021

Contents

ABOUT COVER

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Krish Ragunath, Fatih Altintoprak, Sahin Coban

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
April 28, 2021

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 12-24

DOI: 10.37126/aige.v2.i2.12

ISSN 2689-7164 (online)

MINIREVIEWS

Application of deep learning in image recognition and diagnosis of gastric cancer

Yu Li, Da Zhou, Tao-Tao Liu, Xi-Zhong Shen

ORCID number: Yu Li 0000-0002-4685-3418; Da Zhou 0000-0001-8838-1351; Tao-Tao Liu 0000-0002-4623-9012; Xi-Zhong Shen 0000-0003-3748-0709.

Author contributions: TT Liu and D Zhou contributed equally to conceptual development and supervision; Y Li and D Zhou contributed to the data collection and manuscript; XZ Shen supervised the paper; All authors have read and approved the final manuscript.

Supported by National Natural Science Foundation of China, No. 81800510; Shanghai Sailing Program, No. 18YF1415900.

Conflict-of-interest statement: The authors report no conflicts of interest in this work.

Yu Li, Da Zhou, Tao-Tao Liu, Xi-Zhong Shen, Department of Gastroenterology and Hepatology, Zhongshan Hospital Affiliated to Fudan University, Shanghai 200032, China

Corresponding author: Da Zhou, PhD, Doctor, Research Fellow, Department of Gastroenterology and Hepatology, Zhongshan Hospital Affiliated to Fudan University, 180 Feng Lin road, Shanghai 200032, China. mubing2007@foxmail.com

Abstract

In recent years, artificial intelligence has been extensively applied in the diagnosis of gastric cancer based on medical imaging. In particular, using deep learning as one of the mainstream approaches in image processing has made remarkable progress. In this paper, we also provide a comprehensive literature survey using four electronic databases, PubMed, EMBASE, Web of Science, and Cochrane. The literature search is performed until November 2020. This article provides a summary of the existing algorithm of image recognition, reviews the available datasets used in gastric cancer diagnosis and the current trends in applications of deep learning theory in image recognition of gastric cancer. covers the theory of deep learning on endoscopic image recognition. We further evaluate the advantages and disadvantages of the current algorithms and summarize the characteristics of the existing image datasets, then combined with the latest progress in deep learning theory, and propose suggestions on the applications of optimization algorithms. Based on the existing research and application, the label, quantity, size, resolutions, and other aspects of the image dataset are also discussed. The future developments of this field are analyzed from two perspectives including algorithm optimization and data support, aiming to improve the diagnosis accuracy and reduce the risk of misdiagnosis.

Key Words: Endoscope; Artificial intelligence; Algorithm optimization; Data support

Core Tip: Gastric cancer is a life-threatening disease with a high mortality rate. With the development of deep learning in the image processing of gastrointestinal endoscope, the efficiency and accuracy of gastric cancer diagnosis through imaging technology have been greatly improved. At present, there is no comprehensive

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and Hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): A
Grade B (Very good): B
Grade C (Good): 0
Grade D (Fair): D
Grade E (Poor): 0

Received: February 15, 2021
Peer-review started: February 15, 2021
First decision: March 16, 2021
Revised: March 30, 2021
Accepted: April 20, 2021
Article in press: April 20, 2021
Published online: April 28, 2021

P-Reviewer: Nayyar A, Taira K, Tanabe S
S-Editor: Wang JL
L-Editor: A
P-Editor: Wang LL

summary on the graphic recognition method for gastric cancer based on deep learning. In this review, some gastric cancer image databases and mainstream gastric cancer recognition models were summarized to make a prospect for the application of deep learning in this field.

Citation: Li Y, Zhou D, Liu TT, Shen XZ. Application of deep learning in image recognition and diagnosis of gastric cancer. *Artif Intell Gastrointest Endosc* 2021; 2(2): 12-24
URL: https://www.wjgnet.com/2689-7164/full/v2/i2/12.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i2.12

INTRODUCTION

Gastric cancer is a life-threatening disease with a high mortality rate[1]. Globally, more than 900000 individuals develop gastric cancer each year out of which more than 700000 Lose their lives. Gastric cancer is second only to lung cancer in terms of mortality[2]. Unlike the developing countries, the number of diagnosed cases and the mortality rate of this cancer are declining in the developed countries such as those in the EU and North America[3,4].

Around 50% of the world's gastric cancer cases are diagnosed in Southeast Asia[5]. In China, gastric cancer is also second to lung cancer in terms of the number of annual cases, for instance, 424000 new patients are annually diagnosed with gastric cancer, accounting for more than 40% of the global total, out of which 392000 Lose their lives ranking the fifth and the sixth worldwide in annual morbidity and mortality, respectively[6].

The diagnosis of gastric cancer mainly relies on clinical manifestation, pathological images and medical imaging[7]. Compared with other methods such as pathological diagnosis, medical imaging provides a simple non-invasive and reliable method for the diagnosis of gastric cancer which is more accessible and efficient, easier to operate and has almost no side effects for the patients[8].

Doctors make a judgment based on medical imaging which mainly depend on their experience from similar cases, hence, occasional misdiagnosis is inevitable[9,10]. With the rapid development of computer technology and artificial intelligence, deep learning techniques are extremely effective in various branches of image processing and have been used in medical imaging to improve cancer diagnosis[11-13]. Danaee *et al*[14] established a deep learning model for colorectal cancer image recognition, the results showed that the deep learning method can achieve more effective information and is far more efficient than the way of manual extraction. Burke *et al*[15] found that deep learning could classify and predict mutations of NSCLC based on histopathological images, and the recognition efficiency of deep learning was much higher than that of manual recognition. Muhammad Owais *et al*[16] proposed a deep learning model to classify a variety of gastrointestinal diseases by recognizing endoscopic videos. This model can simultaneously extract spatiotemporal features to achieve better classification performance. Experimental results of the proposed models showed superior performance to the latest technology and indicated its potential in clinical application[16].

Endoscopic images are mostly used in gastric cancer diagnosis[17]. Endoscopic images contain a lot of useful structural information which can be used for deep learning algorithm, the algorithm can carry out purposeful image recognition[18]. Most of the image recognition based on gastric cancer diagnosis methods adopt supervised deep learning algorithms, mainly because the monitored network in supervised learning makes full use of the labeled sample data in the training and can obtain more accurate segmentation results[19].

In fact, the purpose of medical image recognition is to identify the tumor and we call this process image segmentation[20-22]. Accurate segmentation of tumor images is an important step in diagnosis, surgical planning and postoperative evaluation[23,24]. Endoscopic images segmentation can provide more comprehensive information for the diagnosis and treatment of gastric cancer, alleviate the doctor's heavy work for reading film and improve the accuracy of diagnosis[25]. However, due to the variety and complexity of gastric tumor types, segmentation has become an important and difficult problem in computer-aided diagnosis. Compared with the traditional

segmentation methods, the deep learning segmentation method of gastric tumor image has achieved obvious improved performance and rapid development[26,27].

As mentioned above, the deep learning method based on supervised learning can fully mine the effective information of existing data. However, when the amount of existing data cannot meet the requirements of model training, it is necessary to find ways to increase the data scale[28]. The deep learning based on unsupervised learning can generate samples, which are similar to the existing samples in dimension and structure, but not identical. At present, relevant research results have been obtained[29]. Researchers use semi-supervised and unsupervised image recognition algorithms to generate samples like training samples, to improve the accuracy of gastric cancer tumor recognition and enhance the robustness of the model[30].

In this paper, deep learning-based diagnosis of gastric cancer based on endoscope images is summarized and analyzed. The adopted segmentation networks in the previous works can be divided into three categories: the supervised network, semi-supervised network, and unsupervised network. The basic idea of the recognition method, the basic structure of the network, the experimental results, as well as their advantages and disadvantages are summarized. The performance of typical methods above-mentioned in recognition is compared. Finally, we hope to provide insights and concluding remarks on the development of deep-learning-based diagnosis of gastric cancer.

RELEVANT DATA SETS AND ALGORITHM EVALUATION INDEXES

Relevant datasets

To promote the progress of image recognition and make an objective comparison of available image recognition methods for gastric cancer diagnosis, we investigate the commonly used datasets including the GR-AIDS provided by Medical Image Computing and Computer Assisted Intervention Society as well as those internal datasets.

The GR-AIDS dataset established by Sun Yat-Sen University Cancer Center consists of 1036496 endoscopic images from 84424 individuals. This dataset is used according to the 8:1:1 pattern, the data is randomly selected for training and internal validation datasets for GR-AIDS development as well as for evaluating GR-AIDS performance[31].

Using clinical data collected from Gil Hospital, Jang Hyung Lee *et al*[32] also established a data set containing 200 normal cases, 367 cancer cases, and 220 ulcer cases. The data was divided into training sets of 180, 200, 337 images and test sets of 20, 30, 20 images. To improve the local contrast of the image and enhance the edge definition in each area of the image, histogram equalization was adopted to further enhance the image, the images' size was adjusted to 224 × 224 pixels [32].

Hirasawa *et al*[32] collected 13,584 endoscopic images of gastric cancer to build an image database. To evaluate the diagnostic accuracy, an independent test set of 2296 gastric images was collected from 69 patients with continuous gastric cancer lesions constructed as convolutional neural network (CNN). The image has an in-plane resolution of 512 × 512[33].

Cho *et al*[34] collected 5017 images from 1269 patients, of which 812 images from 212 patients were used as the test data set. An additional 200 images from 200 patients were collected and used for prospective validation. The resolution of the images is 512 × 512. The information for all major databases is shown in Table 1[34].

Introduction of evaluation indexes

To evaluate the effectiveness of each model in diagnosing gastric cancer, the following evaluation indicators are commonly used in the related literature (Table 2): DICE Similarity Coefficient (DICE, 1945), Jaccard Coefficient (Jaccard, 1912), Volumetric Over-lap Error (VOE), and Relative Volume Difference (RVD).

Here we define the following variables: P and N are used for judgment of the model results, T and F evaluation model of the judgment is correct, FP is on behalf of the false-positive cases, FN represents false-negative cases, TP is on behalf of the real example, TN represents true negative cases[38]. A represents the theory of segmentation, results for comparison with the resulting image. B represents the segmentation results[39]. The relationship among them is shown in Figure 1.

DICE coefficient: DICE coefficient also known as the overlap index, is one of the most commonly used indexes for verification of image segmentation. The DICE coefficient

Table 1 Commonly used databases in image recognition of gastric cancer

Database	Time collected	Number of samples	Resolution	Training set	Test set
GR-AIDS[31]	2019	1036496	512 × 512	829197	103650
Jang Hyung Lee[32]	2019	787	224 × 224	717	70
Toshiaki Hirasawa[33]	2018	13584	512 × 512	13584	2496
Bum-Joo Cho[34]	2019	5017	512 × 512	4205	812
Hiroya Ueyama[35]	2020	7874	512 × 512	5574	2300
Lan Li[36]	2020	2088	512 × 512	1747	341
Mads Sylvest Bergholt[37]	2011	1063	512 × 512	850	213

Table 2 Specific concepts of the main evaluation indicators

Index	Description	Usage	Unit
DICE	Repeat rate between the segmentation results and markers	Commonly	%
RMSD	The root mean square of the symmetrical position surface distance between the segmentation results and the markers	Commonly	mm
VOE	The degree of overlap between the segmentation results and the actual segmentation results represents the error rate	Commonly	%
RVD	The difference in volume between the segmentation results and the markers	Rarely	%

DICE: DICE Similarity Coefficient; RMSD: Root-Mean-Square Deviation; VOE: Volumetric Over-lap Error; RVD: Relative Volume Difference.

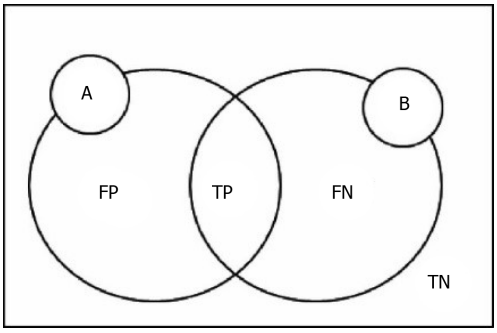

Figure 1 Schematic diagram of each evaluation index relationship. TP: True positive; FP: False-positive; TN: True negative; FN: False negative.

represents the repetition rate between the segmentation results and the markers. The value range of DICE is 0-1, where 0 indicates that the experimental segmentation result significantly deviates from the labeled result, and 1 indicates that the experimental segmentation result completely coincides with the labeled result[40]. DICE coefficient is defined as the following:

$$DICE = (2|A \cap B|)/(|A| + |B|) = (2TP)/(2TP + FP + FN)$$

Jaccard coefficient: Jaccard coefficient represents the similarity and difference between the segmentation result and the standard. The larger the coefficient, the higher the sample similarity. Besides, the Jaccard coefficient and DICE coefficient are correlated[41]. Jaccard coefficient is defined as the following:

$$JAC = (|A \cap B|)/(|A| \cup |B|) = TP/(TP + FP + FN) = DICE/(2 - DICE)$$

VOE: VOE stands for error rate, derived from Jaccard. VOE is represented as %, where 0% indicates complete segmentation. If there is no overlap between the segmentation result and the markers, the VOE is 100%[42]. VOE is defined as the following:

$$VOE = 1 - (|A \cap B|)/(|A| \cup |B|) = 1 - TP/(TP + FP + FN)$$

RVD: RVD represents the noise difference between the segmentation result and the markers. RVD is presented as %, where 0% denotes the same volume between the segmentation result and the markers[42]. The formula is:

$RVD = (|B| - |A|)/|A| = FP/(TP + FN)$

The specific concepts of all indicators are shown in Table 2.

CLASSIFICATION OF THE ALGORITHM

Supervised learning-based diagnosis of gastric cancer

Deep neural networks are often trained based on deep learning algorithms using large labeled datasets (*i.e.*, images in this case)[43]. The network is therefore able to learn how features are related to the target[44]. Since the data is already labeled, this learning method is referred to as supervised learning. Most of the existing studies on diagnosing gastric cancer are based on supervised learning in image recognition tasks[45-47]. This is because the network makes full use of the labeled dataset in the training, hence can obtain more accurate segmentation results.

Recent research works showed that CNN achieves outstanding performance in various image recognition tasks[48,49]. Toshiaki Hirasawa built a CNN-based diagnostic system based on a single-shot Multi-Box detector, Adejub, with a total sensitivity of 92.2% and trained their CNN using 13584 endoscopic images of gastric cancer. The trained CNN correctly called 71 out of 77 cases of gastric cancer, *i.e.*, a total sensitivity of 92.2%, also detected 161 non-cancerous samples as gastric cancer, *i.e.*, a positive predictive value of 30.6%. The CNN also correctly detected 70 of 71 cases of gastric cancer (98.6%) with a diameter of 6 mm or larger, as well as all invasive cancers[33]. Ueyama *et al*[35] also constructed an AI-assisted CNN based computer-aided diagnosis system with narrow band imaging-magnifying endoscopy images.

The above studies show that the CNN-based approach is far more accurate than human in recognition of cancer. This makes us believe that the method based on deep CNN can effectively solve the identification problem of gastric cancer.

However, the issue with the CNN is that only partial features could be extracted[50]. Due to the imbalanced information of gastric cancer image data, extracting the local features does not reflect all the information and might harm the efficiency of the image recognition. To address the problem, Shelhamer *et al*[51] proposed full convolutional neural network (FCN) for image segmentation. This network attempts to recover the category of each pixel from the abstract feature, in other words, instead of image-level classification, the network uses pixel-level classification[51]. This addresses the semantic level image segmentation problem and is the core component of many advanced semantic segmentation models[52,53].

The segmentation method of gastric cancer images based on the FCN network is mainly based on the idea of code-decoding design[54]. In practice, the image is classified at the pixel level and the network is pre-trained with supervision. In this method, the input image can have any arbitrary size and the output of the same size can be generated through effective reasoning and learning[55]. Typical FCN network-based image segmentation architecture for gastric cancer is shown in Figure 2.

The FCN is improved based on the CNN by transforming the last three full connections into three convolutional layers. The success of FCN network is largely attributed to the excellent ability of CNN network to extract hierarchical representation. In the concrete implementation process, the network realizes the segmentation of gastric tumor by down-sampling and up-sampling through convolution-deconvolution operation. The down-sampling path consists of convolution layer and maximum or average pooling layer, which can extract high-level semantic information, but its spatial resolution is often low. The up-sample path consists of convolution and a deconvolution layer (also known as transpose convolution) and uses the output of the down-sample path to predicting the fraction of each class at the pixel level[56,57]. However, the output image of deconvolution operations might be very rough and lost a lot of detail. The skip structure of the FCN network presented in the classified forecast comes from the deep layer (thick) semantic information and information from the appearance of the shallow layer (fine), thus, achieving a more accurate and robust segmentation result. As a deep neural network, FCN has shown good performance in many challenging medical image segmentation tasks, including liver tumor segmentation[58,59].

One of the most important features of the FCN is the use of skip structure. It is used to fuse the feature information of both the high and low layers. Through the cross-layer connection structure, the texture information of the shallow layer and the semantic information of the deep layer of the network are then combined to achieve the precise segmentation task[60,61]. Jang Hyung Lee improved the original FCN framework by applying the pre-trained Inception, Res-Net, and VGG-Net models on

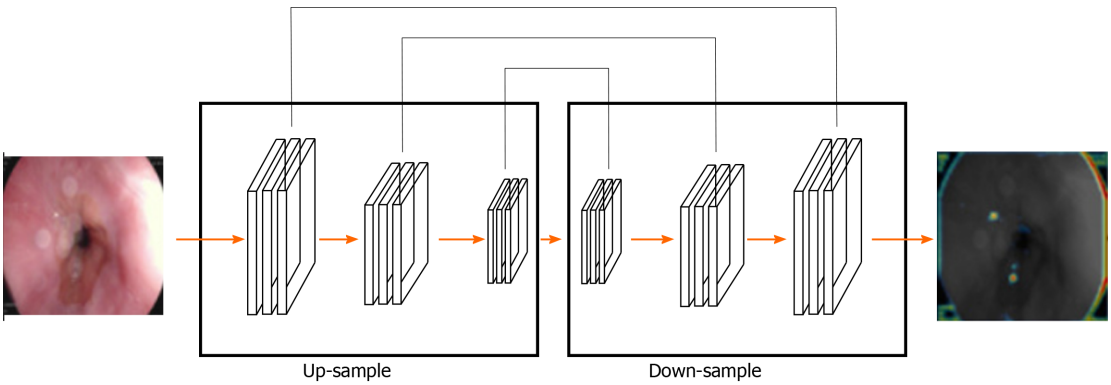

Figure 2 The basic architecture of image segmentation for gastric cancer based on full convolutional network.

ImageNet. The areas under the operating characteristic curves of each receiver are 0.95, 0.97, and 0.85, respectively, hence, Res-Net shows the highest level of performance. Under normal conditions, the classification between normal and ulcer or cancer, is more than 90 percent accurate[32].

The deep network structure leads to the problem of decreased training accuracy[62]. In Sun *et al*[63] the basic form of convolution is replaced with the deformable convolution and Atrous convolution in a specific layer to adapt to the non-rigid characteristics and large receiving fields. The Atrous space pyramid pooling module and the semantic-level embedded network based on encoder/decoder are used for multi-scale segmentation. Besides, they proposed a lightweight decoder to fuse the context information and further used dense up-sampled convolution for the boundary optimization at the end of the decoder. The model achieves 91.60% pixel-level accuracy and 82.65% average degree of the intersection[63].

Cho *et al*[34] established the Inception-ResNET-V2 model, which is an FCN model. In this model, they divided the images into five categories: advanced gastric cancer, early gastric cancer, high atypical hyperplasia, low atypical hyperplasia and non-neoplastic. For the above five categories, the Inception- ResNet-v2 model has a weighted average accuracy of 84.6%. The mean area under the curve of the model for differentiating gastric cancer and neoplasm was 0.877 and 0.927, respectively[34].

The above works show that FCN addresses the issue with the CNN hence can extract the local features. This is why the FCN is considered as the mainstream in gastric cancer image classification methods.

In addition to the application of FCN to address the shortcomings of CNN, researchers also tried other approaches such as fusion of multiple CNN methods to obtain an Ensemble of CNN algorithm to get more accurate classification results. Nguyen *et al*[64] trained three different CNN model architectures, including VGG-based, Inception-based Network and Dense-Net. In their study, the VGG-based network was used as a conventional deep CNN for classification problems, which consists of a linear stack of the convolutional layer. The network-based on Dense-net can be used as a very deep CNN with a short path, which is also helpful to train the network and extract more abstract and effective image features easily. The three models were trained separately, the AVERAGE combination rule is then used to combine the classification results of the three CNN-based Models. The final result was 70.369% of overall classification accuracy, 68.452% of sensitivity and 72.571% of specificity. The overall classification accuracy is higher than that generated by the listed model based on a single CNN[64].

Both the use of a fully convolutional network and the fusion of several CNN algorithms are significantly effective in improving the accuracy of gastric cancer image recognition. They are also effective in addressing the issues with the quality of images in the database. Table 3 shows the performance comparison of gastric cancer image recognition by using CNN, FCN, and Ensemble CNN.

Image recognition based on semi-supervised and unsupervised learning in gastric cancer

Most gastric cancer image recognition methods adopt supervised learning algorithms because the monitored network makes full use of the labeled sample data in the training and can obtain more accurate segmentation results. Nevertheless, there are very few accurately labeled image datasets, hence researchers have carried out studies

Table 3 Comparison of recognition performance of convolutional neural network, full convolutional neural network, and ensemble convolutional neural network models

Methods	DICE/%	VOE/%	RMSD/mm
Toshiaki Hirasawa (CNN)	0.5738	0.5977	6.491
Hiroya Ueyama (CNN)	0.6327	0.5373	7.257
Jang Hyung Lee (FCN)	0.8102	0.319	2.468
Bum-Joo Cho (FCN)	0.9350	0.1221	-
Dat Tien Nguyen (ECNN)	0.8947	0.113	-

CNN: Convolutional neural network; FCN: Full convolutional neural network; ECNN: Evolutionary convolutional neural network; DICE: DICE Similarity Coefficient; VOE: Volumetric Over-lap Error; RMSD: Root-Mean-Square Deviation.

based on semi-supervised and unsupervised image recognition algorithms for gastric cancer. In such studies, they trained a small number of samples through generative models to generate similar samples to improve the accuracy and robustness of gastric cancer tumor recognition[65].

Generative adversarial network (GAN) is a generative model proposed by Goodfellow *et al*[66] It uses an unsupervised training method that is trained by adversarial learning. The objective is to estimate the potential distribution of data samples and generate new data samples. GAN is composed of a generation model (Goodfellow *et al*[66], 2014) and a discrimination model (Denton *et al*, 2015). The generation model learns the distribution of a given noise (generally refers to uniform distribution or normal distribution) and synthesizes it, whereas the discrimination model distinguishes the real data from generated data. In theory, the former is trying to produce data that is closed to the real data. The latter is also constantly strengthening the "counterfeit detection" ability[67]. The success of GAN lies in its ability to capture high-level semantic information using adversarial learning techniques. Luc *et al*[68] first applied GAN to image segmentation. However, GAN has several drawbacks: (1) Crash problem: when the generation model crashes, all different inputs are mapped to the same data[69]; and (2) Instability: It causes the same input to produce different outputs. The main reason is due to gradient vanishing problem during the optimization process[66,70].

Although batch normalization is often used to solve the instability of GAN, it is often not enough to achieve optimal stability of GAN performance. Therefore, many GAN derived models have emerged to solve these gaps, *e.g.*, conditional GAN, deep convolutional GAN, information maxi-mizing GAN, Wassertein GAN, *etc*[71]. In the GAN-based image recognition for gastric cancer, the generator is used to perform the segmentation task. The discriminator is then used to train the refining generator. A typical gastric cancer image recognition architecture based on the generative adversarial network is illustrated in Figure 3.

Since its proposal generative adversarial network has been widely considered and rapidly developed in different application areas. In medical image processing, it is very challenging to construct a large enough dataset due to the difficulty of data acquisition and annotation[72]. To overcome this problem, traditional image enhancement technology such as geometric transformation is often used to generate new data. This technique cannot learn biological changes in medical data and can produce images that are not credible[73]. Although GAN is unable to know in advance hypothesis distribution due to the limitation of segmentation performance improvement. it can automatically infer real data sets, further expand the scale and diversity of data, and provide a new method for data expansion, thus improving the efficiency of model training[74,75].

Almalioglu *et al*[76] showed that the poor resolution of the capsule endoscope is a limiting factor in the accuracy of diagnosis. They designed an image synthesis technology based on GAN to enrich the training data. First, the standard data expansion method was used to enlarge the dataset. Then the dataset was used to train GAN and the proposed Endol2h method was used to synthesize gastric cancer images with higher resolution[76]. Wang proposed an unsupervised image classification method for tumors based on prototype migration generated against the network (Prototype Transfer Generative Adversarial Network). Using different data acquisition devices and parameter settings caused differences in the stye of tumor image and data distribution. These differences can be reduced by designing the target domain to

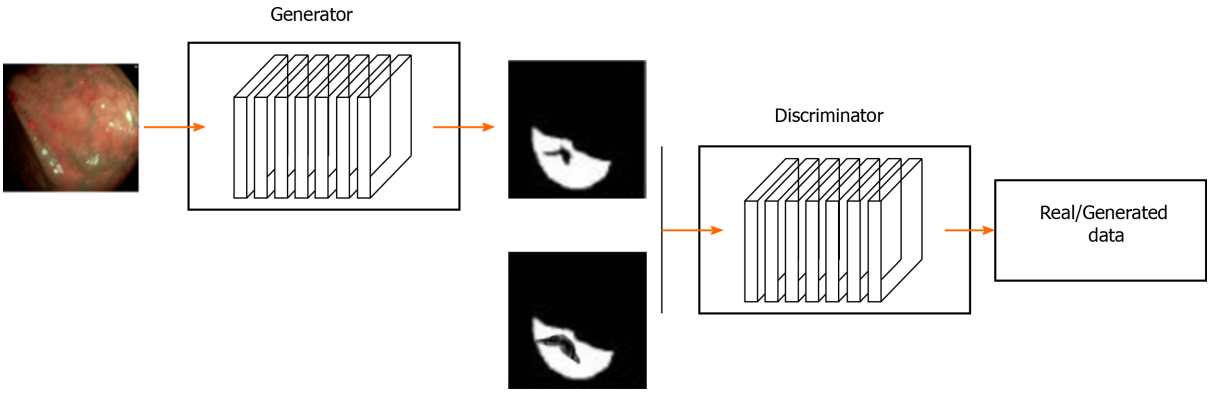

Figure 3 Basic architecture of gastric cancer image recognition methods based on generative adversarial network.

generate network, training process through the domain discriminant and performing generator reconstruction between source domain and target domain. The method achieved an average accuracy of 87.6% for unsupervised breast tumor image dichotomy under different magnifications and shows good scalability[77].

In conclusion, the GAN-based image segmentation method for gastric cancer can generate realistic gastric cancer images through the GAN network in the training stage, thus avoid the imbalance of the training samples. Moreover, due to the amplification of limited labeled sample data, the deep network is well-trained and achieves a high segmentation efficiency. However, there are still many problems in GAN, such as the instability of training and the breakdown of the training network. Therefore, researchers have optimized the original GAN network to reduce data noise or deal with class imbalance and other problems. In order to solve the problem that medical images are often polluted by different amounts and types of praise, T.Y Zhang *et al*. propose a novel Noise Adaptation Generative Adversarial Network (NAGAN), which contains a generator and two discriminators. The generator aims to map the data from source domain to target domain. Among the two discriminators, one discriminator enforces the generated images to have the same noise patterns as those from the target domain, and the second discriminator enforces the content to be preserved in the generated images. They apply the proposed NAGAN on both optical coherence tomography images and ultrasound images. Results show that the method is able to translate the noise style[74]. In the traditional GAN network training, the small number of samples of the minority classes in the training data makes the learning of optimal classification challenging, while the more frequently occurring samples of the majority class hamper the generalization of the classification boundary between infrequently occurring target objects and classes. Mina Rezaei *et al*. developed a novel generative multi-adversarial network, called Ensemble-GAN, for mitigating this class imbalance problem in the semantic segmentation of abdominal images. The Ensemble-GAN framework is composed of a single-generator and a multi-discriminator variant for handling the class imbalance problem to provide a better generalization than existing approaches[73]. In addition, there are other studies on the optimization of GAN network in medical image segmentation. Klages *et al*[78] proposed the patch-based generative adversarial neural network models, this model can significantly reduce errors in data generation. Nuo Tong *et al*[79] proposed the self-paced Dense-Net with boundary constraint for automated multi-organ segmentation on abdominal CT images. Specifically, a learning-based attention mechanism and dense connection block are seamlessly integrated into the proposed self-paced Dense-Net to improve the learning capability and efficiency of the backbone network. In a word, in the process of optimizing GAN network, whether it is optimizing generator or discriminant, the purpose of optimization is to generate new data which is as equal to the real data as possible. Therefore, more studies will be devoted to the optimization of GAN network to provide strong support for improving the image recognition of gastric cancer.

Table 4 shows comparison results of the three current mainstream methods for image recognition of gastric cancer.

Table 4 Comparison of convolutional neural network, full convolutional neural network, and generative adversarial network models

Model features	Contributions	Advantages	Disadvantages	Scope of application
CNN	The topology can be extracted from a two-dimensional image, and the backpropagation algorithm is used to optimize the network structure and solve the unknown parameters in the network	Shared convolution kernel, processing high-dimensional data without pressure; Feature extraction can be done automatically	When the network layer is too deep, the parameters near the input layer will be changed slowly by using BP propagation to modify parameters. A gradient descent algorithm is used to make the training results converge to the local minimum rather than the global minimum. The pooling layer will lose a lot of valuable information	Suitable for data scenarios with similar network structures
FCN	The end-to-end convolutional network is extended to semantic segmentation. The deconvolution layer is used for up-sampling; A skip connection is proposed to improve the roughness of the upper sampling	Can accept any size; Input image; Jump junction; The structure combines fine layers and coarse; Rough layers, generating precise segmentation	The receptive field is too small to obtain the global information;Small storage overhead	Applicable to large sample data
GAN	With adversarial learning criteria, there are two No's: The same network, not a single network	Can produce a clearer, more realistic sample; any generated network can be trained	Training is unstable and difficult to train; GAN is not suitable for processing data in discrete form	Suitable for data generation (*e.g.*, there are not many data sets with labels), image style transfer; Image denoising and restoration; Used to counter attacks

CNN: Convolutional neural network; FCN: Full convolutional neural network; GAN: Generative adversarial network.

CONCLUSION

At the present, the development direction of deep learning in image recognition of gastric cancer mainly focuses on the following aspects: (1) Training of deep learning algorithms relies on the availability of large datasets, because medical images are often difficult to obtain, medical professionals need to spend a lot of time on data collection and annotation which is time-consuming and costly. Besides, medical workers need not only to provide a large amount of data support but also to make use of all the effective information in the data as much as possible. Deep neural networks enable full mining of the information content of the data. Using deep networks seems to be the dominant future research direction in this field; (2) Multimodal gastric image segmentation combined with several different deep neural networks are used to extract the deeper information of the image and improve the accuracy of tumor segmentation and recognition. This is a promising major research direction in this field; and (3) Currently, most of the medical image segmentation techniques use supervised deep learning algorithms. However, for some of the rare diseases lacking a large number of data samples, supervised deep learning algorithms cannot reach their full efficiency. To overcome the issue with the lack of large datasets, some researchers utilize semi-supervised or unsupervised techniques such as GAN and combine the generated adversarial network with other higher performance networks. This might be another emerging research trend in this area.

REFERENCES

1 **Higgins AJ**, Lees P. Arachidonic acid metabolites in carrageenin-induced equine inflammatory exudate. *J Vet Pharmacol Ther* 1984; **7**: 65-72 [PMID: 6423835 DOI: 10.1111/j.1365-2885.1984.tb00881.x]

2 **Torre LA**, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* 2015; **65**: 87-108 [PMID: 25651787 DOI: 10.3322/caac.21262]

3 **Ferlay J**, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, Gavin A, Visser O, Bray F. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer* 2018; **103**: 356-387 [PMID: 30100160 DOI: 10.1016/j.ejca.2018.07.005]

4 **Howlader N NA**, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA. SEER Cancer Statistics Review, 1975-2017. Bethesda: National Cancer Institute, 2019. Available from: https://seer.cancer.gov/csr/1975_2017/

5 **Rahman R**, Asombang AW, Ibdah JA. Characteristics of gastric cancer in Asia. *World J Gastroenterol* 2014; **20**: 4483-4490 [PMID: 24782601 DOI: 10.3748/wjg.v20.i16.4483]

6 **Chen W**, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; **66**: 115-132 [PMID: 26808342 DOI: 10.3322/caac.21338]

7 **Homeida AM**, Cooke RG. Pharmacological aspects of metaldehyde poisoning in mice. *J Vet Pharmacol Ther* 1982; **5**: 77-81 [PMID: 6178838 DOI: 10.1111/j.1365-2885.1982.tb00500.x]

8 **Luo X**, Mori K, Peters TM. Advanced Endoscopic Navigation: Surgical Big Data, Methodology, and Applications. *Annu Rev Biomed Eng* 2018; **20**: 221-251 [PMID: 29505729 DOI: 10.1146/annurev-bioeng-062117-120917]

9 **de Groof J**, van der Sommen F, van der Putten J, Struyvenberg MR, Zinger S, Curvers WL, Pech O, Meining A, Neuhaus H, Bisschops R, Schoon EJ, de With PH, Bergman JJ. The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019; **7**: 538-547 [PMID: 31065371 DOI: 10.1177/2050640619837443]

10 **Qu JY**, Li Z, Su JR, Ma MJ, Xu CQ, Zhang AJ, Liu CX, Yuan HP, Chu YL, Lang CC, Huang LY, Lu L, Li YQ, Zuo XL. Development and Validation of an Automatic Image-Recognition Endoscopic Report Generation System: A Multicenter Study. *Clin Transl Gastroenterol* 2020; **12**: e00282 [PMID: 33395075 DOI: 10.14309/ctg.0000000000000282]

11 **Gan T**, Liu S, Yang J, Zeng B, Yang L. A pilot trial of Convolution Neural Network for automatic retention-monitoring of capsule endoscopes in the stomach and duodenal bulb. *Sci Rep* 2020; **10**: 4103 [PMID: 32139758 DOI: 10.1038/s41598-020-60969-5]

12 **Shin HC**, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging* 2016; **35**: 1285-1298 [PMID: 26886976 DOI: 10.1109/TMI.2016.2528162]

13 **Bernal J**, Kushibar K, Asfaw DS, Valverde S, Oliver A, Martí R, Lladó X. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artif Intell Med* 2019; **95**: 64-81 [PMID: 30195984 DOI: 10.1016/j.artmed.2018.08.008]

14 **Danaee P**, Ghaeini R, Hendrix DA. A Deep Learning Approach for Cancer Detection and Relevant Gene Identification. *Pac Symp Biocomput* 2017; **22**: 219-229 [PMID: 27896977 DOI: 10.1142/9789813207813_0022]

15 **Burke HB**, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE Jr, Marks JR, Winchester DP, Bostwick DG. Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer* 1997; **79**: 857-862 [PMID: 9024725 DOI: 10.1002/(sici)1097-0142(19970215)79:4<857::aid-cncr24>3.0.co;2-y]

16 **Owais M**, Arsalan M, Choi J, Mahmood T, Park KR. Artificial Intelligence-Based Classification of Multiple Gastrointestinal Diseases Using Endoscopy Videos for Clinical Diagnosis. *J Clin Med* 2019; **8** [PMID: 31284687 DOI: 10.3390/jcm8070986]

17 **Lu S**, Cottone CM, Yoon R, Jefferson FA, Sung JM, Okhunov Z, Tapiero S, Patel RM, Landman J, Clayman RV. Endockscope: A Disruptive Endoscopic Technology. *J Endourol* 2019; **33**: 960-965 [PMID: 31195831 DOI: 10.1089/end.2019.0252]

18 **Wang KW**, Dong M. Potential applications of artificial intelligence in colorectal polyps and cancer: Recent advances and prospects. *World J Gastroenterol* 2020; **26**: 5090-5100 [PMID: 32982111 DOI: 10.3748/wjg.v26.i34.5090]

19 **Yasuda Y**, Tokunaga K, Koga T, Sakamoto C, Goldberg IG, Saitoh N, Nakao M. Computational analysis of morphological and molecular features in gastric cancer tissues. *Cancer Med* 2020; **9**: 2223-2234 [PMID: 32012497 DOI: 10.1002/cam4.2885]

20 **Yao Y**, Gou S, Tian R, Zhang X, He S. Automated Classification and Segmentation in Colorectal Images Based on Self-Paced Transfer Network. *Biomed Res Int* 2021; **2021**: 6683931 [PMID: 33542924 DOI: 10.1155/2021/6683931]

21 **Baig R**, Bibi M, Hamid A, Kausar S, Khalid S. Deep Learning Approaches Towards Skin Lesion Segmentation and Classification from Dermoscopic Images - A Review. *Curr Med Imaging* 2020; **16**: 513-533 [PMID: 32484086 DOI: 10.2174/1573405615666190129120449]

22 **Zhang K**, Liu X, Liu F, He L, Zhang L, Yang Y, Li W, Wang S, Liu L, Liu Z, Wu X, Lin H. An Interpretable and Expandable Deep Learning Diagnostic System for Multiple Ocular Diseases: Qualitative Study. *J Med Internet Res* 2018; **20**: e11144 [PMID: 30429111 DOI: 10.2196/11144]

23 **Weng S**, Xu X, Li J, Wong STC. Combining deep learning and coherent anti-Stokes Raman scattering imaging for automated differential diagnosis of lung cancer. *J Biomed Opt* 2017; **22**: 1-10 [PMID: 29086544 DOI: 10.1117/1.JBO.22.10.106017]

24 **Xu Y**, Jia Z, Wang LB, Ai Y, Zhang F, Lai M, Chang EI. Large scale tissue histopathology image classification, segmentation, and visualization *via* deep convolutional activation features. *BMC Bioinformatics* 2017; **18**: 281 [PMID: 28549410 DOI: 10.1186/s12859-017-1685-x]

25 **Yonekura A**, Kawanaka H, Prasath VBS, Aronow BJ, Takase H. Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network. *Biomed Eng Lett* 2018; **8**: 321-327 [PMID: 30603216 DOI: 10.1007/s13534-018-0077-0]

26 **Horiuchi Y**, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Convolutional Neural Network for Differentiating Gastric Cancer from Gastritis Using Magnified Endoscopy with Narrow Band Imaging. *Dig Dis Sci* 2020; **65**: 1355-1363 [PMID: 31584138 DOI: 10.1007/s10620-019-05862-6]

27 **Guimarães P**, Keller A, Fehlmann T, Lammert F, Casper M. Deep-learning based detection of gastric precancerous conditions. *Gut* 2020; **69**: 4-6 [PMID: 31375599 DOI: 10.1136/gutjnl-2019-319347]

28 **Wu T**, Tegmark M. Toward an artificial intelligence physicist for unsupervised learning. *Phys Rev E* 2019; **100**: 033311 [PMID: 31639888 DOI: 10.1103/PhysRevE.100.033311]

29 **Zhang R**, Zhang Y, Li X. Unsupervised Feature Selection *via* Adaptive Graph Learning and Constraint. *IEEE Trans Neural Netw Learn Syst* 2020; **PP** [PMID: 33361001 DOI: 10.1109/TNNLS.2020.3042330]

30 **Wang Z**, Li M, Xu Z, Jiang Y, Gu H, Yu Y, Zhu H, Zhang H, Lu P, Xin J, Xu H, Liu C. Improvements to the gastric cancer tumor-node-metastasis staging system based on computer-aided unsupervised clustering. *BMC Cancer* 2018; **18**: 706 [PMID: 29970022 DOI: 10.1186/s12885-018-4623-z]

31 **Luo H**, Xu G, Li C, He L, Luo L, Wang Z, Jing B, Deng Y, Jin Y, Li Y, Li B, Tan W, He C, Seeruttun SR, Wu Q, Huang J, Huang DW, Chen B, Lin SB, Chen QM, Yuan CM, Chen HX, Pu HY, Zhou F, He Y, Xu RH. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol* 2019; **20**: 1645-1654 [PMID: 31591062 DOI: 10.1016/S1470-2045(19)30637-0]

32 **Lee JH**, Kim YJ, Kim YW, Park S, Choi YI, Park DK, Kim KG, Chung JW. Spotting malignancies from gastric endoscopic images using deep learning. *Surg Endosc* 2019; **33**: 3790-3797 [PMID: 30719560 DOI: 10.1007/s00464-019-06677-2]

33 **Hirasawa T**, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J, Tada T. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018; **21**: 653-660 [PMID: 29335825 DOI: 10.1007/s10120-018-0793-2]

34 **Cho BJ**, Bang CS, Park SW, Yang YJ, Seo SI, Lim H, Shin WG, Hong JT, Yoo YT, Hong SH, Choi JH, Lee JJ, Baik GH. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019; **51**: 1121-1129 [PMID: 31443108 DOI: 10.1055/a-0981-6133]

35 **Ueyama H**, Kato Y, Akazawa Y, Yatagai N, Komori H, Takeda T, Matsumoto K, Ueda K, Hojo M, Yao T, Nagahara A, Tada T. Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. *J Gastroenterol Hepatol* 2021; **36**: 482-489 [PMID: 32681536 DOI: 10.1111/jgh.15190]

36 **Li L**, Chen Y, Shen Z, Zhang X, Sang J, Ding Y, Yang X, Li J, Chen M, Jin C, Chen C, Yu C. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. *Gastric Cancer* 2020; **23**: 126-132 [PMID: 31332619 DOI: 10.1007/s10120-019-00992-2]

37 **Bergholt MS**, Zheng W, Lin K, Ho KY, Teh M, Yeoh KG, Yan So JB, Huang Z. In vivo diagnosis of gastric cancer using Raman endoscopy and ant colony optimization techniques. *Int J Cancer* 2011; **128**: 2673-2680 [PMID: 20726002 DOI: 10.1002/ijc.25618]

38 **Liang Q**, Nan Y, Coppola G, Zou K, Sun W, Zhang D, Wang Y, Yu G. Weakly Supervised Biomedical Image Segmentation by Reiterative Learning. *IEEE J Biomed Health Inform* 2019; **23**: 1205-1214 [PMID: 29994489 DOI: 10.1109/JBHI.2018.2850040]

39 **El-Khatib H**, Popescu D, Ichim L. Deep Learning-Based Methods for Automatic Diagnosis of Skin Lesions. *Sensors (Basel)* 2020; **20** [PMID: 32245258 DOI: 10.3390/s20061753]

40 **Pereira S**, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans Med Imaging* 2016; **35**: 1240-1251 [PMID: 26960222 DOI: 10.1109/TMI.2016.2538465]

41 **Polanski WH**, Zolal A, Sitoci-Ficici KH, Hiepe P, Schackert G, Sobottka SB. Comparison of Automatic Segmentation Algorithms for the Subthalamic Nucleus. *Stereotact Funct Neurosurg* 2020; **98**: 256-262 [PMID: 32369819 DOI: 10.1159/000507028]

42 **Liu T**, Liu J, Ma Y, He J, Han J, Ding X, Chen CT. Spatial feature fusion convolutional network for liver and liver tumor segmentation from CT images. *Med Phys* 2021; **48**: 264-272 [PMID: 33159809 DOI: 10.1002/mp.14585]

43 **Cuocolo R**, Caruso M, Perillo T, Ugga L, Petretta M. Machine Learning in oncology: A clinical appraisal. *Cancer Lett* 2020; **481**: 55-62 [PMID: 32251707 DOI: 10.1016/j.canlet.2020.03.032]

44 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

45 **Zhou T**, Han G, Li BN, Lin Z, Ciaccio EJ, Green PH, Qin J. Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method. *Comput Biol Med* 2017; **85**: 1-6 [PMID: 28412572 DOI: 10.1016/j.compbiomed.2017.03.031]

46 **Tan JW**, Wang L, Chen Y, Xi W, Ji J, Xu X, Zou LK, Feng JX, Zhang J, Zhang H. Predicting Chemotherapeutic Response for Far-advanced Gastric Cancer by Radiomics with Deep Learning Semi-automatic Segmentation. *J Cancer* 2020; **11**: 7224-7236 [PMID: 33193886 DOI: 10.7150/jca.46704]

47 **Cho BJ**, Bang CS, Lee JJ, Seo CW, Kim JH. Prediction of Submucosal Invasion for Gastric Neoplasms in Endoscopic Images Using Deep-Learning. *J Clin Med* 2020; **9** [PMID: 32549190 DOI: 10.3390/jcm9061858]

48 **Krizhevsky A**, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems 25 (NIPS 2012). Red Hook: Curran Associates, 2012

49 **Karpathy A**, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L. Large-Scale Video Classification with Convolutional Neural Networks. In: 2014 IEEE Conference on Computer Vision

and Pattern Recognition; 2014 June 23-28; Columbus, USA. IEEE, 2014 [DOI: 10.1109/CVPR.2014.223]

50 **Xiang Y**, Lin Z, Meng J. Automatic QRS complex detection using two-level convolutional neural network. *Biomed Eng Online* 2018; **17**: 13 [PMID: 29378580 DOI: 10.1186/s12938-018-0441-4]

51 **Shelhamer E**, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans Pattern Anal Mach Intell* 2016; **39**: 1 [DOI: 10.1109/tpami.2016.2572683]

52 **Kim HK**, Yoo KY, Park JH, Jung HY. Asymmetric Encoder-Decoder Structured FCN Based LiDAR to Color Image Generation. *Sensors (Basel)* 2019; **19** [PMID: 31694330 DOI: 10.3390/s19214818]

53 **Zhu H**, Adeli E, Shi F, Shen D; Alzheimer's Disease Neuroimaging Initiative. FCN Based Label Correction for Multi-Atlas Guided Organ Segmentation. *Neuroinformatics* 2020; **18**: 319-331 [PMID: 31898145 DOI: 10.1007/s12021-019-09448-5]

54 **Guo X**, Nie R, Cao J, Zhou D, Qian W. Fully Convolutional Network-Based Multifocus Image Fusion. *Neural Comput* 2018; **30**: 1775-1800 [PMID: 29894654 DOI: 10.1162/neco_a_01098]

55 **Zhou X**, Takayama R, Wang S, Hara T, Fujita H. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. *Med Phys* 2017; **44**: 5221-5233 [PMID: 28730602 DOI: 10.1002/mp.12480]

56 **Wang R**, Cao S, Ma K, Zheng Y, Meng D. Pairwise learning for medical image segmentation. *Med Image Anal* 2021; **67**: 101876 [PMID: 33197863 DOI: 10.1016/j.media.2020.101876]

57 **Xue J**, He K, Nie D, Adeli E, Shi Z, Lee SW, Zheng Y, Liu X, Li D, Shen D. Cascaded MultiTask 3-D Fully Convolutional Networks for Pancreas Segmentation. *IEEE Trans Cybern* 2021; **51**: 2153-2165 [PMID: 31869812 DOI: 10.1109/TCYB.2019.2955178]

58 **Wu W**, Wu S, Zhou Z, Zhang R, Zhang Y. 3D Liver Tumor Segmentation in CT Images Using Improved Fuzzy *C*-Means and Graph Cuts. *Biomed Res Int* 2017; **2017**: 5207685 [PMID: 29090220 DOI: 10.1155/2017/5207685]

59 **Baazaoui A**, Barhoumi W, Zagrouba E. Semi-Automated Segmentation of Single and Multiple Tumors in Liver CT Images Using Entropy-Based Fuzzy Region Growing. *IRBM* 2017; **38**: 98-108 [DOI: 10.1016/j.irbm.2017.02.003]

60 **Öztürk Ş**, Özkaya U. Skin Lesion Segmentation with Improved Convolutional Neural Network. *J Digit Imaging* 2020; **33**: 958-970 [PMID: 32378058 DOI: 10.1007/s10278-020-00343-z]

61 **Oda M**, Tanaka K, Takabatake H, Mori M, Natori H, Mori K. Realistic endoscopic image generation method using virtual-to-real image-domain translation. *Healthc Technol Lett* 2019; **6**: 214-219 [PMID: 32038860 DOI: 10.1049/htl.2019.0071]

62 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929. e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

63 **Sun M**, Zhang G, Dang H, Qi X, Zhou X, Chang Q. Accurate Gastric Cancer Segmentation in Digital Pathology Images Using Deformable Convolution and Multi-Scale Embedding Networks. *IEEE Access* 2019; **7**: 75530-75541 [DOI: 10.1109/ACCESS.2019.2918800]

64 **Nguyen DT**, Lee MB, Pham TD, Batchuluun G, Arsalan M, Park KR. Enhanced Image-Based Endoscopic Pathological Site Classification Using an Ensemble of Deep Learning Models. *Sensors (Basel)* 2020; **20** [PMID: 33105736 DOI: 10.3390/s20215982]

65 **Teramoto A**, Tsukamoto T, Yamada A, Kiriyama Y, Imaizumi K, Saito K, Fujita H. Deep learning approach to classification of lung cytological images: Two-step training using actual and synthesized images by progressive growing of generative adversarial networks. *PLoS One* 2020; **15**: e0229951 [PMID: 32134949 DOI: 10.1371/journal.pone.0229951]

66 **Goodfellow IJ**, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. Montreal: MIT Press, 2014: 2672-2680

67 **Enokiya Y**, Iwamoto Y, Chen YW, Han XH. Automatic Liver Segmentation Using U-Net with Wasserstein GANs. *Int J Image Graph* 2019; **7**: 94-101 [DOI: 10.18178]

68 **Luc P**, Couprie C, Chintala S, Verbeek J. Semantic segmentation using adversarial networks. 2016 Preprint. Available from: arXiv:1611.08408

69 **Borji A**. Pros and cons of GAN evaluation measures. *CVIU* 2019; **179**: 41-65 [DOI: 10.1016/j.cviu.2018.10.009]

70 **Metz L**, Poole B, Pfau D, Sohl-Dickstein J. Unrolled generative adversarial networks. 2016 Preprint. Available from: arXiv:1611.02163

71 **Poorneshwaran JM**, Santhosh Kumar S, Ram K, Joseph J, Sivaprakasam M. Polyp Segmentation using Generative Adversarial Network. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; **2019**: 7201-7204 [PMID: 31947496 DOI: 10.1109/EMBC.2019.8857958]

72 **Tang C**, Zhang W, Wang L, Cai A, Liang N, Li L, Yan B. Generative adversarial network-based sinogram super-resolution for computed tomography imaging. *Phys Med Biol* 2020; **65**: 235006 [PMID: 33053522 DOI: 10.1088/1361-6560/abc12f]

73 **Rezaei M**, Näppi JJ, Lippert C, Meinel C, Yoshida H. Generative multi-adversarial network for striking the right balance in abdominal image segmentation. *Int J Comput Assist Radiol Surg* 2020; **15**: 1847-1858 [PMID: 32897490 DOI: 10.1007/s11548-020-02254-4]

74 **Zhang T**, Cheng J, Fu H, Gu Z, Xiao Y, Zhou K, Gao S, Zheng R, Liu J. Noise Adaptation Generative Adversarial Network for Medical Image Analysis. *IEEE Trans Med Imaging* 2020; **39**: 1149-1159 [PMID: 31567075 DOI: 10.1109/TMI.2019.2944488]

75 **Han L**, Huang Y, Dou H, Wang S, Ahamad S, Luo H, Liu Q, Fan J, Zhang J. Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network. *Comput Methods Programs Biomed* 2020; **189**: 105275 [PMID: 31978805 DOI: 10.1016/j.cmpb.2019.105275]

76 **Almalioglu Y**, Bengisu Ozyoruk K, Gokce A, Incetan K, Irem Gokceler G, Ali Simsek M, Ararat K, Chen RJ, Durr NJ, Mahmood F, Turan M. EndoL2H: Deep Super-Resolution for Capsule Endoscopy. *IEEE Trans Med Imaging* 2020; **39**: 4297-4309 [PMID: 32795966 DOI: 10.1109/TMI.2020.3016744]

77 **Wang D**. Research on Key Technologies of Medical Image Classification Based on Unsupervised and Semi-supervised Framework. M.D. Thesis, Jilin University. 2020 Available from: http://cdmd.cnki.com.cn/Article/CDMD-10183-1020754186.htm

78 **Klages P**, Benslimane I, Riyahi S, Jiang J, Hunt M, Deasy JO, Veeraraghavan H, Tyagi N. Patch-based generative adversarial neural network models for head and neck MR-only planning. *Med Phys* 2020; **47**: 626-642 [PMID: 31733164 DOI: 10.1002/mp.13927]

79 **Tong N**, Gou S, Niu T, Yang S, Sheng K. Self-paced DenseNet with boundary constraint for automated multi-organ segmentation on abdominal CT images. *Phys Med Biol* 2020; **65**: 135011 [PMID: 32657281 DOI: 10.1088/1361-6560/ab9b57]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 25-35

DOI: 10.37126/aige.v2.i2.25

ISSN 2689-7164 (online)

MINIREVIEWS

Application of artificial intelligence to endoscopy on common gastrointestinal benign diseases

Hang Yang, Bing Hu

ORCID number: Hang Yang 0000-0002-7235-8162; Bing Hu 0000-0002-9898-8656.

Author contributions: All authors participated in the work; Yang H contributed to the design and draft of the manuscript; Hu B contributed to reviewing the manuscript; Yang H and Bing H contributed to revising the manuscript.

Supported by the 1·3·5 Project for Disciplines of Excellence Clinical Research Incubation Project, West China Hospital, Sichuan University, China, No. 20HXFH016.

Conflict-of-interest statement: The authors declare that they have no competing interests.

Hang Yang, Bing Hu, Department of Gastroenterology, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, China

Corresponding author: Bing Hu, MBBS, MD, Professor, Department of Gastroenterology, West China Hospital, Sichuan University, No. 37 Guo Xue Xiang, Wu Hou District, Chengdu 610041, Sichuan Province, China. hubingnj@163.com

Abstract

Artificial intelligence (AI) has been widely involved in every aspect of healthcare in the preclinical stage. In the digestive system, AI has been trained to assist auxiliary examinations including histopathology, endoscopy, ultrasonography, computerized tomography, and magnetic resonance imaging in detection, diagnosis, classification, differentiation, prognosis, and quality control. In the field of endoscopy, the application of AI, such as automatic detection, diagnosis, classification, and invasion depth, in early gastrointestinal (GI) cancers has received wide attention. There is a paucity of studies of AI application on common GI benign diseases based on endoscopy. In the review, we provide an overview of AI applications to endoscopy on common GI benign diseases including in the esophagus, stomach, intestine, and colon. It indicates that AI will gradually become an indispensable part of normal endoscopic detection and diagnosis of common GI benign diseases as clinical data, algorithms, and other related work are constantly repeated and improved.

Key Words: Artificial intelligence; Endoscopy; Common gastrointestinal benign diseases

Core Tip: In endoscopy, the application of artificial intelligence in early gastrointestinal cancer has been widely concerned. We provide a general conclusion of artificial intelligence endoscopy applications in common gastrointestinal benign diseases, such as Barrett's esophagus, atrophic gastritis, and colonic polyp. Studies indicate high accuracies and efficiencies. Further related work is needed to boost the real application of artificial intelligence in common gastrointestinal benign diseases in the future.

Citation: Yang H, Hu B. Application of artificial intelligence to endoscopy on common

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: March 5, 2021
Peer-review started: March 5, 2021
First decision: March 14, 2021
Revised: March 17, 2021
Accepted: April 20, 2021
Article in press: April 20, 2021
Published online: April 28, 2021

P-Reviewer: Azimi P
S-Editor: Wang JL
L-Editor: Filipodia
P-Editor: Wang LL

gastrointestinal benign diseases. *Artif Intell Gastrointest Endosc* 2021; 2(2): 25-35
URL: https://www.wjgnet.com/2689-7164/full/v2/i2/25.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i2.25

INTRODUCTION

Artificial intelligence (AI) is essentially a process of learning human thinking and transferring human experience based on mathematics and statistics. Iteration of algorithm, rising data, and improving computing power are cores of AI. Machine learning (ML) is a subset of AI[1], and deep learning is a subset of ML to realize ML[2], where multiple algorithms are structured together in complex layers. Artificial neural networks are one of the most common algorithms of AI[3]. Convolutional neural networks (CNNs) are a kind of supervised deep learning algorithm[4]. Its modified format is defined as deep convolutional neural networks[5]. Recognizing images based on artificial neural networks/CNNs promotes AI penetrating in medicine. Computer-aided diagnosis (CAD) systems are designed to interpret medical images using advances of AI from ML to deep learning[6].

In the field of gastroenterology, diseases of the liver, pancreases, and full digestive tract have been involved. Examples include a deep learning model based on computed tomography images to stage liver fibrosis, a deep learning model constructed to differentiate between precancerous lesions and pancreatic cancers, and a deep learning model used in endoscopy to detect early gastrointestinal (GI) cancers. A study covered five kinds of gastric diseases and showed the diagnostic specificity of the CNNs was higher than that of the endoscopists for early gastric cancer and high-grade intraepithelial neoplasia images (91.2% *vs* 86.7%). The diagnostic accuracy of the CNNs was close to those of the endoscopists for lesion-free, early gastric cancer and high-grade intraepithelial neoplasia, peptic ulcer (PU), advanced gastric cancer (GC), and gastric submucosal tumor images. The CNNs had an image recognition time of 42 s for all the test set images[7]. In this review, the application and research of AI on common GI benign lesions based on endoscopy were concluded.

LITERATURE SEARCH

This review aimed to make a qualitative only review of the application of AI on common GI benign diseases. We searched the PubMed database for articles that were published in the last 5 years using the term combinations of artificial intelligence and common GI benign lesions [Barrett's esophagus (BE), esophageal varices (EV), atrophic gastritis (AG), PU, gastric polyp, small bowel capsule endoscopy, colonic polyp/adenoma, and inflammatory bowel diseases (IBDs)]. Articles based on radiological images or other samples, review articles, research articles of early or advanced GI cancers or other cancers, and articles only related to either GI benign diseases or AI were excluded. Two authors independently extracted data. Any disagreement was resolved by discussion until consensus was reached or by consulting a third author. Endoscopic-related results were qualitatively concluded in Table 1. The flowchart was presented in Figure 1.

SEARCH RESULTS

Initially, a total of 555 articles were identified. After manually screening and reading, only research articles related to the application of AI to common GI benign lesions (BE, EV, AG, PU, gastric polyp, small bowel capsule endoscopy, colonic polyp/adenoma, and IBDs) based on different endoscopic images or tissue slides from endoscopic biopsies were included. Finally, 35 studies were tabulated in Table 1. Six studies demonstrated the application of AI on esophageal benign diseases (5 BE and 1 EV). Seven studies were about gastric benign diseases (3 AG, 3 PU, and 1 polyp). Seven studies were about intestinal diseases. Fifteen studies were about colonic benign diseases (11 polyp/adenoma and 4 IBDs).

Table 1 Application of artificial intelligence on common gastrointestinal benign diseases

Ref.	Aim and disease	Prospective/retrospective	AI method	Endoscopy image	Training dataset	Validation dataset	Result sensitivity	Result specificity	Result accuracy/AUC
Esophageal benign diseases									
de Groof *et al*[12]	Detecting Barrett'sneoplasia	Retrospective	CAD	WLI images	40 images	A leave one out cross validation	92%	95%	85%[1]
Jisu *et al*[39]	Distinguishing BE	Retrospective	CNNs	Endomicroscopic images	262 images	Image distortion methods			80.77%[1]
Ebigbo *et al*[40]	Distinguishing BE	Retrospective	CNNs (ResNet)	WLI images	129 images	62 images	83.7%	100.0%	89.9%[1]
Sehgal *et al*[41]	Detecting dysplasia in BE	Retrospective	ML (decision trees)	Video recordings(AAC)	40 patients with NDBE and DBE		97%	88%	92%[1]
de Groof *et al*[14]	Detecting Barrett'sneoplasia	Retrospective	CNN (CAD (ResNet-UNet))	WLI images	494364 images	1704 images (early stage neoplasia in BE and NDBE from 669 patients)	90%	88%	89%[1]
Dong *et al*[16]	Screening high risk EV	Retrospective	ML (Random forest)		238 patients	109 patients			Training set (0.84); Validation set (0.82)
Gastric benign diseases									
Zhang *et al*[42]	Diagnosing CAG	Retrospective	CNNs (DenseNet)	WLI images	5470 images	Five-fold cross validation	94.5%	94.0%	94.2%[1]
Guimarães *et al*[43]	DiagnosingCAG	Retrospective	CNNs (VGG16)	WLI images	200 images	70 images(ten-fold cross validation)			93%[1]/0.98
Horiuchi *et al*[44]	Differentiating CAG	Retrospective	CNNs (GoogLeNet)	ME-NBI images	1078 images	107 images	95.4%	71.0%	85.3%[1]/0.85
Zhang *et al*[7]	Diagnosing PU	Retrospective	CNNs (ResNet34)	WLI images	4200 images	228 images	78.9%	88.4%	86.4%[1]
Lee *et al*[45]	Differentiating PU	Retrospective	CNNs (ResNet-50/ Inception v3/VGG16 model)	WLI images	200 images	20 images			92.6%[1]/85.24%[1] /91.2%[1]
Namikawa *et al*[46]	Classifying gastriccancers and ulcers	Retrospective	CNNs (SSD)	WLI/NBI/chromoendoscopy images	373 images	720 images	93.3%	99.0%	93.3 %[1]
Zhang *et al*[26]	Detecting GP	Retrospective	CNNs (SSD-GPNet)	WLI images	404 images	50 images			93.92%[1]
Intestinal benign diseases									

Hwang *et al* [29]	Classifying hemorrhagic and ulcerations	Retrospective	CNNs (VGGNet)	Capsule endoscopy	7556 images	5760 images	Model 1 *vs* Model 2; 97.61% *vs* 95.07%	Model 1 *vs* Model 2; 96.04% *vs* 98.18%	Model 1 *vs* Model 2; 96.83%[1] *vs* 96.62%[1]
Aoki *et al*[47]	Detecting erosions and ulcerations	Retrospective	CNNs (SSD)	Capsule endoscopy	5360 images	10440 images	88.2%	90.9%	90.8%[1]/0.958
Aoki *et al*[48]	Detecting erosions and ulcerations	Retrospective	CNNs (SSD)	Capsule endoscopy		20 videos			
Ding *et al*[49]	Detecting small bowel diseases	Retrospective	CNNs (ResNet)	Capsule endoscopy	158235 images	5000 patients	99.88% per patient99.90% per lesion	100% per patient100 % per lesion	
Fan *et al*[50]	Detecting erosions and ulcerations	Retrospective	CNNs (AlexNet)	Capsule endoscopy	Ulcer 2000; Erosion 2720	Ulcer 500; Erosion 690	Ulcer: 96.80%; Erosion: 93.67%	Ulcer: 94.79%; Erosion: 95.98%	Ulcer: 95.16%[1]; Erosion: 95.34%[1]/0.98
Leenhardt *et al* [51]	Detecting small bowel angiectasia	Retrospective	CNNs	Capsule endoscopy	300 videos with angiectasia	300 videos with angiectasia	100%	96%	
Tsuboi *et al*[52]	Detecting small bowel angiectasia	Retrospective	CNNs (SSD)	Capsule endoscopy	141 patients	28 patients	98.8%	98.4%	0.998
Colonic benign diseases									
Lui *et al*[34]	Detecting missed colonic lesions	Retrospective and prospective	R-FCN (ResNet101)	Endoscopic videos (WLI)	52 videos	Real-time AI detected at least 1 missed adenoma in 14 patients (26.9%) and increased the total number of adenomas detected by 23.6%.			
Rodriguez-Diaz *et al*[53]	Histologically classifying CP	Retrospective	CAD	NBI	745 images +65000 images		96%	84%	
Komeda *et al* [54]	Diagnosing CP	Retrospective	CNNs-CAD	WLI/NBI/ chromoendoscopy images	1200 images	10-fold cross validation			75.1%[1]
Akbari *et al*[55]	Classifying CP	Retrospective	FCNs	WLI images	200 images	300 images			
Chen *et al*[56]	Classifying diminutive CP	Retrospective	DCNNs-CAD	NBI images	96 images + 188 images		96.3%	78.1%	
Gong *et al*[57]	Detecting CA	Prospective	DCNNs	WLI images	DCNNs system (n = 355) or unassisted (control) colonoscopy (n = 349)		58 (16%) of 35527 (8%) of 349		
Byrne *et al*[58]	Differentiating adenomatous and hyperplastic polyps	Retrospective	DCNNs	Videos and NBI images	223 polyp videos	40 videos	98%	83%	
Mori *et al*[59]	Identifying diminutive CP	Prospective	CAD	NBI/stained images	791 consecutive patients undergoing colonoscopy and 23 endoscopists				Pathologic prediction rate of 98.1%[1]
Misawa *et al*[60]	DetectingCP	Retrospective	CAD	WLI images	105 positive and 306 negative	50 positive and 85 negative videos	90.0%	63.3%	76.5%[1]

					videos				
Taunk *et al*[61]	Classifying polyp histology	Retrospective	CAD	pCLE images	125 images	189 images	95%	94%	94%[1]
Wang *et al*[62]	Detecting CA	Prospective	CAD	WLI images	484 patients in the CADe group and 478 in the sham group		165 (34%) of 484; 132 (28%) of 478		
Tong *et al*[63]	Differentiating UC, CD, and ITB	Retrospective	CNNs/RF	WLI images	6399 consecutive patients (5128 UC, 875 CD and 396 ITB)		RF (UC 97%, CD 65%, and ITB 68%); CNN (UC 99%, CD 87%, and ITB 52%)	RF (UC 97%, CD 53%, and ITB 76%); CNN (UC 97%, CD 83%, and ITB 81%)	RF (UC 0.97, CD 0.58, and ITB 0.72); CNN (UC 0.98, CD 0.85, and ITB 0.63)
Ozawa *et al* [36]	Diagnosing UC	Retrospective	CAD	WLI images	26304 images	3981 images			0.86 (Mayo 0); 0.98 (Mayo 0–1)
Stidham *et al*[37]	Grading the severity of ulcerative colitis	Retrospective	CNNs	WLI images	2465 patients	308 patients	83.0%	96.0%	0.966
Maeda *et al*[38]	Identifying histologic inflammation associated with UC	Retrospective	CAD	Endocytoscopic images	87 patients	100 patients	74%	97%	91%[1]

[1]Results accuracy. AAC: Acetic acid chromoendoscopy; AI: Artificial intelligence; AUC: Area under the curve; BE: Barrett's esophagus; CA: Colorectal adenomas; CAD: Computer-aided diagnosis; CAG: Chronic atrophic gastritis; CD: Crohn's disease; CNN: Convolutional neural network; CP: Colorectal polyp; DBE: Dysplastic Barrett's esophagus; DCNNs: Deep convolutional neural networks; EV: Esophageal Varices; FCNs: Fully convolutional networks; GP: Gastric polyp; ITB: Intestinal tuberculosis; ME-NBI: Magnifying narrow-band imaging; ML: Machine learning; NBI: Narrow-band imaging; NDBE: Non-dysplastic Barrett's esophagus; pCLE: Probe-based confocal laser endomicroscopy; PU: Peptic ulcer; RF: Random forest; R-FCNs: Region-based fully connected convolutional neural networks; SSD: Single shot detector; UC: Ulcerative colitis; WLI: White-light imaging.

AI AND ESOPHAGEAL BENIGN DISEASES: BARRETT'S ESOPHAGUS AND ESOPHAGEAL VARICES

BE is a precursor to esophageal adenocarcinoma. Intestinal metaplasia and gastric metaplasia are two pathological subclasses of BE. Intestinal metaplasia can progress to esophageal cancer. The ablation of dysplastic BE will reduce the risk of progression to cancer[8]. Endoscopic surveillance, including white-light imaging (WLI), narrow-band imaging, and chromoendoscopy, is performed to detect dysplasia in BE. Approximately 5% of the esophageal mucosa is found at risk by random biopsies sample[9].

Recently, AI has been applied in some studies of BE. For example, CAD based on deep learning and different algorithms trained by WLI and endomicroscopic images to detect, diagnose, and distinguish BE with achievable results (the accuracy from 80.77% to 92%, specificity from 88% to 100%, and sensitivity from 83.7% to 97%) (Table 1). On pathology, CAD with wide area transepithelial sampling could increase the detection of high-grade dysplasia/esophageal adenocarcinoma (absolute increase: 14.4%)[10]. Deep convolutional neural networks were used in the whole-slide tissue histopathology images-based diagnosis of dysplastic and non-dysplastic BE[11]. Moreover, distinguishing BE adenocarcinoma by AI methods has been studied based on different endoscopic images such as WLI and volumetric laser endomicroscopic

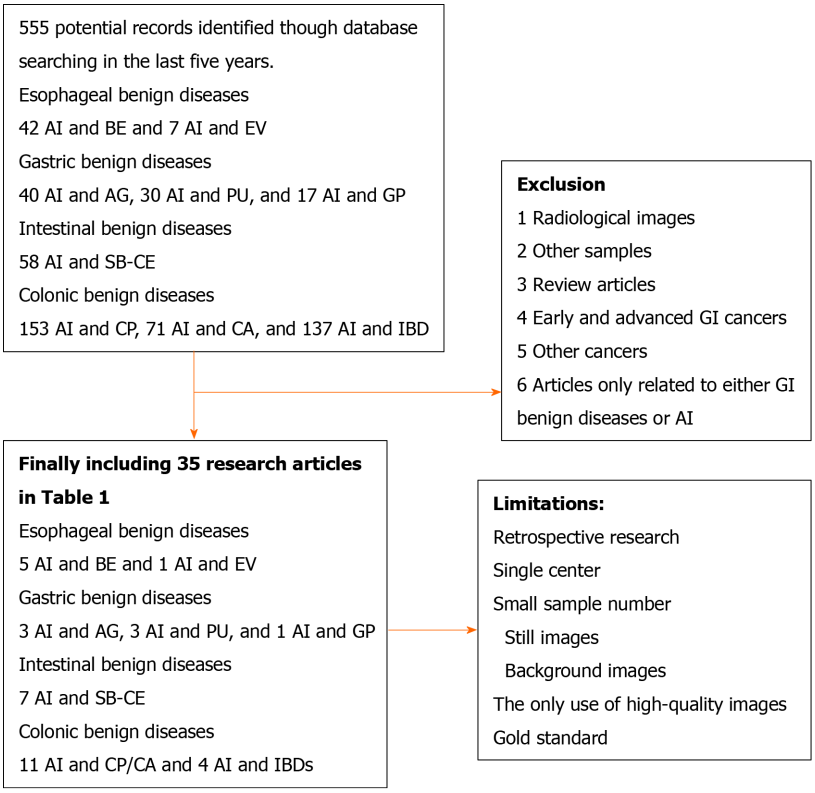

Figure 1 Flow chart of study selection and logic arrangement of review. AG: Atrophic gastritis; AI: Artificial intelligence; BE: Barrett's esophagus; CA: Colonic adenoma; CP: Colonic polyp; EV: Esophageal varices; GI: Gastrointestinal; GP: Gastric polyp; IBDs: Inflammatory bowel diseases; PU: Peptic ulcer; SB-CE: Small bowel capsule endoscopy.

images with accuracy from 88% to 92%, specificity from 88% to 93%, and sensitivity from 90% to 95%[12-14].

As another common esophageal benign disease, EV are associated with cirrhosis and portal hypertension, and variceal hemorrhage is a substantial cause of mortality[15]. However, related AI research is limited. A score system based on ML was built on the data of 238 patients with cirrhosis to reliably identify patients with varices that needed treatments and achieved an area under the curve (AUC) from 0.75 to 0.84 in different groups[16]. Another study of the index of spleen volume-to-platelet ratio based on deep learning-measured spleen volume on computed tomography to assess high-risk varices in B-viral compensated cirrhosis had a sensitivity of 69.4% and specificity of 78.5%[17]. There is little research of AI on esophagitis, although it is also a common esophageal disease associated with BE and esophageal cancer.

AI AND GASTRIC BENIGN LESIONS: ATROPHIC GASTRITIS, PEPTIC ULCER, AND POLYP

Gastritis, peptic ulcer, polyp and adenoma, and vascular lesion are common gastric benign diseases. The detection and diagnosis of these lesions account for a large part of daily endoscopic work. If AI can be applied in this field, then the rate of detection and accuracy will be improved. Moreover, the rapid identification of simple lesions can fill the lack of endoscopists and reduce the workload.

Early diagnosis of chronic AG, a precancerous lesion, is important to prevent the occurrence and development of GC. AI-assisted detection and diagnosis has been related to endoscopic images (Table 1), histological images[18,19], and X-ray images[20,21]. The accuracy was from 85.3% to 94.2%, the specificity was from 71% to 94%, and the sensitivity was from 94.5% to 95.4%. *Helicobacter pylori* infection, as a dominant cause of chronic AG and GC, has also been detected *via* AI methods based on endoscopic images, such as CNNs (GoogLeNet) and CNNs (ResNet-50 model), which achieved an accuracy up to 93.8% in a considerably short time of less than 200 s[22-24].

A CNN method was constructed to diagnose PU and differentiate GC from PU mainly based on WLI, narrow-band imaging, and chromoendoscopic images with an accuracy from 85.2% to 93.3%, specificity from 88.4% to 99%, and sensitivity from 78.9% to 93.3% (Table 1). In addition, a ML model was built on six parameters, such as age and the presence of PU, to predict recurrent ulcer bleeding within 1 year with an AUC of 0.775 and an accuracy of 84.3%[25].

There were only a few applications of AI on detecting gastric hyperplastic polyps and adenomas. A 93.92% accuracy was achieved when detecting polyps by CNNs (SSD-GPNet) based on WLI images[26]. A CNN method was trained to detect adenomas and showed an AUC of 0.99 based on histopathology whole-slide images[27]. Research and application of AI on gastric benign lesions are limited, although these diseases make up a considerable part of daily work. Some of them are usually prone to severe outcomes and risks despite the relative ease to diagnose. Indeed, the study of AI on this aspect will assist endoscopists to improve early detection rates and bring the opportunity of early treatment to benefit patients.

AI AND INTESTINAL DISEASES: CAPSULE ENDOSCOPY

The application of AI in small bowel diseases has been concentrated on capsule endoscopy. It includes image enhancement using ML algorithms to reduce artifact interference as well as three-dimensional luminal map reconstruction and localization[28]. AI-assisted capsule endoscopy in detecting ulcer, erosion, bleeding, polyps, parasite, diverticulum, and angiectasia with an accuracy more than 90.0%, specificity from 90.9% to 100%, and sensitivity from 88.2% to 100% in a short time (about 6 min) (Table 1). Furthermore, a gradient class activation map was used to visualize and detect lesions by CNNs-VGGNet to improve the classification and localization[29]. In addition, a CNN method based on conventional abdominal radiographs was trained to detect high-grade small bowel obstruction with an AUC of 0.84, a sensitivity of 83.8%, and a specificity of 68.1%[30]. In another study, it achieved an AUC of 0.971, a sensitivity of 91.4%, and a specificity of 91.9% using region-based CNNs[31]. The limited research indicated CNNs could recognize specific images among a large variety with high efficiency and accuracy. The application of AI will relieve the clinical workload as capsule endoscopy reading is a time-consuming process.

AI AND COLONIC BENIGN LESIONS: POLYP, ADENOMA, AND IBDS

A 1.0% increase of adenoma detection rate has been associated with a 3.0% decrease in the risk of interval colorectal cancer[32]. To improve colorectal polyp and adenoma detection, AI has been widely applied in the detection, real-time histological classification, segmentation, localization, and distinguishing of diminutive polyps and adenomas based on different methods trained by videos and images in retrospective or prospective and in multicenter or single center clinical trials (Table 1). Deep learning was also used to automatically classify colorectal polyps on histopathologic slides[33]. For the internal evaluation, the accuracy of the deep CNN method was 93.5%, which was comparable to the pathologists accuracy of 91.4%. On the external test, it achieved an accuracy of 87.0%, which was comparable to the pathologists accuracy of 86.6%. The application of AI in colorectal polyps has gained more concerns and practice, and it is deeper and closer to the clinical use to further increase the detection rate of polyps. For example, real-time AI detected at least one missed adenoma in 14 patients (26.9%) and increased the total number of adenomas detected by 23.6%[34].

AI methods have been trained in grading endoscopic disease severity of patients with ulcerative colitis and in predicting remission in patients with moderate to severe Crohn's disease[35]. For example, a CNN-CAD system based on GoogLeNet was robustly promising to identify normal mucosa (Mayo 0) and mucosal healing state with an accuracy of 0.86 of Mayo 0 and of 0.98 of Mayo 0-1[36]. Another similar system could differentiate remission (Mayo 0 or 1) from moderate or severe disease (Mayo 2 or 3) with an AUC of 0.966, a specificity of 96.0%, and a sensitivity of 83.0%[37]. A CAD was constructed to identify the presence of histologic inflammation associated with ulcerative colitis using endocytoscopy with an accuracy of 91%, a specificity of 97%, and a sensitivity 74%[38] (Table 1).

FUTURE PERSPECTIVES OF AI APPLICATION ON COMMON GI BENIGN LESIONS

We summarized the application and research of AI on common GI benign diseases. Limited studies are promising as most of the studies showed comparatively high accuracies and efficiencies. As studies of AI application on gastroenterology continue to increase, there are several areas of interest that will hold significant value in the future. First, the technical integration of AI systems will be important to optimize clinical workflow. New AI applications can easily "read in" data from a video input, allowing the systems to use the data for training and real time decision support. Second, AI systems will continue to expand the clinical applications. Some promising studies have demonstrated how AI can improve our performance on clinical tasks such as polyp identification, detection of small bowel bleeding, and endoscopic recognition of *Helicobacter pylori* infection, *etc*. More research, especially randomized controlled trials, on how to train and validate up-to-date algorithms will be continued on the present work to find more precise methods and identify new clinical tasks after practice. Third, further research will be needed to describe the most effective training methods for physician practices beginning to adopt AI technology because AI will be an indispensable helper of normal endoscopic detection and diagnosis of common GI benign lesions in the future.

CONCLUSION

Although AI is a relatively new technology, it has the potential to ease the daily workload of radiologists, pathologists, and sonographers. In endoscopy, AI related to early GI cancers and precancerous lesions has garnered more research than common GI benign diseases, despite the latter occupying a large proportion of daily work and being easier to detect and diagnose than early cancers. If models and diagnosing routes based on AI targeted at common GI benign diseases are well developed, then it will bring great benefits to patients and endoscopists, especially in primary hospitals where medical resources are lacking and core work is mainly focused on early diagnosis and treatment of common GI benign diseases. Furthermore, AI methods and technology targeted at common benign diseases will be easier for endoscopists to adopt professional education. More research is needed to overcome the challenges of integrating AI into the detection of common GI benign diseases by endoscopy, but the future is promising.

REFERENCES

1 **Bi Q**, Goodman KE, Kaminsky J, Lessler J. What is Machine Learning? *Am J Epidemiol* 2019; **188**: 2222-2239 [PMID: 31509183 DOI: 10.1093/aje/kwz189]

2 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

3 **Zador AM**. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat Commun* 2019; **10**: 3770 [PMID: 31434893 DOI: 10.1038/s41467-019-11786-6]

4 **Mieloszyk RJ**, Bhargava P. Convolutional Neural Networks: The Possibilities are Almost Endless. *Curr Probl Diagn Radiol* 2018; **47**: 129-130 [PMID: 29477265 DOI: 10.1067/j.cpradiol.2018.01.008]

5 **O'Toole AJ**, Castillo CD, Parde CJ, Hill MQ, Chellappa R. Face Space Representations in Deep Convolutional Neural Networks. *Trends Cogn Sci* 2018; **22**: 794-809 [PMID: 30097304 DOI: 10.1016/j.tics.2018.06.006]

6 **Ahmad OF**, Soares AS, Mazomenos E, Brandao P, Vega R, Seward E, Stoyanov D, Chand M, Lovat LB. Artificial intelligence and computer-aided diagnosis in colonoscopy: current evidence and future directions. *Lancet Gastroenterol Hepatol* 2019; **4**: 71-80 [PMID: 30527583 DOI: 10.1016/S2468-1253(18)30282-6]

7 **Zhang L**, Zhang Y, Wang L, Wang J, Liu Y. Diagnosis of gastric lesions through a deep convolutional neural network. *Dig Endosc* 2020 [PMID: 32961597 DOI: 10.1111/den.13844]

8 **Soh YSA**, Lee YY, Gotoda T, Sharma P, Ho KY; Asian Barrett's Consortium. Challenges to diagnostic standardization of Barrett's esophagus in Asia. *Dig Endosc* 2019; **31**: 609-618 [PMID: 30892742 DOI: 10.1111/den.13402]

9 **ASGE Technology Committee**. , Thosani N, Abu Dayyeh BK, Sharma P, Aslanian HR, Enestvedt BK, Komanduri S, Manfredi M, Navaneethan U, Maple JT, Pannala R, Parsi MA, Smith ZL, Sullivan SA, Banerjee S. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE Preservation and Incorporation of Valuable Endoscopic Innovations thresholds for adopting

real-time imaging-assisted endoscopic targeted biopsy during endoscopic surveillance of Barrett's esophagus. *Gastrointest Endosc* 2016; **83**: 684-98. e7 [PMID: 26874597 DOI: 10.1016/j.gie.2016.01.007]

10 **Vennalaganti PR**, Kaul V, Wang KK, Falk GW, Shaheen NJ, Infantolino A, Johnson DA, Eisen G, Gerson LB, Smith MS, Iyer PG, Lightdale CJ, Schnoll-Sussman F, Gupta N, Gross SA, Abrams J, Haber GB, Chuttani R, Pleskow DK, Kothari S, Goldblum JR, Zhang Y, Sharma P. Increased detection of Barrett's esophagus-associated neoplasia using wide-area trans-epithelial sampling: a multicenter, prospective, randomized trial. *Gastrointest Endosc* 2018; **87**: 348-355 [PMID: 28757316 DOI: 10.1016/j.gie.2017.07.039]

11 **Sali R**, Moradinasab N, Guleria S, Ehsan L, Fernandes P, Shah TU, Syed S, Brown DE. Deep Learning for Whole-Slide Tissue Histopathology Classification: A Comparative Study in the Identification of Dysplastic and Non-Dysplastic Barrett's Esophagus. *J Pers Med* 2020; **10** [PMID: 32977465 DOI: 10.3390/jpm10040141]

12 **de Groof J**, van der Sommen F, van der Putten J, Struyvenberg MR, Zinger S, Curvers WL, Pech O, Meining A, Neuhaus H, Bisschops R, Schoon EJ, de With PH, Bergman JJ. The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019; **7**: 538-547 [PMID: 31065371 DOI: 10.1177/2050640619837443]

13 **Swager AF**, van der Sommen F, Klomp SR, Zinger S, Meijer SL, Schoon EJ, Bergman JJGHM, de With PH, Curvers WL. Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy. *Gastrointest Endosc* 2017; **86**: 839-846 [PMID: 28322771 DOI: 10.1016/j.gie.2017.03.011]

14 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929. e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

15 **Kovacs TOG**, Jensen DM. Varices: Esophageal, Gastric, and Rectal. *Clin Liver Dis* 2019; **23**: 625-642 [PMID: 31563215 DOI: 10.1016/j.cld.2019.07.005]

16 **Dong TS**, Kalani A, Aby ES, Le L, Luu K, Hauer M, Kamath R, Lindor KD, Tabibian JH. Machine Learning-based Development and Validation of a Scoring System for Screening High-Risk Esophageal Varices. *Clin Gastroenterol Hepatol* 2019; **17**: 1894-1901. e1 [PMID: 30708109 DOI: 10.1016/j.cgh.2019.01.025]

17 **Lee CM**, Lee SS, Choi WM, Kim KM, Sung YS, Lee S, Lee SJ, Yoon JS, Suk HI. An index based on deep learning-measured spleen volume on CT for the assessment of high-risk varix in B-viral compensated cirrhosis. *Eur Radiol* 2021; 31: 3355-3365 [PMID: 33128186 DOI: 10.1007/s00330-020-07430-3]

18 **Steinbuss G**, Kriegsmann K, Kriegsmann M. Identification of Gastritis Subtypes by Convolutional Neuronal Networks on Histological Images of Antrum and Corpus Biopsies. *Int J Mol Sci* 2020; **21** [PMID: 32932860 DOI: 10.3390/ijms21186652]

19 **Martin DR**, Hanson JA, Gullapalli RR, Schultz FA, Sethi A, Clark DP. A Deep Learning Convolutional Neural Network Can Recognize Common Patterns of Injury in Gastric Pathology. *Arch Pathol Lab Med* 2020; **144**: 370-378 [PMID: 31246112 DOI: 10.5858/arpa.2019-0004-OA]

20 **Li Z**, Togo R, Ogawa T, Haseyama M. Chronic gastritis classification using gastric X-ray images with a semi-supervised learning method based on tri-training. *Med Biol Eng Comput* 2020; **58**: 1239-1250 [PMID: 32221796 DOI: 10.1007/s11517-020-02159-z]

21 **Kanai M**, Togo R, Ogawa T, Haseyama M. Chronic atrophic gastritis detection with a convolutional neural network considering stomach regions. *World J Gastroenterol* 2020; **26**: 3650-3659 [PMID: 32742133 DOI: 10.3748/wjg.v26.i25.3650]

22 **Shichijo S**, Nomura S, Aoyama K, Nishikawa Y, Miura M, Shinagawa T, Takiyama H, Tanimoto T, Ishihara S, Matsuo K, Tada T. Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images. *EBioMedicine* 2017; **25**: 106-111 [PMID: 29056541 DOI: 10.1016/j.ebiom.2017.10.014]

23 **Zheng W**, Zhang X, Kim JJ, Zhu X, Ye G, Ye B, Wang J, Luo S, Li J, Yu T, Liu J, Hu W, Si J. High Accuracy of Convolutional Neural Network for Evaluation of Helicobacter pylori Infection Based on Endoscopic Images: Preliminary Experience. *Clin Transl Gastroenterol* 2019; **10**: e00109 [PMID: 31833862 DOI: 10.14309/ctg.0000000000000109]

24 **Bang CS**, Lee JJ, Baik GH. Artificial Intelligence for the Prediction of Helicobacter Pylori Infection in Endoscopic Images: Systematic Review and Meta-Analysis Of Diagnostic Test Accuracy. *J Med Internet Res* 2020; **22**: e21983 [PMID: 32936088 DOI: 10.2196/21983]

25 **Wong GL**, Ma AJ, Deng H, Ching JY, Wong VW, Tse YK, Yip TC, Lau LH, Liu HH, Leung CM, Tsang SW, Chan CW, Lau JY, Yuen PC, Chan FK. Machine learning model to predict recurrent ulcer bleeding in patients with history of idiopathic gastroduodenal ulcer bleeding. *Aliment Pharmacol Ther* 2019; **49**: 912-918 [PMID: 30761584 DOI: 10.1111/apt.15145]

26 **Zhang X**, Chen F, Yu T, An J, Huang Z, Liu J, Hu W, Wang L, Duan H, Si J. Real-time gastric polyp detection using convolutional neural networks. *PLoS One* 2019; **14**: e0214133 [PMID: 30908513 DOI: 10.1371/journal.pone.0214133]

27 **Iizuka O**, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep Learning Models for

Histopathological Classification of Gastric and Colonic Epithelial Tumours. *Sci Rep* 2020; **10**: 1504 [PMID: 32001752 DOI: 10.1038/s41598-020-58467-9]

28 **Turan M**, Almalioglu Y, Araujo H, Konukoglu E, Sitti M. A non-rigid map fusion-based direct SLAM method for endoscopic capsule robots. *Int J Intell Robot Appl* 2017; **1**: 399-409 [PMID: 29250588 DOI: 10.1007/s41315-017-0036-4]

29 **Hwang Y**, Lee HH, Park C, Tama BA, Kim JS, Cheung DY, Chung WC, Cho YS, Lee KM, Choi MG, Lee S, Lee BI. Improved classification and localization approach to small bowel capsule endoscopy using convolutional neural network. *Dig Endosc* 2020 [PMID: 32640059 DOI: 10.1111/den.13787]

30 **Cheng PM**, Tejura TK, Tran KN, Whang G. Detection of high-grade small bowel obstruction on conventional radiography with convolutional neural networks. *Abdom Radiol (NY)* 2018; **43**: 1120-1127 [PMID: 28828625 DOI: 10.1007/s00261-017-1294-1]

31 **Cheng PM**, Tran KN, Whang G, Tejura TK. Refining Convolutional Neural Network Detection of Small-Bowel Obstruction in Conventional Radiography. *AJR Am J Roentgenol* 2019; **212**: 342-350 [PMID: 30476452 DOI: 10.2214/AJR.18.20362]

32 **Corley DA**, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 2541 [PMID: 24963577 DOI: 10.1056/NEJMc1405329]

33 **Wei JW**, Suriawinata AA, Vaickus LJ, Ren B, Liu X, Lisovsky M, Tomita N, Abdollahi B, Kim AS, Snover DC, Baron JA, Barry EL, Hassanpour S. Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathologic Slides. *JAMA Netw Open* 2020; **3**: e203398 [PMID: 32324237 DOI: 10.1001/jamanetworkopen.2020.3398]

34 **Lui TKL**, Hui CKY, Tsui VWM, Cheung KS, Ko MKL, Foo DCC, Mak LY, Yeung CK, Lui TH, Wong SY, Leung WK. New insights on missed colonic lesions during colonoscopy through artificial intelligence-assisted real-time detection (with video). *Gastrointest Endosc* 2021; **93**: 193-200. e1 [PMID: 32376335 DOI: 10.1016/j.gie.2020.04.066]

35 **Waljee AK**, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, Stidham RW, Zhu J, Higgins PDR. Development and Validation of Machine Learning Models in Prediction of Remission in Patients With Moderate to Severe Crohn Disease. *JAMA Netw Open* 2019; **2**: e193721 [PMID: 31074823 DOI: 10.1001/jamanetworkopen.2019.3721]

36 **Ozawa T**, Ishihara S, Fujishiro M, Saito H, Kumagai Y, Shichijo S, Aoyama K, Tada T. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc* 2019; **89**: 416-421. e1 [PMID: 30367878 DOI: 10.1016/j.gie.2018.10.020]

37 **Stidham RW**, Liu W, Bishu S, Rice MD, Higgins PDR, Zhu J, Nallamothu BK, Waljee AK. Performance of a Deep Learning Model *vs* Human Reviewers in Grading Endoscopic Disease Severity of Patients With Ulcerative Colitis. *JAMA Netw Open* 2019; **2**: e193963 [PMID: 31099869 DOI: 10.1001/jamanetworkopen.2019.3963]

38 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

39 **Jisu Hong**, Bo-Yong Park, Hyunjin Park. Convolutional neural network classifier for distinguishing Barrett's esophagus and neoplasia endomicroscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2017; **2017**: 2892-2895 [PMID: 29060502 DOI: 10.1109/EMBC.2017.8037461]

40 **Ebigbo A**, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA Jr, Papa J, Palm C, Messmann H. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. *Gut* 2020; **69**: 615-616 [PMID: 31541004 DOI: 10.1136/gutjnl-2019-319460]

41 **Sehgal V**, Rosenfeld A, Graham DG, Lipman G, Bisschops R, Ragunath K, Rodriguez-Justo M, Novelli M, Banks MR, Haidry RJ, Lovat LB. Machine Learning Creates a Simple Endoscopic Classification System that Improves Dysplasia Detection in Barrett's Oesophagus amongst Non-expert Endoscopists. *Gastroenterol Res Pract* 2018; **2018**: 1872437 [PMID: 30245711 DOI: 10.1155/2018/1872437]

42 **Zhang Y**, Li F, Yuan F, Zhang K, Huo L, Dong Z, Lang Y, Zhang Y, Wang M, Gao Z, Qin Z, Shen L. Diagnosing chronic atrophic gastritis by gastroscopy using artificial intelligence. *Dig Liver Dis* 2020; **52**: 566-572 [PMID: 32061504 DOI: 10.1016/j.dld.2019.12.146]

43 **Guimarães P**, Keller A, Fehlmann T, Lammert F, Casper M. Deep-learning based detection of gastric precancerous conditions. *Gut* 2020; **69**: 4-6 [PMID: 31375599 DOI: 10.1136/gutjnl-2019-319347]

44 **Horiuchi Y**, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Convolutional Neural Network for Differentiating Gastric Cancer from Gastritis Using Magnified Endoscopy with Narrow Band Imaging. *Dig Dis Sci* 2020; **65**: 1355-1363 [PMID: 31584138 DOI: 10.1007/s10620-019-05862-6]

45 **Lee JH**, Kim YJ, Kim YW, Park S, Choi YI, Park DK, Kim KG, Chung JW. Spotting malignancies from gastric endoscopic images using deep learning. *Surg Endosc* 2019; **33**: 3790-3797 [PMID: 30719560 DOI: 10.1007/s00464-019-06677-2]

46 **Namikawa K**, Hirasawa T, Nakano K, Ikenoyama Y, Ishioka M, Shiroma S, Tokai Y, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Artificial intelligence-based diagnostic system classifying gastric cancers and ulcers: comparison between the original and newly developed systems. *Endoscopy* 2020; **52**: 1077-1083 [PMID: 32503056 DOI: 10.1055/a-1194-8771]

47 **Aoki T**, Yamada A, Aoyama K, Saito H, Tsuboi A, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Tanaka S, Koike K, Tada T. Automatic detection of erosions and ulcerations in

wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest Endosc* 2019; **89**: 357-363. e2 [PMID: 30670179 DOI: 10.1016/j.gie.2018.10.027]

48 **Aoki T**, Yamada A, Aoyama K, Saito H, Fujisawa G, Odawara N, Kondo R, Tsuboi A, Ishibashi R, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Nakahori M, Tanaka S, Koike K, Tada T. Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading. *Dig Endosc* 2020; **32**: 585-591 [PMID: 31441972 DOI: 10.1111/den.13517]

49 **Ding Z**, Shi H, Zhang H, Meng L, Fan M, Han C, Zhang K, Ming F, Xie X, Liu H, Liu J, Lin R, Hou X. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* 2019; **157**: 1044-1054. e5 [PMID: 31251929 DOI: 10.1053/j.gastro.2019.06.025]

50 **Fan S**, Xu L, Fan Y, Wei K, Li L. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. *Phys Med Biol* 2018; **63**: 165001 [PMID: 30033931 DOI: 10.1088/1361-6560/aad51c]

51 **Leenhardt R**, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, Becq A, Marteau P, Histace A, Dray X; CAD-CAP Database Working Group. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc* 2019; **89**: 189-194 [PMID: 30017868 DOI: 10.1016/j.gie.2018.06.036]

52 **Tsuboi A**, Oka S, Aoyama K, Saito H, Aoki T, Yamada A, Matsuda T, Fujishiro M, Ishihara S, Nakahori M, Koike K, Tanaka S, Tada T. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig Endosc* 2020; **32**: 382-390 [PMID: 31392767 DOI: 10.1111/den.13507]

53 **Rodriguez-Diaz E**, Baffy G, Lo WK, Mashimo H, Vidyarthi G, Mohapatra SS, Singh SK. Real-time artificial intelligence-based histologic classification of colorectal polyps with augmented visualization. *Gastrointest Endosc* 2021; **93**: 662-670 [PMID: 32949567 DOI: 10.1016/j.gie.2020.09.018]

54 **Komeda Y**, Handa H, Watanabe T, Nomura T, Kitahashi M, Sakurai T, Okamoto A, Minami T, Kono M, Arizumi T, Takenaka M, Hagiwara S, Matsui S, Nishida N, Kashida H, Kudo M. Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience. *Oncology* 2017; **93** Suppl 1: 30-34 [PMID: 29258081 DOI: 10.1159/000481227]

55 **Akbari M**, Mohrekesh M, Nasr-Esfahani E, Soroushmehr SMR, Karimi N, Samavi S, Najarian K. Polyp Segmentation in Colonoscopy Images Using Fully Convolutional Network. *Annu Int Conf IEEE Eng Med Biol Soc* 2018; **2018**: 69-72 [PMID: 30440343 DOI: 10.1109/EMBC.2018.8512197]

56 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: 29042219 DOI: 10.1053/j.gastro.2017.10.010]

57 **Gong D**, Wu L, Zhang J, Mu G, Shen L, Liu J, Wang Z, Zhou W, An P, Huang X, Jiang X, Li Y, Wan X, Hu S, Chen Y, Hu X, Xu Y, Zhu X, Li S, Yao L, He X, Chen D, Huang L, Wei X, Wang X, Yu H. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020; **5**: 352-361 [PMID: 31981518 DOI: 10.1016/S2468-1253(19)30413-3]

58 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. 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 [PMID: 29066576 DOI: 10.1136/gutjnl-2017-314547]

59 **Mori Y**, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018; **169**: 357-366 [PMID: 30105375 DOI: 10.7326/M18-0249]

60 **Misawa M**, Kudo SE, Mori Y, Cho T, Kataoka S, Yamauchi A, Ogawa Y, Maeda Y, Takeda K, Ichimasa K, Nakamura H, Yagawa Y, Toyoshima N, Ogata N, Kudo T, Hisayuki T, Hayashi T, Wakamura K, Baba T, Ishida F, Itoh H, Roth H, Oda M, Mori K. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. *Gastroenterology* 2018; **154**: 2027-2029. e3 [PMID: 29653147 DOI: 10.1053/j.gastro.2018.04.003]

61 **Taunk P**, Atkinson CD, Lichtenstein D, Rodriguez-Diaz E, Singh SK. Computer-assisted assessment of colonic polyp histopathology using probe-based confocal laser endomicroscopy. *Int J Colorectal Dis* 2019; **34**: 2043-2051 [PMID: 31696259 DOI: 10.1007/s00384-019-03406-y]

62 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

63 **Tong Y**, Lu K, Yang Y, Li J, Lin Y, Wu D, Yang A, Li Y, Yu S, Qian J. Can natural language processing help differentiate inflammatory intestinal diseases in China? *BMC Med Inform Decis Mak* 2020; **20**: 248 [PMID: 32993636 DOI: 10.1186/s12911-020-01277-w]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 36-49

DOI: 10.37126/aige.v2.i2.36

ISSN 2689-7164 (online)

MINIREVIEWS

Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged?

James Weiquan Li, Tiing Leong Ang

ORCID number: James Weiquan Li 0000-0002-5241-4278; Tiing Leong Ang 0000-0001-9993-8549.

Author contributions: Li JW performed the literature search and wrote the manuscript; Ang TL performed the literature search and reviewed the manuscript.

Conflict-of-interest statement: The authors declare no conflict of interests for this article.

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: Singapore

James Weiquan Li, Tiing Leong Ang, Department of Gastroenterology and Hepatology, Changi General Hospital, Singapore 529889, Singapore

Corresponding author: Tiing Leong Ang, FASGE, FRCP, MBBS, Chief Physician, Professor, Department of Gastroenterology and Hepatology, Changi General Hospital, 2 Simei Street 3, Singapore 529889, Singapore. ang.tiing.leong@singhealth.com.sg

Abstract

Research in artificial intelligence (AI) in gastroenterology has increased over the last decade. Colonoscopy represents the most widely published field with regards to its use in gastroenterology. Most studies to date center on polyp detection and characterization, as well as real-time evaluation of adequacy of mucosal exposure for inspection. This review article discusses how advances in AI has bridged certain gaps in colonoscopy. In addition, the gaps formed with the development of AI that currently prevent its routine use in colonoscopy will be explored.

Key Words: Artificial intelligence; Endoscopy; Colonoscopy; Detection; Diagnosis

Core Tip: The use of artificial intelligence (AI) for colonoscopy has been studied most extensively for polyp detection and characterization. Despite advances made in this field, AI systems studied for these purposes represent only the machine learning domain of AI, and individual machine learning algorithms used in these studies are each focused on performing a very narrow task. While they may bridge existing gaps in polyp detection and real-time optical diagnosis of colorectal polyps, the introduction of AI into colonoscopy will also mean that there are new gaps that must be bridged for AI systems to be routinely used in clinical practice.

Citation: Li JW, Ang TL. Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged? *Artif Intell Gastrointest Endosc* 2021; 2(2): 36-49

URL: https://www.wjgnet.com/2689-7164/full/v2/i2/36.htm

DOI: https://dx.doi.org/10.37126/aige.v2.i2.36

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: March 19, 2021
Peer-review started: March 19, 2021
First decision: March 26, 2021
Revised: March 27, 2021
Accepted: April 20, 2021
Article in press: April 20, 2021
Published online: April 28, 2021

P-Reviewer: Viswanath YK
S-Editor: Wang JL
L-Editor: Filipodia
P-Editor: Wang LL

INTRODUCTION

The use of artificial intelligence (AI) in gastroenterology has gained momentum in the past decade. This is reflected in the increasing number of publications in the field of AI in endoscopy, most of which have been centered on colonoscopy. This is understandable as the unique role of colonoscopy in the prevention and management of colorectal cancer (CRC), together with the unmet needs in this field, has created the perfect milieu for the introduction of AI into world of endoscopy.

CRC represents one of the leading causes of cancer-related morbidity and mortality worldwide[1,2]. Colonoscopy decreases CRC-related mortality[3,4], with a 1% increase in adenoma detection rate (ADR) estimated to decrease interval CRC by 3%[5]. As such, a key barrier to overcome is the adenoma miss rate (AMR), which has been estimated in a meta-analysis to be a high as 22% overall, with a higher AMR when diminutive adenomas are considered[6]. Another unmet need in colonoscopy is the need for accuracy in the optical diagnosis of colonic polyps in relation to their actual histology. Up to 90% of lesions detected on colonoscopy consist of diminutive (≤ 5 mm) and small (6-9 mm) polyps, with the progression rates to advanced adenomas or CRC postulated to be low based on evidence from available studies[7]. It is therefore no surprise that most of the literature to date has focused on computer-assisted detection (CADe)[8,9] and computer-assisted diagnosis (CADx)[10-12] applications in colonoscopy.

This review article evaluates the areas in colonoscopy where AI may be a bridge for certain gaps in clinical practice. It will also explore in detail the current limitations and pitfalls in the application of AI in colonoscopy, highlighting how despite the proliferation of literature on this topic and what it promises to offer, AI may be a new gap in endoscopy which clinicians need to work to bridge.

LITERATURE SEARCH

We performed a comprehensive literature search in the PubMed, MEDLINE and EMBASE (up to March 17, 2021) electronic databases to identify relevant clinical trials that evaluated the roles of AI systems in colonoscopy. Electronic searches were also supplemented with manual searches of the references in the included studies and review articles.

AI TERMINOLOGY IN COLONOSCOPY

What does the term AI mean in colonoscopy?

The term "artificial intelligence" was first coined by John McCarthy in 1956 at the Dartmouth Summer Research Project. In essence, it is a branch in computer science where computer systems are designed to perform tasks which would ordinarily require human intelligence. This definition is extremely broad and often confuses clinicians to what exactly the capabilities, and by inference, the limitations of AI are in their respective fields[13]. There is therefore a need to define what AI means in colonoscopy as this is a prerequisite for meaningful discussion of its role in colonoscopy.

Published and ongoing studies incorporating AI in the context of colonoscopy involve the machine learning (ML) domain of AI. ML refers to the use of algorithms, which form predictive and descriptive models based on analysis of input data provided by investigators (the training set)[14]. These algorithms undergo multiple iterations of these models with the goal of performing a specific task, the aim of which is to come to a specified classification output (*e.g.*, polyp or no polyp) when the algorithms are tested on an unseen set of data (the test set). In practical terms and in the context of colonoscopy, this is achieved using either handcrafted models or deep learning (DL).

A useful mental model in understanding the scope of and roles which AI plays in colonoscopy is to regard the progress made in this field as "waves"[15]. It is crucial to understand that the methods, technologies, and results from earlier AI studies are not obsolete the moment a "better" or "faster" computer system is available based on results we as clinicians are familiar with such as the ADR and adenoma per colonoscopy (APC), or technical matrices that we may gravitate towards such as the processing speed of an algorithm. Rather, these "waves" are continuously interacting

and building on top of each other, and as a result, have a strong influence on the development of later technologies. The earlier "waves" remain relevant and may sometimes harbor solutions to certain issues faced with CADe and CADx support tools, which will be discussed later in this article. Having this mental model also helps us better understand the intrinsic biases present in all forms of ML regardless of advancements made in AI, which is essential for critical appraisal of literature surrounding AI in clinical practice.

AI terminology relevant to colonoscopy

Commonly used terms in AI which are relevant to this review article will be discussed here. This list is not meant to be exhaustive and is meant instead to highlight terms which will help the reader understand the later critiques and solutions offered in this paper.

AI can be categorized very broadly into weak (or narrow) AI and strong AI. The former refers to systems built to solve a specific problem or performing a single task extremely well, without an emphasis on elucidating how human reasoning works. This type of AI operates within significant constraints and a limited context. The latter term, also referred to as artificial general intelligence, aims to build systems which think like humans.

Features in ML refer to the set of numbers which quantitatively summarize and represent in a compact fashion the input data. For example, differences in morphology of polyps as defined in the Paris classification[16] and pit patterns[17] can be converted into different arrays of numbers which an ML algorithm can use to generate a prediction such as "polyp" or "no polyp" in a CADe application. Conventional learning by the ML algorithm may be supervised, where training takes place on labeled data sets, or unsupervised, where commonalities are used to identify groups within data. Supervised learning occurs on pre-established input and output pairs, enabling the ML algorithm to learn predictive mathematical models which can then map the input from unseen data into an outcome of interest (*e.g.*, neoplastic, or hyperplastic). In contrast, unsupervised learning predicts similarities between data points through looking at the underlying structure of the data provided, with no prior knowledge of its significance.

Handcrafted knowledge represents the first "wave" of AI. This consisted of knowledge-based methods where manual extraction and selection of characteristics of an object such as polyp shape and texture, are used to create mathematical models which can achieve a class or numerical output. This is labor-intensive and as a result, are usually implemented on small sets of data. These systems do not have the ability to learn and were of limited clinical use. DL is another form of ML where an artificial neural network (ANN) is used to perform the same task. ANNs are supervised ML models where interconnected artificial neurons form layered networks. Signals travel *via* weighted inputs from artificial neurons in the previous layer to the next layer, which then propagate the signal when a predefined threshold is reached, like how biological neurons work. Classification can be optimized, and the system enhanced by adjustment of the weights given to these inter-neuron connections.

Deep convolutional neural networks (DCNNs) have enabled more hidden layers to be added to the input and output layers of ANN, a development which has been facilitated by advancements made in other areas of computer science as this is computationally expansive. In addition, convolutional layers apply filters (a set of weights) in a systematic fashion to each overlapping part of the input data. In this manner, large numbers of filters can be applied to the training set of data in parallel under the constraints of the intended task, for example classification of an image as having a polyp or not in colonoscopy, allowing information to be extracted directly from images training data to form a feature map. DCNN usually require large amounts of labelled training data, which are derived wither from public databases or private collections in individual institutions.

Hyperparameters in ML refer to all parameters that have been arbitrarily set by the investigator and are used to configure the model for optimal performance at a specific task or on a specific dataset. As opposed to model parameters, which are learned automatically during training of the model, hyperparameters are manually set and affects the learning process and ultimately, the behavior of the model. This is useful in understanding the roles (and potential biases resulting from) the optimization and training process of AI models used in colonoscopy. The training set refers to the initial dataset used to determine optimal parameters after multiple rounds or iterations of adjustments. The validation set is mostly (but not always) a different dataset where these parameters are tested and adjusted. It is also used to optimize the hyperparameters in the model. Lastly, the test set refers to a new set of unseen data which is used

to test the model and its generalizability.

AI: BRIDGING THE GAP IN COLONOSCOPY

AI in the field of colonoscopy has been studied primarily for polyp detection, polyp characterization in terms of predicted histology, and for quality assurance in the performance of colonoscopy.

Polyp detection

The rate of missed polyps was mentioned earlier in the introduction. The AMR in influenced by different factors, among which the endoscopist is considered one of the major determinants[18-21]. These human biases may be due to distraction during colonoscopy, fatigue, or the inability to maintain a sustained level of alertness during withdrawal. These lead to errors in perception where the endoscopist may miss polyps which are visible on the monitor. The role of "second readers" in colonoscopy in increasing ADR[22,23] lends support to the hypothesis that CADe may help increase APC and ADR, and decrease AMR, during colonoscopy.

At the time of writing, there are six randomized controlled trials (RCTs)[24-29] to date that have evaluated the role of CADe in colonoscopy. Hassan *et al*[9] recently performed a systematic review and meta-analysis of five of these studies[24,25,27-29], which consisted of 4354 participants. The pooled ADR was significantly higher in the CADe group compared with the control group (36.6% *vs* 25.2%; relative risk [RR] 1.44; 95% confidence interval [CI]: 1.27-1.62; $P < 0.1$), with all of the included RCTs reporting a significant increase in ADR individually. APC, which is defined as the total number of adenomas found divided by the total number of colonoscopies and has good correlation with ADR[30,31], was also significantly higher in the CADe compared to the control group (0.58 *vs* 0.36; RR 1.70; 95%CI: 1.53-1.89; $P < 0.01$). The mean withdrawal time in the CADe and control groups was shown to be statistically different in this meta-analysis.

An interesting prospective study conducted by Wang *et al*[32] showed that the AMR was decreased with CADe. This study differed from the RCT mentioned above in that tandem colonoscopies were performed. Patients in this study were randomly assigned to colonoscopy with CADe or colonoscopy without CADe by an endoscopist, followed immediately by the other procedure. The study showed that the AMR and polyp miss rates were significantly lower in the CADe colonoscopy group compared to the routine colonoscopy group (13.89% *vs* 40.00%, $P < 0.0001$ and 12.98% *vs* 45.90%; $P < 0.0001$, respectively). These results were also consistent regardless of colonic segments, *i.e.* the AMR was significantly lower in the CADe group in the ascending, transverse, and descending colon.

Polyp characterization (optical prediction of polyp histology)

In contrast to CADe for polyp detection, CADx deals with the interpretation of polyp appearance during colonoscopy to determine the predicted histology. Polyp classification systems such as the Kudo pit pattern[17], Sano *et al*[33], NBI International Colorectal Endoscopic (NICE)[34], and Japan NBI Expert Team (JNET)[35] classifications were developed with the purpose of predicting polyp histology and severity of neoplasia to guide therapy. The use of these classification systems for optical prediction of colorectal polyp histology requires the proper equipment, structured training, and experience in clinical application. Studies have shown wide variation in the sensitivity and specificity of NICE and JNET classifications, with most studies reporting a moderate interobserver agreement at best[36-39].

With the clinical use of CADe, the detection of diminutive polyps is likely to increase exponentially, as demonstrated in the CADe RCT mentioned[24,25,27-29]. Most diminutive polyps tend to be hyperplastic in nature with low malignant potential. The "resect and discard" and "detect and leave" strategies for such polyps were previously studied to address these issues before the emergence of AI but have failed to gain traction due to the need for better quality training and quality assurance in the accurate optical diagnosis of colon polyps[40-42]. The threshold for optical biopsy technologies in high confidence predictions established by the American Society for Gastrointestinal Endoscopy (ASGE) Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI)[43] are deemed appropriate targets for CADx support tools[44]. A systematic review and meta-analysis by ASGE[45] showed that these thresholds were met using NBI only among NBI experts, illustrating the difficulty and practical limitations of replying on the use of these forms of imaging by

endoscopists in general to achieve accurate optical diagnoses of colorectal polyps. Hence, this represents a significant clinical gap which AI has the potential to bridge in colonoscopy.

CADx is postulated to aid in this field of colorectal polyp management by using DL models to increase the accuracy of prediction of polyp histology during colonoscopy[46]. At the time of writing, there are currently no RCT evaluating CADx in colonoscopy. In a study by Jin *et al*[10], a DCNN was trained to differentiate between adenomatous and hyperplastic diminutive colorectal polyps with an overall accuracy of 86.7% using polyp histology as the gold standard. The system was tested on 22 endoscopists with varying expertise such as novice endoscopists, colonoscopy experts with differing levels of expertise in NBI, and NBI-trained experts. The use of CADx markedly improved the accuracy of novice endoscopists in differentiating adenomatous and hyperplastic polyps from 73.8% to 85.6% ($P < 0.05$), which was comparable to the baseline accuracy of NBI-trained experts (87.6%). However, in the colonoscopy expert and NBI-trained expert groups, this increase in accuracy was less impressive (83.8% to 89.0% and 87.6% to 90.0, respectively). The overall time to diagnosis per polyp was also decreased from 3.92 s to 3.37 s; $P = 0.42$).

A review of CADx predictions[47] for diminutive polyp histology which included 9 studies[48-56] showed a pooled sensitivity of 93.5% (95%CI: 90.7%-95.6%) and specificity of 90.8% (95%CI: 86.3%-95.9%), with a pooled area under the curve of 0.98. This pooled analysis of diminutive polyps had a negative predictive value (NPV) of 0.91 (95%CI: 0.89-0.94). This meets the 90% or greater threshold for NPV in adenomatous histology in rectosigmoid diminutive polyps recommended by the ASGE PIVI[43] and thus would in theory support a "diagnose and leave" strategy if these applications are validated in clinical use. However, most of these studies are retrospective in nature or, when conducted prospectively, involved the use of *ex vivo* video or still images.

Few prospective studies on CADx in real-time colonoscopy are currently available in the literature. In a single-center, open-label, prospective study of 791 consecutive patients undergoing colonoscopy in a university hospital, Mori *et al*[54] evaluated the performance of CADx in a clinical setting using endocytoscopy (CF-H290ECI; Olympus Corp, Tokyo, Japan). NBI was applied to visualize the microvascular pattern and methylene blue staining for cellular structure under these ultra-magnifying colonoscopes with 520X optical zoom capability. Of the 466 diminutive polyps found in this study, 250 polyps were in the rectosigmoid colon. The CADx system using endocytoscopy had an NPV for diminutive rectosigmoid adenomas ranging from 93.7% to 96.4% with methylene blue staining and 95.2% to 96.5% with NBI. This is well above the "diagnose and leave" threshold of 90% recommended by the ASGE PIVI[43] described. This prospective study also provides evidence for utilization of CADx for prediction of polyp histology in a clinical setting which may have an impact on decisions on polyp management real-time.

In an earlier study with a similar design by Horiuchi *et al*[56], CADx was evaluated with the use of autofluorescence imaging (AFI) to differentiate diminutive rectosigmoid polyps in real-time colonoscopies. The CADx system used software-based automatic color intensity analysis, which utilized AFI's ability to differentiate polyps based on the ratio of green to red tone intensities and was tested on 258 rectosigmoid polyps in 95 patients undergoing colonoscopy. The CAD-AFI system achieved an NPV for adenomatous polyps of 93.4% (95%CI: 89.0%-96.4%), which again exceeds the 90% "diagnose and leave" threshold[43]. In addition, the NPV using CAD-AFI was comparable to that of diagnoses made by endoscopists using AFI in the study (94.9%; 95%CI: 90.8%-97.5%).

Quality assurance in colonoscopy

Quality indices such as a high cecal intubation rate and adequate withdrawal time have been studied extensively[57,58]. However, these quality indices in colonoscopy performance and reporting are not always adhered to for a variety of factors such as training, lack of real-time feedback and failure of enforcement[59-61]. In an RCT of 704 patients by Gong *et al*[26], which used an AI system called ENDOANGEL, the withdrawal speed and time, as well as the adequacy of mucosal exposure, was monitored in real-time and in an automated fashion. The resulted in a significantly longer withdrawal time in the ENDOANGEL[62] *vs* the control group (mean 6.38 min *vs* 4.76 min, respectively; $P < 0.0001$). This translated into an increased ADR in the ENDOANGEL group and, more significantly, is the only RCT to date which demonstrates an AI system which can increase the rate of detection of adenomas 10 mm or larger in size (10/355 *vs* 1/349, respectively; odds ratio [OR] 9.50, 95%CI: 1.19-75.75; $P = 0.034$). Su *et al*[28] used both a CADe tool together with an automatic quality

control system (AQCS) to increase ADR and APC. The AQUS consisted of a timer on the monitor and audio prompts for the Endoscopist to slow down withdrawal speed when unstable and blurry frames were displayed or when the Boston Bowel Preparation Scale (BPPS) in a colonic segment was < 2. This study showed an improved withdrawal time (7.03 min *vs* 5.68 min; $P < 0.001$) and rate of adequate bowel preparation (87.34% *vs* 80.63%; $P = 0.023$) in the AQCS group in addition to the mentioned significant increase in ADR and APC.

AI: A GAP NEEDING TO BE BRIDGED IN COLONOSCOPY?

While AI has emerged in the world of endoscopy with much promise, there are several significant gaps which need to be bridged before it can be routinely applied in colonoscopy in a clinical setting.

Undefined and unspecified role in clinical environment

A major bridge which needs to be bridged before AI systems can be applied in routine environments is its generalizability. Three of the five CADe RCT[25,27,28] available involved senior endoscopists with extensive experience in colonoscopy. ADR is dependent on several factors, one of which includes experience. A more experienced endoscopist is not only skilled in recognition, but also in scope handling and consequent mucosal exposure during withdrawal. The role of a "second reader" in previous studies[22,23] in increasing small adenoma detection rates suggests that trainees and Nurses, who by inference have less "experience" than the senior endoscopist, have no issues recognizing a polyp visible on screen. In addition, as discussed in the ENDOANGEL study, one of the largest increments in ADR and the only increase in detection of adenomas larger then 10 mm was seen in the RCT by Gong *et al*[26], where real-time feedback on adequacy of mucosal exposure was studied. An obvious but less often mentioned fact is that any CADe algorithm is still completely dependent on the endoscopist to present optimal images with adequately exposed colonic mucosa in each real-time colonoscopy performed in a busy clinical setting. A polyp not visible on the screen will not be detected by a CADe tool, no matter how powerful the algorithm is[33]. This has implications on how generalizable available data is for clinical use, as more studies involving both "high detectors" and "low detectors" are required[25,63].

Most RCT in CADe to date were conducted in single centers. Moreover, except for the study by Wang *et al*[27] where a second monitor was used and visible only to an observer who reported the alerts, the rest of the RCT were non-blinded studies[24-26,28-29]. It is not known what the impact of the latter factor may be in actual clinical practice, as non-blinded endoscopists in these studies may put in more effort in exposing colonic mucosa for inspection when they are under observation. This Hawthorne Effect, together with the single-center experiences of most of these RCT, also limit their generalizability to routine clinical practice. While single monitors are encouraged[44] due to presumed gaze limitations of endoscopists and the need to reduce distractions, it is the opinion of the authors that a dual monitor setting in clinical trials plays a crucial role in achieving a double-blind and objective environment for assessment of the performance of the AI system and to bridge this gap. Furthermore, it resembles tandem colonoscopy in that the performance of the AI system can be compared directly against endoscopists of varying skill levels and experience. Useful information such as the AMR can be determined accurately without the patient having to go through an additional colonoscopy like in a traditional tandem study with this methodology.

Another limitation to the generalizability of the published results of AI systems for polyp detection and characterization is the differences in operational environments of different endoscopy suites and centers. These can vary greatly between institutions, even those located in the same country[64]. Unlike a new endoscopic method or classification system which can be taught or standardized in training or with major society guidelines, different AI algorithms have unique hardware and software requirements which must be fulfilled for technical integration into the operational environment. For instance, some may be fully integrated into the processing unit[65] while others may be web-based applications or require an additional laptop to be linked to the endoscopy stack to function. The latter may require cloud integration support, which in turn is likely to be vendor-specific and has implications in procurement and cybersecurity. This technical integration into the operational environment is key, as the development environment from which these AI systems are derived may be vastly

different[66]. Most clinical trials understandably focus on the clinical aspects like the ADR and APC and the outcomes will inevitably be based on these primary objectives. However, few studies have reported the technical specifications and limitations of the AI systems they are investigating. The rare studies that do report them, do so in varying details, most of which are insufficient for interpretation and contextualization into the operational environment. Moreover, most of the published trials have been conducted in academic or expert centers and in several instances, in the same institutions where the AI algorithm was developed, *i.e.* the development and operational environment are the same[3,47]. Individual institutions may have difficulty integrating these systems due to budgeting constraints, existence of legacy systems which are incompatible with the software and hardware requirements of the AI systems, logistical limitations such as space, and established workflows in endoscopy which does not cater to the introduction of an AI system.

The current scope of AI applications in colonoscopy in the literature is also largely skewed towards to polyp detection, characterization, and assessment of adequacy of mucosal exposure, which is ultimately linked to ADR. When translated to clinical practice, this effectively confines the indications for which AI should be used in colonoscopy to CRC screening or indications where one might expect to find colorectal polyps in the process of performing a colonoscopy. All systems developed in the field of AI in colonoscopy, from handcrafted models to the most complex DCNN, are fundamentally "weak AI." This is a term used to describe AI systems designed to solve a single problem or narrow task[15]. In a clinical setting, indications for colonoscopy are widely variable and the pre-test probability of finding of a polyp may be low. An endoscopist will be able to process the demographic data, clinical course, medical history, clinical condition, laboratory investigations and concerns of the patient and use this information during the colonoscopy. For example, an 85-year-old patient who is troubled by per rectal bleeding has a hugely different indication and clinical index of suspicion than a 50-year-old male with a family history of early CRC. In the former case, the endoscopist's focus may be on looking for angiodysplasia, diverticular disease or hemorrhoids as the etiology. A "strong AI" system would be able to think and adapt like a human and calibrate the weights in its layers to perform the task at hand, determine the appropriate classification output and achieve the correct alarm settings. However, current AI systems will continue looking for polyps and may present a distraction to the Endoscopist if used in this clinical example, prolonging the time taken for colonoscopy in an elderly patient, who may have multiple co-morbidities and for whom resection of small or diminutive adenomas may not have clinical relevance, much less answer the clinical question at hand. A trainee endoscopist or an experienced nurse, on the other hand, would be able to immediately recognize an unusual finding, such as multiple angiodysplasia or extensive diverticular disease, even if they were not formally trained to recognize these abnormalities.

It should be noted that AI has also been studied in colonoscopy outside the context of polyp detection, characterization, and quality assurance. Endocytoscopy has been used with AI to accurately detect persistent histologic inflammation in patients with ulcerative colitis (UC) which was reproducible based on static images[67]. A separate group used a deep neural network to predict endoscopic and histologic remission in UC patients based on evaluation of static images obtained from colonoscopy with high accuracy[68]. However, studies looking at indications other than polyp detection and characterization are few and far between.

Technical biases and lack of technical knowledge among clinicians

There is significant variability and a lack of standardization in reporting of the technical aspects of AI algorithms in clinical trials[69]. In addition, clinicians may not have the technical knowledge to critically appraise AI literature given that this has not been a formal part of training or an emphasis in clinical practice until relatively recently. A "minimum reporting standard" and practical knowledge of terms and potential biases on the part of investigators and clinicians, respectively, is required to bridge these gaps[70-72].

A practical knowledge of commonly used terms and how AI systems are derived is necessary for the clinician to appreciate the technical biases inherent to these algorithms. While the inclusion criteria of patients in clinical trials is clearly defined, the criteria for inclusion of the input data for the AI system during training and validation may not always be included in the methodology. This is crucial as most AI systems for CADe were tested in the same centers where they were developed[73]. This is often due to the ease with which large amounts of data are readily available for training and validation. Although the training, validation, and test datasets may be

different, they could be derived from the same database in a single, often expert, center, which is then split to form these datasets. The nature of the images used could be highly similar in terms of quality (*e.g.*, no confounding fecal material and bubbles and polyps always centered in the image) and labelling (*e.g.*, experts from different centers may mark out the most obvious abnormal area or delineate even the most minute detail which does not look like normal colonic mucosa for sessile serrated polyps depending on their level of skill and the training received, while experts from the same center are more likely to label lesions similarly). Prevalence and variability in presentations of disease may also differ depending on the populations studied, but the sample of images used in training and validating the AI algorithm may not necessarily reflect this natural variability of disease if data from a single center is used in the development of the AI system. This is a form of selection bias, as input data is not selected at random and hence is not fully representative of the study population in which the AI system is meant to function. This could impact the hyperparameters chosen during validation, and lead to overfitting, which occurs when the mathematical model derived is optimized to work on the training data and fits this data too tightly. This would limit its generalizability when new data is presented to the same AI algorithm.

Moreover, the proportion of "positive" to "normal" images used for training is not often mentioned in the published literature. For example, in a CADe application, polyps of various shapes, sizes and colors may be included in the training dataset to expose the AI algorithm to all possible eventualities when presented with an image with even the subtlest polyp. However, the "normal" images used may be disproportionately lower when compared to the natural prevalence of adenomas in the population. In addition, there may not be the same rigor in the selection of "normal" images for training. Variations in degrees of bowel preparation, bubbles, and artefacts due to the light source reflecting off normal colonic mucosa may thus not be reflected in images supplied to the AI algorithm for training. Positive and negative predictive values are determined by the prevalence of disease, and this may result in a higher proportion of false positives per true positive detected in clinical practice, depending on how the ratio of "positive" to "normal" images used in training compares with the true prevalence of the lesion of interest (*e.g.*, polyps) in the study population. This is a factor which needs to be adjusted for in the AI algorithm[74].

A certain form of publication bias may also exist as clinicians who wish to publish on the topic of AI will search for references almost exclusively from medical journals. For example, meta-analysis and systematic reviews on the use of AI in colonoscopy may take a very clinical slant, while publications in computer science and engineering journals which may add technical dept to the chosen topic on AI being discussed will not be included. Even if a search were performed for these articles, the inclusion criteria for the literature search will inevitably involve clinical-based endpoints like ADR and APC, and almost always exclude publications from computer science and engineering journals as a result. The barrier to entry in medical journals for these studies is high, as editors and reviewers, who themselves are clinicians, may not have enough technical knowledge to feel comfortable about accepting these articles for publication, and may also be compounded by fear of a lack of interest or understanding in the readership. On the other hand, AI and ML experts will not be familiar with the clinical aspects or relevance of their research and would not be able to pitch it at a level that would be acceptable to a Medical journal and its readership. This may result in a "reinforcement bias" of sorts, where only certain types of publications from a few expert centers and which revolve around common themes are published repeatedly and in different forms in Medical journals, whereas significant developments in AI and ML which may have the potential for changing clinical practice are missed out. The same technical terms specific to these publications will also be mentioned repeatedly, while novel approaches and new technical terms unfamiliar to clinicians may never see publication in a medical journal. The endoscopy readership may already have been "overfitted" towards polyp detection and characterization in the endoscopy literature[75], while neglecting the fact that, as mentioned, the use of AI in colonoscopy to date has utilized only an extremely limited aspect of AI and in a very narrow clinical context. Including computer science experts in the editorship and as reviewers for Medical journals may help to bridge the gap in these technical and publication biases.

Physician sentiment towards AI

Physician sentiment is a significant determinant on how quickly technologies and recommendations are deployed in a clinical setting. A recently conducted online survey among Gastroenterologists in the United States showed high overall interest in

CADe and perception that it would increase ADR (85.5% and 75.8%, respectively)[76]. However, the same survey also showed that majority of the respondents felt that CADe will prolong the time taken per colonoscopy, despite evidence to the contrary[9,24,25,27-29].

Concerns about operator dependence, or "deskilling", of the Endoscopist due to reliance on CADe and CADx for detection and characterization of polyps, respectively, are also mentioned in this survey[76] and other reviews[44,73]. Another major concern shown in the survey by Wadhwa *et al*[76] was the perceived increase in cost per procedure (75.2%). While concerns such as withdrawal time have been addressed independently in several RCT, others such as operator dependance and cost-effectiveness have not studied. Hence, physician sentiment may be another significant gap in AI which needs to be bridged in the field of colonoscopy.

Medicolegal challenges and future directions

AI algorithms which utilize DL are considered "black box" models, meaning that it is almost impossible to trace the decision-making process which led to the output determined by the algorithm when faced with a specific task (*e.g.*, polyp or no polyp in the image, hyperplastic or adenomatous). One of the major gaps in clinical use of AI systems in colonoscopy is medicolegal liability when a misdiagnosis or missed diagnosis occurs. While a clinician's account of events and the accompanying documentation can be helped up to scrutiny, the black box nature of DL algorithms means that the root cause and mitigating factors surrounding such a case may never be elucidated or even discovered. This has ethical implications in the event of harm to a patient[77], particularly if no clear protocol exists to define how an AI system should interface with its user and what its limits are, as the error may be due to deviation from safe use of the system or from an error of the AI system itself[78].

As AI systems, like other healthcare interventions, may have unpredictable errors, this inability to explain the errors or to detect them as they occur due to their black box nature may result in a perpetuation of systemic errors with unknown clinical implications if they are scaled up rapidly for routine clinical use in all colonoscopies. It is also unknown if the liability rests with the manufacturer, the regulatory body approving its use, or the clinician interfacing with the AI system. Having a reliable and accountable post-deployment surveillance plan is perhaps one of the strategies to minimize this risk.

Lastly, while AI systems have been shown to improve various quality indices associated with colonoscopy, one should remember that they are still limited most of all by our current expertise in this field. A useful example to illustrate this is the fact that there is currently no AI system capable of detecting dysplasia in UC. The availability of DCNN with high computing power and hardware to support the required processing speeds would have made this a rather simple task from an ML point of view. However, the optimal method of surveillance for dysplasia in UC and its optical features do not have the same clinical certainty as colorectal polyps in CRC screening, with resultant discrepancies in surveillance and biopsy practices[79,80]. Moreover, there is wide interobserver variability in the histological diagnosis of dysplasia in UC[81] and an inadequate understanding of its pathogenesis[82]. It is therefore understandable that there would be a paucity of expertly labelled data for "dysplasia" and "non-dysplasia" controls in UC patients for the training of an ML algorithm. Similarly, other potential AI applications in colonoscopy could include localization of diverticular bleeding and an automated scoring system for adequacy of bowel preparation which includes the BPPS[83] and the newly validated Colon Endoscopic Bubble Scale[84]. The clinical expertise and research in these fields must progress sufficiently for an accompanying increase in standardized and labelled data to be available for such future AI systems to be trained on and to materialize.

CONCLUSION

Despite the advances made in the field of AI, most notably for polyp detection and characterization in colonoscopy, there remain significant gaps which need to be bridged before its routine clinical use in colonoscopy.

REFERENCES

1 **Bray F**, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. *Lancet Oncol* 2012; **13**: 790-801 [PMID: 22658655 DOI: 10.1016/S1470-2045(12)70211-5]

2 **Araghi M**, Soerjomataram I, Jenkins M, Brierley J, Morris E, Bray F, Arnold M. Global trends in colorectal cancer mortality: projections to the year 2035. *Int J Cancer* 2019; **144**: 2992-3000 [PMID: 30536395 DOI: 10.1002/ijc.32055]

3 **Kaminski MF**, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, Hoff G, Jover R, Suchanek S, Ferlitsch M, Anderson J, Roesch T, Hultcranz R, Racz I, Kuipers EJ, Garborg K, East JE, Rupinski M, Seip B, Bennett C, Senore C, Minozzi S, Bisschops R, Domagk D, Valori R, Spada C, Hassan C, Dinis-Ribeiro M, Rutter MD. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscopy* 2017; **49**: 378-397 [PMID: 28268235 DOI: 10.1055/s-0043-103411]

4 **Rex DK**, Schoenfeld PS, Cohen J, Pike IM, Adler DG, Fennerty MB, Lieb JG 2nd, Park WG, Rizk MK, Sawhney MS, Shaheen NJ, Wani S, Weinberg DS. Quality indicators for colonoscopy. *Gastrointest Endosc* 2015; **81**: 31-53 [PMID: 25480100 DOI: 10.1016/j.gie.2014.07.058]

5 **Corley DA**, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 2541 [PMID: 24963577 DOI: 10.1056/NEJMc1405329]

6 **van Rijn JC**, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. *Am J Gastroenterol* 2006; **101**: 343-350 [PMID: 16454841 DOI: 10.1111/j.1572-0241.2006.00390.x]

7 **Vleugels JLA**, Hazewinkel Y, Fockens P, Dekker E. Natural history of diminutive and small colorectal polyps: a systematic literature review. *Gastrointest Endosc* 2017; **85**: 1169-1176. e1 [PMID: 28024986 DOI: 10.1016/j.gie.2016.12.014]

8 **Barua I**, Vinsard DG, Jodal HC, Løberg M, Kalager M, Holme Ø, Misawa M, Bretthauer M, Mori Y. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. *Endoscopy* 2021; **53**: 277-284 [PMID: 32557490 DOI: 10.1055/a-1201-7165]

9 **Hassan C**, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, Antonelli G, Yu H, Areia M, Dinis-Ribeiro M, Bhandari P, Sharma P, Rex DK, Rösch T, Wallace M, Repici A. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021; **93**: 77-85. e6 [PMID: 32598963 DOI: 10.1016/j.gie.2020.06.059]

10 **Jin EH**, Lee D, Bae JH, Kang HY, Kwak MS, Seo JY, Yang JI, Yang SY, Lim SH, Yim JY, Lim JH, Chung GE, Chung SJ, Choi JM, Han YM, Kang SJ, Lee J, Chan Kim H, Kim JS. Improved Accuracy in Optical Diagnosis of Colorectal Polyps Using Convolutional Neural Networks with Visual Explanations. *Gastroenterology* 2020; **158**: 2169-2179. e8 [PMID: 32119927 DOI: 10.1053/j.gastro.2020.02.036]

11 **Renner J**, Phlipsen H, Haller B, Navarro-Avila F, Saint-Hill-Febles Y, Mateus D, Ponchon T, Poszler A, Abdelhafez M, Schmid RM, von Delius S, Klare P. Optical classification of neoplastic colorectal polyps - a computer-assisted approach (the COACH study). *Scand J Gastroenterol* 2018; **53**: 1100-1106 [PMID: 30270677 DOI: 10.1080/00365521.2018.1501092]

12 **Kudo SE**, Mori Y, Misawa M, Takeda K, Kudo T, Itoh H, Oda M, Mori K. Artificial intelligence and colonoscopy: Current status and future perspectives. *Dig Endosc* 2019; **31**: 363-371 [PMID: 30624835 DOI: 10.1111/den.13340]

13 **Tang A**, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, Chepelev L, Cairns R, Mitchell JR, Cicero MD, Poudrette MG, Jaremko JL, Reinhold C, Gallix B, Gray B, Geis R; Canadian Association of Radiologists (CAR) Artificial Intelligence Working Group. Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. *Can Assoc Radiol J* 2018; **69**: 120-135 [PMID: 29655580 DOI: 10.1016/j.carj.2018.02.002]

14 **Shalev-Shwartz S**, Ben-David S. Understanding Machine Learning: from theory to algorithms. New York: Cambridge University Press, 2014

15 **Wittenberg T**, Raithel M. Artificial Intelligence-Based Polyp Detection in Colonoscopy: Where Have We Been, Where Do We Stand, and Where Are We Headed? *Visc Med* 2020; **36**: 428-438 [PMID: 33447598 DOI: 10.1159/000512438]

16 Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. *Gastrointest Endosc* 2003; **58**: S3-43 [PMID: 14652541 DOI: 10.1016/s0016-5107(03)02159-x]

17 **Kudo S**, Tamura S, Nakajima T, Yamano H, Kusaka H, Watanabe H. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointest Endosc* 1996; **44**: 8-14 [PMID: 8836710 DOI: 10.1016/s0016-5107(96)70222-5]

18 **Leufkens AM**, van Oijen MG, Vleggaar FP, Siersema PD. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. *Endoscopy* 2012; **44**: 470-475 [PMID: 22441756 DOI: 10.1055/s-0031-1291666]

19 **Rex DK**, Cutler CS, Lemmel GT, Rahmani EY, Clark DW, Helper DJ, Lehman GA, Mark DG. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology* 1997; **112**: 24-28 [PMID: 8978338 DOI: 10.1016/s0016-5085(97)70214-2]

20 **Zhao S**, Wang S, Pan P, Xia T, Chang X, Yang X, Guo L, Meng Q, Yang F, Qian W, Xu Z, Wang Y, Wang Z, Gu L, Wang R, Jia F, Yao J, Li Z, Bai Y. Magnitude, Risk Factors, and Factors Associated

With Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology* 2019; **156**: 1661-1674. e11 [PMID: 30738046 DOI: 10.1053/j.gastro.2019.01.260]

21 **Anderson R**, Burr NE, Valori R. Causes of Post-Colonoscopy Colorectal Cancers Based on World Endoscopy Organization System of Analysis. *Gastroenterology* 2020; **158**: 1287-1299. e2 [PMID: 31926170 DOI: 10.1053/j.gastro.2019.12.031]

22 **Buchner AM**, Shahid MW, Heckman MG, Diehl NN, McNeil RB, Cleveland P, Gill KR, Schore A, Ghabril M, Raimondo M, Gross SA, Wallace MB. Trainee participation is associated with increased small adenoma detection. *Gastrointest Endosc* 2011; **73**: 1223-1231 [PMID: 21481861 DOI: 10.1016/j.gie.2011.01.060]

23 **Lee CK**, Park DI, Lee SH, Hwangbo Y, Eun CS, Han DS, Cha JM, Lee BI, Shin JE. Participation by experienced endoscopy nurses increases the detection rate of colon polyps during a screening colonoscopy: a multicenter, prospective, randomized study. *Gastrointest Endosc* 2011; **74**: 1094-1102 [PMID: 21889137 DOI: 10.1016/j.gie.2011.06.033]

24 **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

25 **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology* 2020; **159**: 512-520. e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

26 **Gong D**, Wu L, Zhang J, Mu G, Shen L, Liu J, Wang Z, Zhou W, An P, Huang X, Jiang X, Li Y, Wan X, Hu S, Chen Y, Hu X, Xu Y, Zhu X, Li S, Yao L, He X, Chen D, Huang L, Wei X, Wang X, Yu H. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020; **5**: 352-361 [PMID: 31981518 DOI: 10.1016/S2468-1253(19)30413-3]

27 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

28 **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc* 2020; **91**: 415-424. e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

29 **Liu WN**, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, Huang J. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; **26**: 13-19 [PMID: 31898644 DOI: 10.4103/sjg.SJG_377_19]

30 **Kahi CJ**, Vemulapalli KC, Johnson CS, Rex DK. Improving measurement of the adenoma detection rate and adenoma per colonoscopy quality metric: the Indiana University experience. *Gastrointest Endosc* 2014; **79**: 448-454 [PMID: 24246797 DOI: 10.1016/j.gie.2013.10.013]

31 **Hilsden RJ**, Bridges R, Dube C, McGregor SE, Naugler C, Rose SM, Rostom A, Heitman SJ. Defining Benchmarks for Adenoma Detection Rate and Adenomas Per Colonoscopy in Patients Undergoing Colonoscopy Due to a Positive Fecal Immunochemical Test. *Am J Gastroenterol* 2016; **111**: 1743-1749 [PMID: 27725649 DOI: 10.1038/ajg.2016.449]

32 **Wang P**, Liu P, Glissen Brown JR, Berzin TM, Zhou G, Lei S, Liu X, Li L, Xiao X. Lower Adenoma Miss Rate of Computer-Aided Detection-Assisted Colonoscopy vs Routine White-Light Colonoscopy in a Prospective Tandem Study. *Gastroenterology* 2020; **159**: 1252-1261. e5 [PMID: 32562721 DOI: 10.1053/j.gastro.2020.06.023]

33 **Sano Y**, Ikematsu H, Fu KI, Emura F, Katagiri A, Horimatsu T, Kaneko K, Soetikno R, Yoshida S. Meshed capillary vessels by use of narrow-band imaging for differential diagnosis of small colorectal polyps. *Gastrointest Endosc* 2009; **69**: 278-283 [PMID: 18951131 DOI: 10.1016/j.gie.2008.04.066]

34 **Rex DK**. Narrow-band imaging without optical magnification for histologic analysis of colorectal polyps. *Gastroenterology* 2009; **136**: 1174-1181 [PMID: 19187781 DOI: 10.1053/j.gastro.2008.12.009]

35 **Sano Y**, Tanaka S, Kudo SE, Saito S, Matsuda T, Wada Y, Fujii T, Ikematsu H, Uraoka T, Kobayashi N, Nakamura H, Hotta K, Horimatsu T, Sakamoto N, Fu KI, Tsuruta O, Kawano H, Kashida H, Takeuchi Y, Machida H, Kusaka T, Yoshida N, Hirata I, Terai T, Yamano HO, Kaneko K, Nakajima T, Sakamoto T, Yamaguchi Y, Tamai N, Nakano N, Hayashi N, Oka S, Iwatate M, Ishikawa H, Murakami Y, Yoshida S, Saito Y. Narrow-band imaging (NBI) magnifying endoscopic classification of colorectal tumors proposed by the Japan NBI Expert Team. *Dig Endosc* 2016; **28**: 526-533 [PMID: 26927367 DOI: 10.1111/den.12644]

36 **Hewett DG**, Kaltenbach T, Sano Y, Tanaka S, Saunders BP, Ponchon T, Soetikno R, Rex DK. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. *Gastroenterology* 2012; **143**: 599-607. e1 [PMID: 22609383 DOI: 10.1053/j.gastro.2012.05.006]

37 **Repici A**, Ciscato C, Correale L, Bisschops R, Bhandari P, Dekker E, Pech O, Radaelli F, Hassan C. Narrow-band Imaging International Colorectal Endoscopic Classification to predict polyp histology: REDEFINE study (with videos). *Gastrointest Endosc* 2016; **84**: 479-486. e3 [PMID: 26928372 DOI:

10.1016/j.gie.2016.02.020]
38 **Komeda Y**, Kashida H, Sakurai T, Asakuma Y, Tribonias G, Nagai T, Kono M, Minaga K, Takenaka M, Arizumi T, Hagiwara S, Matsui S, Watanabe T, Nishida N, Chikugo T, Chiba Y, Kudo M. Magnifying Narrow Band Imaging (NBI) for the Diagnosis of Localized Colorectal Lesions Using the Japan NBI Expert Team (JNET) Classification. *Oncology* 2017; **93** Suppl 1: 49-54 [PMID: 29258091 DOI: 10.1159/000481230]
39 **Kobayashi S**, Yamada M, Takamaru H, Sakamoto T, Matsuda T, Sekine S, Igarashi Y, Saito Y. Diagnostic yield of the Japan NBI Expert Team (JNET) classification for endoscopic diagnosis of superficial colorectal neoplasms in a large-scale clinical practice database. *United European Gastroenterol J* 2019; **7**: 914-923 [PMID: 31428416 DOI: 10.1177/2050640619845987]
40 **Kandel P**, Wallace MB. Should We Resect and Discard Low Risk Diminutive Colon Polyps. *Clin Endosc* 2019; **52**: 239-246 [PMID: 30661337 DOI: 10.5946/ce.2018.136]
41 **Neumann H**, Neumann Sen H, Vieth M, Bisschops R, Thieringer F, Rahman KF, Gamstätter T, Tontini GE, Galle PR. Leaving colorectal polyps in place can be achieved with high accuracy using blue light imaging (BLI). *United European Gastroenterol J* 2018; **6**: 1099-1105 [PMID: 30228899 DOI: 10.1177/2050640618769731]
42 **von Renteln D**, Kaltenbach T, Rastogi A, Anderson JC, Rösch T, Soetikno R, Pohl H. Simplifying Resect and Discard Strategies for Real-Time Assessment of Diminutive Colorectal Polyps. *Clin Gastroenterol Hepatol* 2018; **16**: 706-714 [PMID: 29174789 DOI: 10.1016/j.cgh.2017.11.036]
43 **Rex DK**, Kahi C, O'Brien M, Levin TR, Pohl H, Rastogi A, Burgart L, Imperiale T, Ladabaum U, Cohen J, Lieberman DA. 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 [PMID: 21353837 DOI: 10.1016/j.gie.2011.01.023]
44 **Vinsard DG**, Mori Y, Misawa M, Kudo SE, Rastogi A, Bagci U, Rex DK, Wallace MB. Quality assurance of computer-aided detection and diagnosis in colonoscopy. *Gastrointest Endosc* 2019; **90**: 55-63 [PMID: 30926431 DOI: 10.1016/j.gie.2019.03.019]
45 **ASGE Technology Committee**. , Abu Dayyeh BK, Thosani N, Konda V, Wallace MB, Rex DK, Chauhan SS, Hwang JH, Komanduri S, Manfredi M, Maple JT, Murad FM, Siddiqui UD, Banerjee S. 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 [PMID: 25597420 DOI: 10.1016/j.gie.2014.12.022]
46 **Song EM**, Park B, Ha CA, Hwang SW, Park SH, Yang DH, Ye BD, Myung SJ, Yang SK, Kim N, Byeon JS. Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model. *Sci Rep* 2020; **10**: 30 [PMID: 31913337 DOI: 10.1038/s41598-019-56697-0]
47 **Lui TKL**, Guo CG, Leung WK. Accuracy of artificial intelligence on histology prediction and detection of colorectal polyps: a systematic review and meta-analysis. *Gastrointest Endosc* 2020; **92**: 11-22. e6 [PMID: 32119938 DOI: 10.1016/j.gie.2020.02.033]
48 **Mori Y**, Kudo SE, Chiu PW, Singh R, Misawa M, Wakamura K, Kudo T, Hayashi T, Katagiri A, Miyachi H, Ishida F, Maeda Y, Inoue H, Nimura Y, Oda M, Mori K. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. *Endoscopy* 2016; **48**: 1110-1118 [PMID: 27494455 DOI: 10.1055/s-0042-113609]
49 **Kominami Y**, Yoshida S, Tanaka S, Sanomura Y, Hirakawa T, Raytchev B, Tamaki T, Koide T, Kaneda K, Chayama K. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; **83**: 643-649 [PMID: 26264431 DOI: 10.1016/j.gie.2015.08.004]
50 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. 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 [PMID: 29066576 DOI: 10.1136/gutjnl-2017-314547]
51 **Gross S**, Trautwein C, Behrens A, Winograd R, Palm S, Lutz HH, Schirin-Sokhan R, Hecker H, Aach T, Tischendorf JJ. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc* 2011; **74**: 1354-1359 [PMID: 22000791 DOI: 10.1016/j.gie.2011.08.001]
52 **Lui TKL**, To WP, Ko KK. Superiority of the artificial intelligence image classifier for histological prediction of diminutive colorectal polyps based on non-magnifying endoscopic images. *Hong Kong Med J* 2019; **25**: 31
53 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: 29042219 DOI: 10.1053/j.gastro.2017.10.010]
54 **Mori Y**, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018; **169**: 357-366 [PMID: 30105375 DOI: 10.7326/M18-0249]
55 **Kudo SE**, Misawa M, Mori Y, Hotta K, Ohtsuka K, Ikematsu H, Saito Y, Takeda K, Nakamura H, Ichimasa K, Ishigaki T, Toyoshima N, Kudo T, Hayashi T, Wakamura K, Baba T, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Artificial Intelligence-assisted System Improves Endoscopic Identification of

Colorectal Neoplasms. *Clin Gastroenterol Hepatol* 2020; **18**: 1874-1881. e2 [PMID: 31525512 DOI: 10.1016/j.cgh.2019.09.009]

56 **Horiuchi H**, Tamai N, Kamba S, Inomata H, Ohya TR, Sumiyama K. Real-time computer-aided diagnosis of diminutive rectosigmoid polyps using an auto-fluorescence imaging system and novel color intensity analysis software. *Scand J Gastroenterol* 2019; **54**: 800-805 [PMID: 31195905 DOI: 10.1080/00365521.2019.1627407]

57 **Mathews SC**, Zhao N, Holub JL, Lieberman D. Improvement in colonoscopy quality metrics in clinical practice from 2000 to 2014. *Gastrointest Endosc* 2019; **90**: 651-655. e3 [PMID: 31207221 DOI: 10.1016/j.gie.2019.06.004]

58 **Clark BT**, Rustagi T, Laine L. What level of bowel prep quality requires early repeat colonoscopy: systematic review and meta-analysis of the impact of preparation quality on adenoma detection rate. *Am J Gastroenterol* 2014; **109**: 1714-23; quiz 1724 [PMID: 25135006 DOI: 10.1038/ajg.2014.232]

59 **de Jonge V**, Sint Nicolaas J, Cahen DL, Moolenaar W, Ouwendijk RJ, Tang TJ, van Tilburg AJ, Kuipers EJ, van Leerdam ME; SCoPE Consortium. Quality evaluation of colonoscopy reporting and colonoscopy performance in daily clinical practice. *Gastrointest Endosc* 2012; **75**: 98-106 [PMID: 21907986 DOI: 10.1016/j.gie.2011.06.032]

60 **Leyden JE**, Doherty GA, Hanley A, McNamara DA, Shields C, Leader M, Murray FE, Patchett SE, Harewood GC. Quality of colonoscopy performance among gastroenterology and surgical trainees: a need for common training standards for all trainees? *Endoscopy* 2011; **43**: 935-940 [PMID: 21997723 DOI: 10.1055/s-0030-1256633]

61 **Coe SG**, Panjala C, Heckman MG, Patel M, Qumseya BJ, Wang YR, Dalton B, Tran P, Palmer W, Diehl N, Wallace MB, Raimondo M. Quality in colonoscopy reporting: an assessment of compliance and performance improvement. *Dig Liver Dis* 2012; **44**: 660-664 [PMID: 22579446 DOI: 10.1016/j.dld.2012.03.022]

62 **Zhou J**, Wu L, Wan X, Shen L, Liu J, Zhang J, Jiang X, Wang Z, Yu S, Kang J, Li M, Hu S, Hu X, Gong D, Chen D, Yao L, Zhu Y, Yu H. A novel artificial intelligence system for the assessment of bowel preparation (with video). *Gastrointest Endosc* 2020; **91**: 428-435. e2 [PMID: 31783029 DOI: 10.1016/j.gie.2019.11.026]

63 **Hassan C**, Wallace MB, Sharma P, Maselli R, Craviotto V, Spadaccini M, Repici A. New artificial intelligence system: first validation study versus experienced endoscopists for colorectal polyp detection. *Gut* 2020; **69**: 799-800 [PMID: 31615835 DOI: 10.1136/gutjnl-2019-319914]

64 **Day LW**, Bhuket T, Inadomi JM, Yee HF. Diversity of endoscopy center operations and practice variation across California's safety-net hospital system: a statewide survey. *BMC Res Notes* 2013; **6**: 233 [PMID: 23767938 DOI: 10.1186/1756-0500-6-233]

65 **Weigt J**, Repici A, Antonelli G, Afifi A, Kliegis L, Correale L, Hassan C, Neumann H. Performance of a new integrated CADe/CADx system for detection and characterization of colorectal neoplasia. *Endoscopy* 2021 [PMID: 33494106 DOI: 10.1055/a-1372-0419]

66 **Kelly CJ**, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019; **17**: 195 [PMID: 31665002 DOI: 10.1186/s12916-019-1426-2]

67 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

68 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: 32060000 DOI: 10.1053/j.gastro.2020.02.012]

69 **Nagendran M**, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS, Maruthappu M. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 2020; **368**: m689 [PMID: 32213531 DOI: 10.1136/bmj.m689]

70 **Collins GS**, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet* 2019; **393**: 1577-1579 [PMID: 31007185 DOI: 10.1016/S0140-6736(19)30037-6]

71 **CONSORT-AI and SPIRIT-AI Steering Group**. Reporting guidelines for clinical trials evaluating artificial intelligence interventions are needed. *Nat Med* 2019; **25**: 1467-1468 [PMID: 31551578 DOI: 10.1038/s41591-019-0603-3]

72 **Rivera SC**, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ* 2020; **370**: m3210 [PMID: 32907797 DOI: 10.1136/bmj.m3210]

73 **Mori Y**, Neumann H, Misawa M, Kudo SE, Bretthauer M. Artificial intelligence in colonoscopy - Now on the market. What's next? *J Gastroenterol Hepatol* 2021; **36**: 7-11 [PMID: 33179322 DOI: 10.1111/jgh.15339]

74 **Park SH**, Han K. Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology* 2018; **286**: 800-809 [PMID: 29309734 DOI: 10.1148/radiol.2017171920]

75 **Ang TL**, Carneiro G. Artificial intelligence in gastrointestinal endoscopy. *J Gastroenterol Hepatol* 2021; **36**: 5-6 [PMID: 33448513 DOI: 10.1111/jgh.15344]

76 **Wadhwa V**, Alagappan M, Gonzalez A, Gupta K, Brown JRG, Cohen J, Sawhney M, Pleskow D, Berzin TM. Physician sentiment toward artificial intelligence (AI) in colonoscopic practice: a survey of US gastroenterologists. *Endosc Int Open* 2020; **8**: E1379-E1384 [PMID: 33015341 DOI: 10.1055/a-1223-1926]

77 **Habli I**, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ* 2020; **98**: 251-256 [PMID: 32284648 DOI: 10.2471/BLT.19.237487]

78 **Wiens J**, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, Jung K, Heller K, Kale D, Saeed M, Ossorio PN, Thadaney-Israni S, Goldenberg A. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019; **25**: 1337-1340 [PMID: 31427808 DOI: 10.1038/s41591-019-0548-6]

79 **Singh K**, Al Khoury A, Kurti Z, Gonczi L, Reinglas J, Verdon C, Kohen R, Bessissow T, Afif W, Wild G, Seidman E, Bitton A, Lakatos PL. High Adherence to Surveillance Guidelines in Inflammatory Bowel Disease Patients Results in Low Colorectal Cancer and Dysplasia Rates, While Rates of Dysplasia are Low Before the Suggested Onset of Surveillance. *J Crohns Colitis* 2019; **13**: 1343-1350 [PMID: 30918959 DOI: 10.1093/ecco-jcc/jjz066]

80 **Watanabe T**, Ajioka Y, Mitsuyama K, Watanabe K, Hanai H, Nakase H, Kunisaki R, Matsuda K, Iwakiri R, Hida N, Tanaka S, Takeuchi Y, Ohtsuka K, Murakami K, Kobayashi K, Iwao Y, Nagahori M, Iizuka B, Hata K, Igarashi M, Hirata I, Kudo SE, Matsumoto T, Ueno F, Watanabe G, Ikegami M, Ito Y, Oba K, Inoue E, Tomotsugu N, Takebayashi T, Sugihara K, Suzuki Y, Watanabe M, Hibi T. Comparison of Targeted vs Random Biopsies for Surveillance of Ulcerative Colitis-Associated Colorectal Cancer. *Gastroenterology* 2016; **151**: 1122-1130 [PMID: 27523980 DOI: 10.1053/j.gastro.2016.08.002]

81 **Allende D**, Elmessiry M, Hao W, DaSilva G, Wexner SD, Bejarano P, Berho M, Al-Qadasi M. Inter-observer and intra-observer variability in the diagnosis of dysplasia in patients with inflammatory bowel disease: correlation of pathological and endoscopic findings. *Colorectal Dis* 2014; **16**: 710-8; discussion 718 [PMID: 24836541 DOI: 10.1111/codi.12667]

82 **Beaugerie L**, Itzkowitz SH. Cancers complicating inflammatory bowel disease. *N Engl J Med* 2015; **372**: 1441-1452 [PMID: 25853748 DOI: 10.1056/NEJMra1403718]

83 **Lai EJ**, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. *Gastrointest Endosc* 2009; **69**: 620-625 [PMID: 19136102 DOI: 10.1016/j.gie.2008.05.057]

84 **Taveira F**, Hassan C, Kaminski MF, Ponchon T, Benamouzig R, Bugajski M, de Castelbajac F, Cesaro P, Chergui H, Goran L, Minelli Grazioli L, Janičko M, Januszewicz W, Lamonaca L, Lenz J, Negreanu L, Repici A, Spada C, Spadaccini M, State M, Szlak J, Veseliny E, Dinis-Ribeiro M, Areia M. The Colon Endoscopic Bubble Scale (CEBuS): a two-phase evaluation study. *Endoscopy* 2020 [PMID: 33285583 DOI: 10.1055/a-1331-4325]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com

ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 June 28; 2(3): 50-94

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 3 June 28, 2021

OPINION REVIEW

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 3 June 28, 2021

Contents

ABOUT COVER

Associate Editor of *Artificial Intelligence in Gastrointestinal Endoscopy*, Mohamed El Kassas, BM BCh, MD, MSc, Associate Professor, Endemic medicine department, Faculty of medicine, Helawan University, Cairo 11448, Egypt. m_elkassas@hq.helwan.edu.eg

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Lin-YuTong Wang*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jin-Lei Wang*.

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Fatih Altintoprak, Sahin Coban, Krish Ragunath

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
June 28, 2021

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 June 28; 2(3): 50-62

DOI: 10.37126/aige.v2.i3.50

ISSN 2689-7164 (online)

OPINION REVIEW

Current situation and prospect of artificial intelligence application in endoscopic diagnosis of *Helicobacter pylori* infection

Yi-Fan Lu, Bin Lyu

ORCID number: Yi-Fan Lu 0000-0002-3623-1332; Bin Lyu 0000-0002-6247-571X.

Author contributions: Lu YF contributed to bibliographic retrieval, data compilation, methodology, software, and manuscript drafting; Lyu B reviewed and proofread the manuscript; all authors contributed to manuscript editing; all authors have read and approved the final manuscript.

Supported by National Natural Science Foundation of China, No. 81770535 and No. 81970470.

Conflict-of-interest statement: The authors declare no conflict of interests for this article.

Yi-Fan Lu, Bin Lyu, Department of Gastroenterology, The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou 310006, Zhejiang Province, China

Corresponding author: Bin Lyu, MD, Chief Doctor, Professor, Department of Gastroenterology, First Affiliated Hospital of Zhejiang Chinese Medical University, No. 54 Youdian Road, Shangcheng District, Hangzhou 310006, Zhejiang Province, China. lvbin@medmail.com.cn

Abstract

With the appearance and prevalence of deep learning, artificial intelligence (AI) has been broadly studied and made great progress in various fields of medicine, including gastroenterology. *Helicobacter pylori* (*H. pylori*), closely associated with various digestive and extradigestive diseases, has a high infection rate worldwide. Endoscopic surveillance can evaluate *H. pylori* infection situations and predict the risk of gastric cancer, but there is no objective diagnostic criteria to eliminate the differences between operators. The computer-aided diagnosis system based on AI technology has demonstrated excellent performance for the diagnosis of *H. pylori* infection, which is superior to novice endoscopists and similar to skilled. Compared with the visual diagnosis of *H. pylori* infection by endoscopists, AI possesses voluminous advantages: High accuracy, high efficiency, high quality control, high objectivity, and high-effect teaching. This review summarizes the previous and recent studies on AI-assisted diagnosis of *H. pylori* infection, points out the limitations, and puts forward prospect for future research.

Key Words: Artificial intelligence; *Helicobacter pylori*; Endoscopy; Diagnosis; Deep learning; Machine learning

Core Tip: In recent years, artificial intelligence (AI) has been rapidly developed and applied in various fields of medicine, including gastroenterology. We witnessed the promising application of AI in endoscopic diagnosis of *Helicobacter pylori* infection. In this review, we summarize the advantages of AI, point out the limitations of current studies, and put forward the direction of future research.

Citation: Lu YF, Lyu B. Current situation and prospect of artificial intelligence application in

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): 0
Grade D (Fair): 0
Grade E (Poor): 0

Received: May 2, 2021
Peer-review started: May 2, 2021
First decision: May 19, 2021
Revised: June 1, 2021
Accepted: June 18, 2021
Article in press: June 18, 2021
Published online: June 28, 2021

P-Reviewer: Yasuda T
S-Editor: Gao CC
L-Editor: Wang TQ
P-Editor: Wang LYT

endoscopic diagnosis of *Helicobacter pylori* infection. *Artif Intell Gastrointest Endosc* 2021; 2(3): 50-62
URL: https://www.wjgnet.com/2689-7164/full/v2/i3/50.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i3.50

INTRODUCTION

Helicobacter pylori (*H. pylori*) is a Gram-negative bacterium that infects the human stomach and is closely associated with a variety of diseases, including chronic gastritis, peptic ulcer, gastric adenocarcinoma, mucosa-associated lymphoid tissue lymphoma, and other digestive diseases, as well as extradigestive diseases of the blood system, nervous system, cardiovascular system, skin, and ophthalmology[1,2]. The International Agency for Research on Cancer has categorized *H. pylori* as a group 1 carcinogen. A recent systematic review and meta-analysis pooling 410879 participants showed that the overall prevalence of *H. pylori* infection worldwide was 44.3% [95% confidence interval (CI): 40.9-47.7][3]. Therefore, accurate diagnosis of *H. pylori* infection is extremely important for the prevention and treatment of related diseases. Currently, various diagnostic methods are available for detecting *H. pylori* infections (non-invasive and invasive methods)[4], but endoscopic evaluation to determine the *H. pylori* infection status is an irreplaceable method, which can assist in the screening of early gastric cancer.

Artificial intelligence (AI) is a technology science that studies and develops the theory, method, technology, and application system that is used to simulate, extend, and expand human intelligence. With the emergence and development of deep learning (DL), the application of AI in medicine has also been enthusiastically explored and extensively studied[5-8]. Numerous research studies, using AI technology to identify or distinguish images in different medical fields including gastroenterology, radiology, neurology, orthopedics, pathology, and ophthalmology, have been published[9].

In this review, we focus on the application of AI in the field of endoscopic diagnosis of *H. pylori* infection and discuss future prospect.

SIGNIFICANCE OF ENDOSCOPIC DIAGNOSIS OF *H. PYLORI* INFECTION

Most patients with gastric cancer have or have had *H. pylori* infection[10,11]. A large number of studies have indicated that the eradication of *H. pylori* can effectively reduce the risk of gastric cancer[12-14]. However, the study conducted by Mabe *et al* [15] showed that people after *H. pylori* eradication still have a higher risk of developing gastric cancer than people who have not been infected with *H. pylori*. Therefore, even after *H. pylori* eradication, regular endoscopic and histological surveillance is strongly recommended[16,17]. In consequence, endoscopic assessment of *H. pylori* infection status (non-infection, past infection, and current infection) has become increasingly important.

The Kyoto classification of gastritis was proposed, which is used to assess the status of *H. pylori* infection and more accurately evaluate the risk of gastric cancer[18]. According to the characteristics of the gastric mucosa under endoscopy, the gastric mucosa can be divided into the following three situations: *H. pylori*-uninfected gastric mucosa, *H. pylori*-infected gastric mucosa, and *H. pylori*-past infected gastric mucosa [18,19]. It should be noted that the Kyoto classification score is the sum of scores for five endoscopic features (atrophy, intestinal metaplasia, enlarged folds, nodularity, and diffuse redness with or without regular arrangement of collecting venules) and ranges from 0 to 8. The scoring system demonstrated excellent ability to evaluate *H. pylori* infection and predict the risk of gastric cancer[20]. However, above endoscopic features do not have objective indicators, and there is the potential for interobserver or intraobserver variability in the optical diagnosis of *H. pylori*-infected mucosa[21]. In other words, for endoscopic diagnosis of *H. pylori* infection, the diagnostic consistency among endoscopists is not ideal. Moreover, professional endoscopists can determine *H. pylori* infection with punctilious visual inspection of the mucosa during endoscopic examination, but novices need a large amount of time to perform this task effectively.

The significance of endoscopic surveillance is not limited to determining whether *H. pylori* is infected, not, or past, but can make an overall evaluation of the stomach. First

of all, the classical Kimura-Takemoto classification is still widely used today to help endoscopists classify the atrophic pattern of the stomach by observing the endoscopic atrophic border[22]. Second, most gastric cancers develop from *H. pylori* associated gastritis. This can occur *via* a multistep pathway of precancerous lesions — in particular, atrophic gastritis, intestinal metaplasia, and dysplasia/intraepithelial neoplasia[16]. We can use histological staging systems such as OLGA and OLGIM to make an assessment of gastric cancer risk by the severity and extent of atrophy and intestinal metaplasia[23-25]. Finally, when one detection method shows *H. pylori* negativity, but there are typical signs of *H. pylori* infection under endoscopy, another different method should be selected for confirmation in this case to avoid missed diagnosis.

WHAT IS AI?

Physicians and endoscopists may be confused about the precise concept of AI, machine learning (ML), and DL. AI is a macro concept with many branches (*e.g.*, Planning and Scheduling, Expert Systems, Multi-Agent Systems, and Evolutionary Computation). In general, there are three approaches to AI: Symbolism (rule based, such as IBM Watson), connectionism (network and connection based, such as DL), and Bayesian (based on the Bayesian theorem)[26]. In AI, computers can imitate humans and display intelligence similar to that of humans.

ML is a subset of AI, which is a method to realize AI. ML is defined as a set of methods that automatically detect patterns in data, and then utilize the uncovered patterns to predict future data or enable decision making under uncertain conditions [27]. ML is approximately divided into supervised and unsupervised methods. Unsupervised learning occurs when the purpose is to identify groups within data according to commonalities, with no *a priori* knowledge of the number of groups or their significance. Supervised learning occurs when training data contain individuals represented as input–output pairs. Input comprises individual descriptors while output comprises outcomes of interest to be predicted — either a class for classification tasks or a numerical value for regression tasks. Then, the supervised ML algorithm learns predictive models that whereafter allow to map new inputs to outputs[28]. The most basic practice of ML [*e.g.*, support vector machine (SVM), random forest, and Gaussian mixture models] is to use algorithms to parse data so as to learn from them, and then make decisions and predictions about events in the real world. Today's ML has made great achievements in computer vision and other fields; however, it has its limitations, requiring a certain amount of manual instruction in the process. The image recognition rate of ML is enough to realize commercialization, but it is still very low in certain fields, which is why image recognition skills are still not as good as human capabilities[29].

DL [*e.g.*, artificial neural network, deep neural network (DNN), convolutional neural network (CNN), and recurrent neural network] is a process in which the computer collects, analyzes, and processes the required data quickly while performing certain tasks, without having to accept the formal data, which is a technique to achieve ML. DL has the characteristics of autonomous learning; once the training data set is provided, the program can extract the key features and quantities by using back-propagation algorithm and changing the internal parameters of each neural network layer, without human instructions[30]. Compared with the conventional hand-crafted algorithm, the recently developed DL algorithm can automatically extract and learn the discriminative features of images, and then classify these images[31]. DL has the potential to automatically detect lesions, classify lesions, prompt differential diagnosis, and write preliminary medical reports, which will be realized in the near future.

CNN is a DNN based on the principle that the visual cortex of the human brain processes and recognizes images, which is now the most popular network architecture for DL for images[29]. CNN uses the multiple network layers (consecutive convolutional layers followed by pooling layers) to extract the key features from an image and provide a final classification through the fully connected layers as the output[30]. Compared to other DL structures, CNN is a prevalent method for image recognition because of its excellent performance in both video and audio applications. For example, CNN performs best in image classification in large image repositories such as ImageNet[32]. Additionally, CNN is easier to train than other DL techniques and has the advantage of using fewer parameters.

In recent years, AI has flourished in the field of gastroenterology, with applications throughout the digestive tract, especially in image recognition and classification. van

der Sommen *et al*[33] reported an automated computer algorithm for the detection of early neoplasia in Barrett's esophagus based on 100 images from 44 patients with Barrett's esophagus. At per-image level, the sensitivity and specificity of the algorithm were both 0.83, and at the patient level, 0.86 and 0.87, respectively. Everson *et al*[34] trained a CNN to classify intrapapillary capillary loops for the real time prediction of early squamous cell cancer of the esophagus, demonstrating strong diagnostic performance with a sensitivity of 93.7% and accuracy of 91.7%, which is comparable to an expert panel of endoscopists. Xu *et al*[35] established a deep CNN system to detect gastric precancerous conditions (including gastric atrophy and intestinal metaplasia) by image-enhanced endoscopy (IEE). In the internal test set, the multicenter external test set, and the prospective video test set, the diagnostic accuracy for gastric atrophy was 0.901, 0.864, and 0.878, and that of intestinal metaplasia was 0.908, 0.859, and 0.898, respectively. To assist endoscopists in distinguishing early gastric cancer, Kanesaka *et al*[36] studied a computer-aided diagnosis (CAD) system utilizing SVM technology to facilitate the use of magnifying narrow band imaging (NBI), which revealed an accuracy of 96.3%, sensitivity of 96.7%, and specificity of 95%. Since capsule endoscopic image viewing and diagnosis is an extremely time-consuming process, Park *et al*[37] developed an AI-assisted reading model based on the Inception-Resnet-V2 model to identify different types of lesions and evaluate the clinical significance of this model. The results showed that the model not only helped the operator to improve the lesion detection rates, but also reduced the reading time. Urban *et al*[38] constructed a deep CNN model, including 8641 images from 2000 patients, to locate and identify colorectal polyps, which revealed an area under the receiver operating characteristic curve of 0.991 and accuracy of 96.4%. Also, several studies have proved the feasibility and prospect of AI-assisted endoscopy in the diagnosis of *H. pylori* infection.

AI-ASSISTED ENDOSCOPIC DIAGNOSIS OF *H. PYLORI* INFECTION

As early as 2004, Huang *et al*[39] independently developed a CAD model based on a refined feature selection with neural network (RFSNN) technique which is planned for predicting *H. pylori*-related gastric histological features. A total of 104 dyspeptic patients were enrolled in this study and all subjects were prospectively evaluated by endoscopy and gastric biopsy. The authors used endoscopic images and histological features of 30 patients (15 with and 15 without *H. pylori* infection) to train the RFSNN model, and then used image parameters of the remaining 74 patients to construct a predictive model of *H. pylori* infection. At the same time, six endoscopic physicians (three novices and three skilled seniors) were invited to predict the histological features of the gastric antrum from endoscopic images. The results showed that the sensitivity and specificity for detecting *H. pylori* infection were 85.4% and 90.9%, respectively, when the RFSNN model included images of the same patient's antrum, body, and cardia for analysis. Together, the accuracy of the six endoscopists in predicting *H. pylori* infection was 67.5%, 64.8%, 72.9%, 74.3%, 79.7%, and 81.1%, respectively (the first three were novices and the second three were skilled elderly). Obviously, the accuracy of RFSNN model in predicting *H. pylori* infection by the antrum images was 85.1% higher than that of endoscopists. Notably, the prediction system has a high sensitivity and specificity in the diagnosis of atrophy and intestinal metaplasia, which was also superior to that of endoscopists. This RFSNN system provides real-time and comprehensive information about the stomach during endoscopy and has the potential to overcome the shortcomings of the localized biopsy. For various reasons, white-light endoscopy was used throughout the study, instead of IEE, which is more conducive to the diagnosis of *H. pylori* infection. As an early study of AI in diagnosing *H. pylori* infection, this paper provides reference data and innovative ideas for subsequent studies.

In 2008, Huang *et al*[40] conducted a further study in the field of AI-assisted endoscopy in the diagnosis of *H. pylori* infection. They designed a CAD system combining SVM and sequential forward floating selection (SFFS) to diagnose gastric histology of *H. pylori* using the features of white-light endoscopic images. This study aimed to use SFFS to select the most suitable feature to describe the relationship between histology and a large number of candidate image features, and then use SVM for classification. A total of 236 dyspepsia patients were enrolled in this study, 130 of whom were defined as *H. pylori*-infected patients using histological examination as the gold standard. The results showed that the accuracy of diagnosing *H. pylori* infection was 87.8%, 87.6%, and 86.7%, respectively, when the SVM with SFFS system was used

to analyze the images of the antrum, body, and cardia. Compared with SVM without SFFS, the SVM with SFFS system had a higher diagnostic accuracy in most cases. This indicates that it is of great significance to use SFFS for screening before the classification of image features, which not only improves the diagnostic accuracy by excluding features with low correlation, but also reduces the time of training and testing system. Furthermore, 1000 repeated tests were carried out on the classification results, which proved the experiment reliability. In addition, the authors compared the new diagnostic system with the previous system[39] that used a neural network with feature selection to detect *H. pylori* infection, and it was shown that the new system had a higher classification rate. It is a pity that both studies classified *H. pylori* infection status only as infected and uninfected, and the authors did not consider cases where the infection disappeared or was eradicated with drugs.

In 2017, Shichijo *et al*[41] developed two deep CNN systems, one based on 32208 unclassified images either positive or negative for *H. pylori* (as a development data set) and the other based on images classified according to eight anatomical locations (cardia, upper body, middle body, lesser curvature, angle, lower body, antrum, and pylorus). Then, the test data set included a total of 11481 images from 397 patients (72 *H. pylori* positive and 325 negative). Patients who tested positive on any of these assays (including blood or urine anti-*H. pylori* immunoglobulin (Ig) G levels, fecal antigen test, or urease breath test) were classified as *H. pylori* positive. To compare the diagnostic performance of the two CNNs, 23 endoscopists were invited to evaluate the test data sets, together. According to their experience, the endoscopists were divided into three groups: "Certified group," "relatively experienced group," and "beginner group". The test results showed that for the first CNN constructed with unclassified images, the area under the receiver operating curve (ROC) curve (AUC) was 0.89 at a cut off value of 0.43. The sensitivity, specificity, accuracy, and diagnostic time of the first CNN were 81.9%, 83.4%, 83.1% and 3.3 min, respectively. These values for the secondary CNN were 88.9%, 87.4%, 87.7%, and 3.2 min, respectively, and the AUC was 0.93 at a cutoff value of 0.34. Furthermore, these values for the overall endoscopists were 79.0%, 83.2%, 82.4%, and 230.1 min, respectively. After statistical analysis, there was no difference in sensitivity, specificity, or accuracy between the first CNN and the 23 endoscopists in the diagnosis of *H. pylori* infection. However, the secondary CNN which was constructed with categorized images according to the location of the stomach was found to have a significantly higher accuracy than the endoscopists (by 5.3%; 95%CI: 0.3-10.2). Besides, the board-certified group was found to have a significantly higher specificity (89.3% *vs* 76.3%, $P < 0.001$) and accuracy (88.6% *vs* 75.6%, $P < 0.001$) than the beginner group. Similarly, a significant difference was observed between the relatively experienced group and the beginner group. In brief, the diagnostic ability of the second CNN is almost as good as that of a skilled endoscopist. In terms of diagnosis time, CNN even completely surpassed the endoscopists. However, still images were adopted to construct CNN algorithm in this study, and whether real-time diagnosis could be realized based on dynamic images remains to be researched.

One weakness of this study was that it did not include the situation after the eradication of *H. pylori*. To address this issue, the authors soon conducted a new study to further elaborate on the role of AI in assessing *H. pylori* infection status. A deep CNN which was constructed by Shichijo *et al*[42] in 2019 was pre-trained and fine-tuned on a dataset of 98564 endoscopic images from 5236 patients (742 *H. pylori*-positive, 3649 *H. pylori*-negative, and 845 *H. pylori*-eradicated). As in the previous study, this AI-based diagnostic system was developed using classified images following eight regions of the stomach (cardia, upper body, middle body, lesser curvature, angle, lower body, antrum, and pylorus). An independent test data set including a total of 23699 images from 847 patients (70 *H. pylori* positive, 493 *H. pylori*-negative, and 284 *H. pylori*-eradicated) was prepared to evaluate the diagnostic accuracy of the constructed CNN. According to the statistical analysis, the proportions of accurate diagnoses were 80% (465/582) for negative, 84% (147/174) for eradicated, and 48% (44/91) for positive. The performance of this diagnostic system is comparable to that of skilled endoscopists who, in one study, diagnosed these statuses in 88.9%, 55.8%, and 62.1% of cases, respectively[43]. Subsequently, the authors assessed the diagnostic ability of CNN for distinguishing *H. pylori* positive from eradicated (excluding *H. pylori* negative patients). Among 70 positive patients, the CNN diagnosed correctly as positive in 46 (66%), while out of 284 eradicated patients, the CNN diagnosed correctly as eradicated in 243 (86%). Nevertheless, this study did not take into account the time after *H. pylori* eradication, but the histological features of atrophic gastritis may disappear a few years after eradication[44]. Then, endoscopic features also change possibly in the diagnosis.

In 2019, Zheng *et al*[45] designed a novel computer-aided decision support system combined with a CNN model (ResNet-50, a state-of-the-art CNN consisting of 50 Layers). This system was expected to be used to retrospectively evaluate *H. pylori* infection based on white-light images (WLI) of the stomach. Totally 1507 patients (11729 gastric images) including 847 with *H. pylori* infection as the derivation cohort were used to train the algorithm. The authors created three DL models: (1) Single gastric image for all gastric images; (2) Single gastric image by different gastric locations (fundus, corpus, angularis, and antrum); and (3) Mmultiple gastric images for the same patient. Afterwards, 452 patients (3755 images) including 310 with *H. pylori* infection as the validation cohort were used to evaluate the diagnostic accuracy CNN for the evaluation of *H. pylori* infection. The evaluation results showed that for a single gastric image, the AUC, sensitivity, specificity, and accuracy were 0.93, 81.4%, 90.1%, and 84.5%, respectively. When evaluating a single gastric image by different anatomical locations, the AUCs from high to low were 0.94 (corpus), 0.91 (angularis), 0.90 (antrum), and 0.82 (fundus). According to statistical analysis, the CNN model using a single corpus image had the highest AUC ($P < 0.01$) compared with the antrum or fundus. More importantly, when multiple stomach images per patient were applied to the CNN model, the AUC, sensitivity, specificity, and accuracy were as high as 0.97, 91.6%, 98.6% and 93.8%, respectively. Consequently, the CNN model using multiple gastric images had a higher AUC compared with a single gastric image ($P < 0.001$) or body gastric image ($P < 0.001$). When selecting endoscopic images to be included in this study, images of poor quality (*i.e.*, blurred images, excessive mucus, food residue, bleeding, and/or insufficient air insufflation) were excluded, which however could not be avoided in the actual operation of endoscopy. Therefore, the CNN's ability to recognize low-quality images needs to be further exploited.

In 2020, Yoshii *et al*[19] established a prediction model based on an ML procedure to prospectively evaluate *H. pylori* infection status (non-infection, past infection, and current infection) and compared it with general assessment by seven well-experienced endoscopists using the Kyoto classification of gastritis. The study recruited a total of 498 subjects (315 non-infection, 104 past infection, and 79 current infection) and the gold standard for determining the *H. pylori* infection status was the history of eradication therapy and the presence of *H. pylori* IgG antibody. The results showed that the overall diagnostic accuracy rate of the seven endoscopists was 82.9%. The diagnostic accuracy of the prediction model without *H. pylori* eradication history was 88.6% and with eradication history was 93.4%. Obviously, the results improved in the model with eradication history. There was no significant difference in diagnostic accuracy between the predictive model and skilled endoscopists. One of the limitations of this study was that only one test method was used to evaluate current status of *H. pylori* infection. In addition, urea breath test or fecal antigen test would evaluate current situation of *H. pylori* infection more surpassingly than that of *H. pylori* IgG antibody levels, especially in patients with an *H. pylori* antibody titer of 3-10 U/mL.

All of the above studies used WLI to build the CAD systems based on AI technology. Besides, some reports have shown the potential of image-enhanced endoscopies (IEEs) in diagnoses of *H. pylori* infection, such as blue laser imaging (BLI), linked color imaging (LCI), and NBI[46-48]. In 2018, Nakashima *et al*[49] built an AI diagnostic system based on a deep CNN algorithm for prospective diagnosis of *H. pylori* infection. A total of 222 subjects (105 *H. pylori*-positive) were recruited and received esophagogastroduodenoscopy and a serum test for *H. pylori* IgG antibodies. A serum *H. pylori* IgG antibody titer ≥ 10 U/mL was considered positive for *H. pylori* infection, while a titer < 3.0 U/mL was considered negative. In addition, subjects with serum *H. pylori* IgG antibody titers between 3.0 and 9.9 U/mL were excluded. In this study, 162 subjects (1944 images) including 75 with *H. pylori* infection were enrolled as a training group for AI training. For the remaining 60 subjects (30 *H. pylori*-positive and 30 *H. pylori*-negative), one WLI, one BLI-bright, and one LCI image of the lesser curvature of the gastric body were collected as a test group to evaluate the diagnostic performance of AI. According to statistical analysis, the AUC, sensitivity, and specificity for WLI were 0.66, 66.7%, and 60.0%, respectively. These indicators were 0.96, 96.7%, and 86.7% for BLI-bright, and 0.95, 96.7%, and 83.3% for LCI, respectively. The AUCs obtained for BLI-bright and LCI were markedly larger than that for WLI ($P < 0.01$). Obviously, this new AI diagnostic system was efficiently adapted to those laser IEEs rather than WLI; hence, it demonstrated an excellent ability to diagnose *H. pylori* infection using the IEEs. It is a pity that patients with a history of *H. pylori* eradication therapy were not included in this study, because this AI system is only an elementary tool and cannot fully evaluate the complex features of the stomach.

In 2020, Yasuda *et al*[21] constructed an automatic diagnosis system based on the SVM algorithm for *H. pylori* infection using LCI images. The authors expected to use this system to retrospectively diagnose *H. pylori* infection and compared its accuracy with that of endoscopists. In this study, endoscopic images of 32 patients (128 images in total) were included as training data, and four images were collected from each patient from the lesser (angle-lower body and middle-upper body) and greater (angle-lower body and middle-upper body) curvature. The diagnosis of *H. pylori* infection was based on more than two different tests: A histological examination, a serum antibody test, a stool antigen test, and/or a 13C-urea breath test. Regarding *H. pylori* infection of the subjects, 14 cases were *H. pylori* positive and 18 were negative. The authors used 525 LCI images from 105 patients (42 *H. pylori* infected, 46 post-eradication, and 17 uninfected) collected from the lesser (angle-lower body and middle-upper body) and greater (angle-lower body and middle-upper body) curvature and the fornix to evaluate the diagnostic capabilities of the system. It was worth noting that for the *H. pylori* post-eradicated subjects, more than 1 year (average of 5.6 years) had passed since *H. pylori* was successfully eradicated after undergoing endoscopy. At the same time, three doctors with different experiences (A, an expert involved in the development of LCI; B, a gastroenterology specialist; and C, a senior resident) also evaluated the same LCI images. The results showed that the accuracy of the AI system, A, B, and C in the diagnosis of *H. pylori* infection was 87.6%, 90.5%, 89.5%, and 86.7%, respectively. Accuracy of the AI system was higher than that of the inexperienced doctor (doctor C), but there was no significant difference between the diagnosis of the doctors and the AI system ($P > 0.05$). According to the sub-analysis of the patients divided with respect to state of *H. pylori* infection, the accuracy of the AI system, doctors A, B ,and C in the diagnosis of *H. pylori* post-eradication were 82.6%, 87.0%, 89.1%, and 76.1%, respectively. According to the sub-analysis of AI diagnosis for each image of stomach area, accuracy of the lesser curvature of the middle-upper body (88.6%) was significantly higher than that of the fornix (69.5%) and the greater curvature of the middle-upper body (73.3%). However, due to the small number of samples included in this study, there may be a risk of large sampling error.

LIMITATIONS AND FUTURE DIRECTION

The above studies show to a great extent that the application of AI in endoscopic diagnosis of *H. pylori* infection is practical, feasible, and promising. The detailed information of these studies is shown in Table 1. Compared with the manual identification and diagnosis by endoscopists, the CAD system based on AI technology has many irreplaceable advantages: (1) High accuracy: According to the current studies, AI is better than novice endoscopists in the diagnosis of *H. pylori* infection in terms of sensitivity, specificity, and accuracy, and is almost comparable to skilled endoscopists; (2) High efficiency: Thanks to today's highly developed computers, AI can classify thousands of endoscopic images in minutes, which can take a great deal of time and energy on the part of endoscopists. At the same time, the efficient image recognition lays a foundation for the real-time diagnosis of *H. pylori* infection under endoscopy; (3) High quality control: Some studies have found that adenoma detection rate decreases gradually with the extension of the working hours of endoscopists. This also suggests that endoscopist fatigue may lead to a decrease in the effectiveness of screening colonoscopy[50,51]. However, the CAD system based on AI technology is not disturbed by external factors and provides excellent quality control; (4) High objectivity: As we all know, it is completely subjective for endoscopists to judge *H. pylori* infection by observing the features of the gastric mucosa under endoscopy. Although the decision-making power is still in the hands of endoscopists, AI assisted endoscopy can help to provide an objective second opinion as a reference[52]; and (5) High-effect teaching: AI is capable of undertaking the teaching work of skilled endoscopists, and provides novices with more accessible, convenient, and objective guidance.

However, the application of AI in endoscopic diagnosis of *H. pylori* infection is still in the preliminary research stage at present, which has many limitations to be overcome. It is promising to put this technology into real clinical practice, but much research and further refinement are needed before that can happen. First of all, all of the above studies are single-center studies and most of them only used images from a single endoscopic device. Different images at different endoscopy centers may not guarantee compatibility and extensibility of the CAD system developed by the researchers and limit the generalization of the results. Next, so far, most of the studies

Table 1 Characteristics of current studies about AI-assisted endoscopic diagnosis of *Helicobacter pylori* infection

Ref.	Type of AI	Type of endoscopy	Training set	Validation set	AUC	Sensitivity (%)	Specificity (%)	Accuracy (%)
Huang *et al*[39], 2004	RFSNN	WLI	30 patients	74 patients	NA	85.4	90.9	NA
Huang *et al*[40], 2008	SVM with SFFS	WLI	236 patients	236 patients	NA	82.6 (antrum); 89.1 (body); 100 (cardia)	94.0 (antrum); 85.8 (body); 72.0 (cardia)	87.8 (antrum); 87.6 (body); 86.7 (cardia)
	SVM without SFFS	WLI	236 patients	236 patients	NA	98.5 (antrum); 98.7 (body); 99.1 (cardia)	70.8 (antrum); 71.5 (body); 70.3 (cardia)	86.3 (antrum); 86.4 (body); 86.0 (cardia)
Shichijo *et al* [41], 2017	CNN (first)	WLI	1750 patients, 32208 images	397 patients, 11481 images	0.89	81.9	83.4	83.1
	CNN (second, constructed according to anatomical locations)	WLI	1750 patients, 32208 images	397 patients, 11481 images	0.93	88.9	87.4	87.7
Shichijo *et al* [42], 2019	CNN	WLI	5236 patients, 98564 images	847 patients, 23699 images	NA	NA	NA	48 (*H. pylori*-positive); 84 (*H. pylori*-eradicated); 80 (*H. pylori*-negative)
Zheng *et al*[45], 2019	CNN (first, single image for all image)	WLI	1507 patients, 76146 images	452 patients, 3755 images	0.93	81.4	90.1	84.5
	CNN (second, single image by different locations)	WLI	1507 patients, 76146 images	452 patients, 3755 images	0.90 (antrum); 0.91 (angularis); 0.94 (corpus); 0.82 (fundus)	76.1 (antrum); 78.8 (angularis); 81.6 (corpus); 72.4 (fundus)	88.5 (antrum); 90.5 (angularis); 92.1 (corpus); 80.5 (fundus)	80.3 (antrum); 82.8 (angularis); 85.6 (corpus); 75.3 (fundus)
	CNN (third, multiple images per patient)	WLI	1507 patients, 76146 images	452 patients, 3755 images	0.97	91.6	98.6	93.8
Yoshii *et al*[19], 2020	ML (model without *H. pylori* eradication history)	WLI	NA	498 patients	NA	91.6 (non-infection); 75.0 (past infection); 59.5 (current infection)	88.6 (non-infection); 89.9 (past infection); 94.7 (current infection)	88.6
	ML (model with *H. pylori* eradication history)	WLI	NA	498 patients	NA	94.0 (non-infection); 94.0 (past infection); 88.1 (current infection)	93.4 (non-infection); 100.0 (past infection); 94.7 (current infection)	93.4
Nakashima *et al* [49], 2018	CNN	WLI	162 patients, 1944 images	60 patients, 60 images	0.66	66.7	60.0	NA
	CNN	BLI-bright	162 patients, 1944 images	60 patients, 60 images	0.96	96.7	86.7	NA
	CNN	LCI	162 patients, 1944 images	60 patients, 60 images	0.95	96.7	83.3	NA
Yasuda *et al* [21], 2020	SVM	LCI	32 patients, 128 images	105 patients, 525 images	NA	90.4	85.7	87.6%

AI: Artificial intelligence; AUC: Area under curve; BLI: Blue laser imaging; CNN: Convolutional neural network; *H. pylori*: *Helicobacter pylori*; LCI: Linked color imaging; ML: Machine learning; NA: Not applicable; RFSNN: Refined feature selection with neural network; SFFS: Sequential forward floating selection; SVM: Support vector machine; WLI: White-light imaging.

have adopted a retrospective method which could be subject to considerable selection bias. As it is, images of high quality or with distinct features of *H. pylori* infection may be preferred for inclusion in studies, which probably lead to exaggerated diagnostic performance of AI and overestimation of the accuracy.

In addition, researchers and endoscopists need to be aware of potential pitfalls and biases in AI research, such as overfitting, spectrum bias, data snooping bias, straw man bias, and P-hacking bias, which can be reduced or eliminated through rigorous research design and appropriate methods[53]. Overfitting occurs when the AI algorithm modulates itself too much on the training dataset and the developed prediction system does not generalize well to new datasets. The translation, rotation, scaling, and clipping of the original endoscopic images to enlarge datasets may be one of the causes of overfitting. Spectrum bias occurs when the training dataset does not adequately represent the range of patients who will be applied in clinical practice (target population)[54]. External validation using independent datasets for model development, collected in a way that minimizes the spectrum bias, is necessary to prove the real performance of an AI algorithm and is important in the verification of any diagnostic or predictive model[55,56]. It is a pity that there is no study that utilized external validation for the performance of an established AI system in this review. It is worth noting that AI has one unavoidable disadvantage that needs to be addressed: "Black box" nature (lack of interpretability), which means that AI technology cannot explain the decision-making processes. But precise interpretability, which can provide diagnostic evidence, assist reduce bias, and build social acceptance, is extremely important in clinical practice. Some methods, such as class activation map, can supplement the "black box" features, hoping to be applied to future research [57].

Besides, some studies only divided *H. pylori* infection status into infected and uninfected, without considering *H. pylori* post-eradication, which is not in line with the clinical reality. Some studies only used single diagnostic method as the gold standard to judge *H. pylori* infection, which will lead to a great loss of diagnostic accuracy. Some studies included a small quantity of subjects and images, which may cause large errors and affect the credibility of the conclusions. IEE has great potential to improve the diagnosis rate of *H. pylori* infection, but there are few studies on the construction of CAD system based on AI using IEE images. What's more, all of the studies in this review were conducted in Asia, and racial difference cannot be avoided.

Finally, before any new technology is introduced into medical practice, ethical problems cannot be avoided and need to be properly solved, including AI technology. AI is not perfect, making no perfect predictions. If a CAD system based on AI technology misdiagnoses or misses diagnoses, who will be held accountable — the endoscopist, medical institution, or manufacturer? What is the attitude of endoscopists

towards the results of AI diagnosis? Question and reject the AI, learn from it, or accept the diagnosis indiscriminately? In the era of AI, how to build a harmonious doctor-patient relationship?

Anyway, in the future, we should expect a "perfect study", a multicenter, large sample, generalized, and prospective study, which has strict inclusion/exclusion criteria, a suitable gold standard for diagnosis and external validation of third-party independent datasets, using high quality datasets to establish a high diagnostic accuracy, and the stability of the CAD system based on AI technology to judge the *H. pylori* infection status. More importantly, ethical principles and laws and regulations related to AI technology need to be improved to protect everyone's legitimate interests. However, it should be pointed out that AI will not completely replace physicians, but will increase diagnostic accuracy, improve diagnostic efficiency, and reduce the burden on physicians. Health care workers need to consider patients' preferences, environment, and ethics before making decisions, which AI cannot replace[58].

CONCLUSION

The era of AI is coming, with both opportunities and challenges. AI is undoubtedly a greatly excellent assistant, which can help endoscopists to evaluate *H. pylori* infection status more quickly, accurately and easily under the endoscope. At the same time, there are some issues as well as ethical considerations that need to be addressed before AI is applied in clinical practice.

REFERENCES

1 **Fischbach W**, Malfertheiner P. Helicobacter Pylori Infection. *Dtsch Arztebl Int* 2018; **115**: 429-436 [PMID: 29999489 DOI: 10.3238/arztebl.2018.0429]

2 **Gravina AG**, Zagari RM, De Musis C, Romano L, Loguercio C, Romano M. *Helicobacter pylori* and extragastric diseases: A review. *World J Gastroenterol* 2018; **24**: 3204-3221 [PMID: 30090002 DOI: 10.3748/wjg.v24.i29.3204]

3 **Zamani M**, Ebrahimtabar F, Zamani V, Miller WH, Alizadeh-Navaei R, Shokri-Shirvani J, Derakhshan MH. Systematic review with meta-analysis: the worldwide prevalence of Helicobacter pylori infection. *Aliment Pharmacol Ther* 2018; **47**: 868-876 [PMID: 29430669 DOI: 10.1111/apt.14561]

4 **Makristathis A**, Hirschl AM, Mégraud F, Bessède E. Review: Diagnosis of Helicobacter pylori infection. *Helicobacter* 2019; **24** Suppl 1: e12641 [PMID: 31486244 DOI: 10.1111/hel.12641]

5 **Esteva A**, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115-118 [PMID: 28117445 DOI: 10.1038/nature21056]

6 **Yeung S**, Downing NL, Fei-Fei L, Milstein A. Bedside Computer Vision - Moving Artificial Intelligence from Driver Assistance to Patient Safety. *N Engl J Med* 2018; **378**: 1271-1273 [PMID: 29617592 DOI: 10.1056/NEJMp1716891]

7 **Gulshan V**, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016; **316**: 2402-2410 [PMID: 27898976 DOI: 10.1001/jama.2016.17216]

8 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. 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 [PMID: 29066576 DOI: 10.1136/gutjnl-2017-314547]

9 **Topol EJ**. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44-56 [PMID: 30617339 DOI: 10.1038/s41591-018-0300-7]

10 **Ono S**, Kato M, Suzuki M, Ishigaki S, Takahashi M, Haneda M, Mabe K, Shimizu Y. Frequency of Helicobacter pylori -negative gastric cancer and gastric mucosal atrophy in a Japanese endoscopic submucosal dissection series including histological, endoscopic and serological atrophy. *Digestion* 2012; **86**: 59-65 [PMID: 22722747 DOI: 10.1159/000339176]

11 **Matsuo T**, Ito M, Takata S, Tanaka S, Yoshihara M, Chayama K. Low prevalence of Helicobacter pylori-negative gastric cancer among Japanese. *Helicobacter* 2011; **16**: 415-419 [PMID: 22059391 DOI: 10.1111/j.1523-5378.2011.00889.x]

12 **Uemura N**, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, Taniyama K, Sasaki N, Schlemper RJ. Helicobacter pylori infection and the development of gastric cancer. *N Engl J Med* 2001; **345**: 784-789 [PMID: 11556297 DOI: 10.1056/NEJMoa001999]

13 **Kamada T**, Hata J, Sugiu K, Kusunoki H, Ito M, Tanaka S, Inoue K, Kawamura Y, Chayama K,

Haruma K. Clinical features of gastric cancer discovered after successful eradication of Helicobacter pylori: results from a 9-year prospective follow-up study in Japan. *Aliment Pharmacol Ther* 2005; **21**: 1121-1126 [PMID: 15854174 DOI: 10.1111/j.1365-2036.2005.02459.x]

14 **Fukase K**, Kato M, Kikuchi S, Inoue K, Uemura N, Okamoto S, Terao S, Amagai K, Hayashi S, Asaka M; Japan Gast Study Group. Effect of eradication of Helicobacter pylori on incidence of metachronous gastric carcinoma after endoscopic resection of early gastric cancer: an open-label, randomised controlled trial. *Lancet* 2008; **372**: 392-397 [PMID: 18675689 DOI: 10.1016/S0140-6736(08)61159-9]

15 **Mabe K**, Takahashi M, Oizumi H, Tsukuma H, Shibata A, Fukase K, Matsuda T, Takeda H, Kawata S. Does Helicobacter pylori eradication therapy for peptic ulcer prevent gastric cancer? *World J Gastroenterol* 2009; **15**: 4290-4297 [PMID: 19750572 DOI: 10.3748/wjg.15.4290]

16 **Sugano K**, Tack J, Kuipers EJ, Graham DY, El-Omar EM, Miura S, Haruma K, Asaka M, Uemura N, Malfertheiner P; faculty members of Kyoto Global Consensus Conference. Kyoto global consensus report on Helicobacter pylori gastritis. *Gut* 2015; **64**: 1353-1367 [PMID: 26187502 DOI: 10.1136/gutjnl-2015-309252]

17 **Correa P**. A human model of gastric carcinogenesis. *Cancer Res* 1988; **48**: 3554-3560 [PMID: 3288329]

18 **Haruma K**, Kato M, Inoue K, Murakami K, Kamada T. Kyoto classification of gastritis. Tokyo: Nihon Medical Center, 2017

19 **Yoshii S**, Mabe K, Watano K, Ohno M, Matsumoto M, Ono S, Kudo T, Nojima M, Kato M, Sakamoto N. Validity of endoscopic features for the diagnosis of Helicobacter pylori infection status based on the Kyoto classification of gastritis. *Dig Endosc* 2020; **32**: 74-83 [PMID: 31309632 DOI: 10.1111/den.13486]

20 **Toyoshima O**, Nishizawa T, Koike K. Endoscopic Kyoto classification of *Helicobacter pylori* infection and gastric cancer risk diagnosis. *World J Gastroenterol* 2020; **26**: 466-477 [PMID: 32089624 DOI: 10.3748/wjg.v26.i5.466]

21 **Yasuda T**, Hiroyasu T, Hiwa S, Okada Y, Hayashi S, Nakahata Y, Yasuda Y, Omatsu T, Obora A, Kojima T, Ichikawa H, Yagi N. Potential of automatic diagnosis system with linked color imaging for diagnosis of Helicobacter pylori infection. *Dig Endosc* 2020; **32**: 373-381 [PMID: 31398276 DOI: 10.1111/den.13509]

22 **Kimura K**, Takemoto T. An Endoscopic Recognition of the Atrophic Border and its Significance in Chronic Gastritis. *Endoscopy* 1969; **1**: 3

23 **Rugge M**, Meggio A, Pennelli G, Piscioli F, Giacomelli L, De Pretis G, Graham DY. Gastritis staging in clinical practice: the OLGA staging system. *Gut* 2007; **56**: 631-636 [PMID: 17142647 DOI: 10.1136/gut.2006.106666]

24 **Capelle LG**, de Vries AC, Haringsma J, Ter Borg F, de Vries RA, Bruno MJ, van Dekken H, Meijer J, van Grieken NC, Kuipers EJ. The staging of gastritis with the OLGA system by using intestinal metaplasia as an accurate alternative for atrophic gastritis. *Gastrointest Endosc* 2010; **71**: 1150-1158 [PMID: 20381801 DOI: 10.1016/j.gie.2009.12.029]

25 **Rugge M**, Correa P, Di Mario F, El-Omar E, Fiocca R, Geboes K, Genta RM, Graham DY, Hattori T, Malfertheiner P, Nakajima S, Sipponen P, Sung J, Weinstein W, Vieth M. OLGA staging for gastritis: a tutorial. *Dig Liver Dis* 2008; **40**: 650-658 [PMID: 18424244 DOI: 10.1016/j.dld.2008.02.030]

26 **Lee JG**, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol* 2017; **18**: 570-584 [PMID: 28670152 DOI: 10.3348/kjr.2017.18.4.570]

27 **Robert C**. Machine Learning, a Probabilistic Perspective. *Chance* 2014; **27**: 62-63 [DOI: 10.1080/09332480.2014.914768]

28 **Shalev-Shwartz S**, Ben-David S. Understanding machine learning: From theory to algorithms. Cambridge university press, 2014

29 **Min JK**, Kwak MS, Cha JM. Overview of Deep Learning in Gastrointestinal Endoscopy. *Gut Liver* 2019; **13**: 388-393 [PMID: 30630221 DOI: 10.5009/gnl18384]

30 **Takiyama H**, Ozawa T, Ishihara S, Fujishiro M, Shichijo S, Nomura S, Miura M, Tada T. Automatic anatomical classification of esophagogastroduodenoscopy images using deep convolutional neural networks. *Sci Rep* 2018; **8**: 7497 [PMID: 29760397 DOI: 10.1038/s41598-018-25842-6]

31 **Mori Y**, Kudo SE, Mohmed HEN, Misawa M, Ogata N, Itoh H, Oda M, Mori K. Artificial intelligence and upper gastrointestinal endoscopy: Current status and future perspective. *Dig Endosc* 2019; **31**: 378-388 [PMID: 30549317 DOI: 10.1111/den.13317]

32 **Krizhevsky A**, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. *ACM* 2017; **60**: 84-90 [DOI: 10.1145/3065386]

33 **van der Sommen F**, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, Bergman JJ, de With PH, Schoon EJ. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. *Endoscopy* 2016; **48**: 617-624 [PMID: 27100718 DOI: 10.1055/s-0042-105284]

34 **Everson MA**, Garcia-Peraza-Herrera L, Wang HP, Lee CT, Chung CS, Hsieh PH, Chen CC, Tseng CH, Hsu MH, Vercauteren T, Ourselin S, Kashin S, Bisschops R, Pech O, Lovat L, Wang WL, Haidry RJ. A clinically interpretable convolutional neural network for the real-time prediction of early squamous cell cancer of the esophagus: comparing diagnostic performance with a panel of expert European and Asian endoscopists. *Gastrointest Endosc* 2021 [PMID: 33549586 DOI: 10.1016/j.gie.2021.01.043]

35 **Xu M**, Zhou W, Wu L, Zhang J, Wang J, Mu G, Huang X, Li Y, Yuan J, Zeng Z, Wang Y, Huang L,

Liu J, Yu H. Artificial intelligence in diagnosis of gastric precancerous conditions by image-enhanced endoscopy: a multicenter, diagnostic study (with video). *Gastrointest Endosc* 2021 [PMID: 33722576 DOI: 10.1016/j.gie.2021.03.013]

36 **Kanesaka T**, Lee TC, Uedo N, Lin KP, Chen HZ, Lee JY, Wang HP, Chang HT. Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc* 2018; **87**: 1339-1344 [PMID: 29225083 DOI: 10.1016/j.gie.2017.11.029]

37 **Park J**, Hwang Y, Nam JH, Oh DJ, Kim KB, Song HJ, Kim SH, Kang SH, Jung MK, Jeong Lim Y. Artificial intelligence that determines the clinical significance of capsule endoscopy images can increase the efficiency of reading. *PLoS One* 2020; **15**: e0241474 [PMID: 33119718 DOI: 10.1371/journal.pone.0241474]

38 **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078. e8 [PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037]

39 **Huang CR**, Sheu BS, Chung PC, Yang HB. Computerized diagnosis of Helicobacter pylori infection and associated gastric inflammation from endoscopic images by refined feature selection using a neural network. *Endoscopy* 2004; **36**: 601-608 [PMID: 15243882 DOI: 10.1055/s-2004-814519]

40 **Huang CR**, Chung PC, Sheu BS, Kuo HJ, Popper M. Helicobacter pylori-related gastric histology classification using support-vector-machine-based feature selection. *IEEE Trans Inf Technol Biomed* 2008; **12**: 523-531 [PMID: 18632332 DOI: 10.1109/TITB.2007.913128]

41 **Shichijo S**, Nomura S, Aoyama K, Nishikawa Y, Miura M, Shinagawa T, Takiyama H, Tanimoto T, Ishihara S, Matsuo K, Tada T. Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images. *EBioMedicine* 2017; **25**: 106-111 [PMID: 29056541 DOI: 10.1016/j.ebiom.2017.10.014]

42 **Shichijo S**, Endo Y, Aoyama K, Takeuchi Y, Ozawa T, Takiyama H, Matsuo K, Fujishiro M, Ishihara S, Ishihara R, Tada T. Application of convolutional neural networks for evaluating Helicobacter pylori infection status on the basis of endoscopic images. *Scand J Gastroenterol* 2019; **54**: 158-163 [PMID: 30879352 DOI: 10.1080/00365521.2019.1577486]

43 **Watanabe K**, Nagata N, Shimbo T, Nakashima R, Furuhata E, Sakurai T, Akazawa N, Yokoi C, Kobayakawa M, Akiyama J, Mizokami M, Uemura N. Accuracy of endoscopic diagnosis of Helicobacter pylori infection according to level of endoscopic experience and the effect of training. *BMC Gastroenterol* 2013; **13**: 128 [PMID: 23947684 DOI: 10.1186/1471-230X-13-128]

44 **Kodama M**, Murakami K, Okimoto T, Sato R, Uchida M, Abe T, Shiota S, Nakagawa Y, Mizukami K, Fujioka T. Ten-year prospective follow-up of histological changes at five points on the gastric mucosa as recommended by the updated Sydney system after Helicobacter pylori eradication. *J Gastroenterol* 2012; **47**: 394-403 [PMID: 22138891 DOI: 10.1007/s00535-011-0504-9]

45 **Zheng W**, Zhang X, Kim JJ, Zhu X, Ye G, Ye B, Wang J, Luo S, Li J, Yu T, Liu J, Hu W, Si J. High Accuracy of Convolutional Neural Network for Evaluation of Helicobacter pylori Infection Based on Endoscopic Images: Preliminary Experience. *Clin Transl Gastroenterol* 2019; **10**: e00109 [PMID: 31833862 DOI: 10.14309/ctg.0000000000000109]

46 **Nishikawa Y**, Ikeda Y, Murakami H, Hori SI, Hino K, Sasaki C, Nishikawa M. Classification of atrophic mucosal patterns on Blue LASER Imaging for endoscopic diagnosis of Helicobacter pylori-related gastritis: A retrospective, observational study. *PLoS One* 2018; **13**: e0193197 [PMID: 29596454 DOI: 10.1371/journal.pone.0193197]

47 **Takeda T**, Asaoka D, Nojiri S, Nishiyama M, Ikeda A, Yatagai N, Ishizuka K, Hiromoto T, Okubo S, Suzuki M, Nakajima A, Nakatsu Y, Komori H, Akazawa Y, Nakagawa Y, Izumi K, Matsumoto K, Ueyama H, Sasaki H, Shimada Y, Osada T, Hojo M, Kato M, Nagahara A. Linked Color Imaging and the Kyoto Classification of Gastritis: Evaluation of Visibility and Inter-Rater Reliability. *Digestion* 2020; **101**: 598-607 [PMID: 31302654 DOI: 10.1159/000501534]

48 **Okubo M**, Tahara T, Shibata T, Nakamura M, Kamiya Y, Yoshioka D, Maeda Y, Yonemura J, Ishizuka T, Arisawa T, Hirata I. Usefulness of magnifying narrow-band imaging endoscopy in the Helicobacter pylori-related chronic gastritis. *Digestion* 2011; **83**: 161-166 [PMID: 21266810 DOI: 10.1159/000321799]

49 **Nakashima H**, Kawahira H, Kawachi H, Sakaki N. Artificial intelligence diagnosis of *Helicobacter pylori* infection using blue laser imaging-bright and linked color imaging: a single-center prospective study. *Ann Gastroenterol* 2018; **31**: 462-468 [PMID: 29991891 DOI: 10.20524/aog.2018.0269]

50 **Lee CK**, Cha JM, Kim WJ. Endoscopist Fatigue May Contribute to a Decline in the Effectiveness of Screening Colonoscopy. *J Clin Gastroenterol* 2015; **49**: e51-e56 [PMID: 25110871 DOI: 10.1097/MCG.0000000000000175]

51 **Lee A**, Iskander JM, Gupta N, Borg BB, Zuckerman G, Banerjee B, Gyawali CP. Queue position in the endoscopic schedule impacts effectiveness of colonoscopy. *Am J Gastroenterol* 2011; **106**: 1457-1465 [PMID: 21448145 DOI: 10.1038/ajg.2011.87]

52 **Hoogenboom SA**, Bagci U, Wallace MB. AI in gastroenterology. The current state of play and the potential. How will it affect our practice and when? *Tech Gastrointest Endosc* 2019; 150634

53 **England JR**, Cheng PM. Artificial Intelligence for Medical Image Analysis: A Guide for Authors and Reviewers. *AJR Am J Roentgenol* 2019; **212**: 513-519 [PMID: 30557049 DOI: 10.2214/AJR.18.20490]

54 **Park SH**, Han K. Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology* 2018; **286**: 800-809 [PMID: 29309734 DOI: 10.1148/radiol.2017171920]

55 **Steyerberg EW**. Overfitting and optimism in prediction models. In: Steyerberg EW. Clinical Prediction Models. Cham: Springer, 2019: 95-112

56 **Yang YJ**, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; **25**: 1666-1683 [PMID: 31011253 DOI: 10.3748/wjg.v25.i14.1666]

57 **Philbrick KA**, Yoshida K, Inoue D, Akkus Z, Kline TL, Weston AD, Korfiatis P, Takahashi N, Erickson BJ. What Does Deep Learning See? *AJR Am J Roentgenol* 2018; **211**: 1184-1193 [PMID: 30403527 DOI: 10.2214/AJR.18.20331]

58 **Le Berre C**, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, Danese S, Peyrin-Biroulet L. Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology* 2020; **158**: 76-94. e2 [PMID: 31593701 DOI: 10.1053/j.gastro.2019.08.058]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 June 28; 2(3): 63-70

DOI: 10.37126/aige.v2.i3.63

ISSN 2689-7164 (online)

MINIREVIEWS

Progress and prospects of artificial intelligence in colonoscopy

Rui-Gang Wang

ORCID number: Rui-Gang Wang 0000-0002-9053-5329.

Author contributions: The author Wang RG completed the review independently.

Supported by Digestive Medical Coordinated Development Center of Beijing Municipal Administration of Hospitals, No. XXT17.

Conflict-of-interest statement: There is no conflict of interest in this article.

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Rui-Gang Wang, Department of Gastroenterology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing 102218, China

Corresponding author: Rui-Gang Wang, Doctor, Department of Gastroenterology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, No. 168 Litang Road, Changping District, Beijing 102218, China. doctorwrg@163.com

Abstract

Artificial intelligence (AI) is a branch of computer science. As a new technological science, it mainly develops and expands human intelligence through the research of intelligence theory, methods and technology. In the medical field, AI has bright application prospects (for example: imaging, diagnosis and treatment). The exploration of robotic gastroscopy and colonoscopy systems is not only a bold attempt, but also an inevitable trend of AI in the development of digestive endoscopy in the future. Based on the current research findings, this article summarizes the research progress of colonoscopy, and looking forward for the application of AI in colonoscopy.

Key Words: Artificial intelligence; Colonoscopy; Application; Gastrointestinal; Endoscopy

Core Tip: Artificial intelligence is a new technological science that studies and develops theories, methods, technologies and application systems for simulating and expanding human intelligence. This article will systematically review the exploration and application of artificial intelligence technology in colonoscopy, and look forward to the development direction of intelligent colonoscopy.

Citation: Wang RG. Progress and prospects of artificial intelligence in colonoscopy. *Artif Intell Gastrointest Endosc* 2021; 2(3): 63-70

URL: https://www.wjgnet.com/2689-7164/full/v2/i3/63.htm

DOI: https://dx.doi.org/10.37126/aige.v2.i3.63

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: April 29, 2021
Peer-review started: April 29, 2021
First decision: May 19, 2021
Revised: May 29, 2021
Accepted: June 23, 2021
Article in press: June 23, 2021
Published online: June 28, 2021

P-Reviewer: Bedny I
S-Editor: Gao CC
L-Editor: Filipodia
P-Editor: Wang LYT

INTRODUCTION

Artificial intelligence (AI) is a new technological science that studies and develops theories, methods, technologies and application systems for simulating and expanding human intelligence. It relates to many fields, for instance, computer science, cybernetics, information theory, and neuroscience. The first AI seminar at Dartmouth College in 1956 marked the birth of the AI, but the development of AI has experienced several ups and downs. AI has achieved results both theoretically and practically in these cycles. It has made solid progress in the world, especially when scientists made breakthrough progress in deep learning.

In its more than 60 years of development, AI has been used in computer vision, natural language processing, data mining, automatic speech recognition. The applications of intelligent robot, automatic programming, and expert systems are becoming increasingly mature, making AI one of the three cutting-edge technologies in the 21st century.

AI is hailed as the stethoscope of the 21st century[1]. With the strengthening of people's health awareness, preventive and precise treatments have been paid more attention at the same time. The improvement of medical standards and the improvement of medical equipment have made the process of patients' visits produce increasingly medical data. Image recognition, speech/semantic recognition, and expert system have received more and more attention in the medical field, smart medical products have gradually emerged[2-4]. A large amount of image data and diagnostic data are used to simulate the mind and diagnostic process of medical experts especially in the field of medical image recognition, AI is expected to partially replace traditional empirical diagnosis so as to provide a more reliable diagnosis and treatment plan.

AI HELPS BREAK THROUGH THE BOTTLENECK OF COLONOSCOPY

In recent years, the incidence of colorectal adenoma, colorectal cancer, and inflammatory bowel disease has increased significantly[5-7], causing great harm to human's health. Colonoscopy is the first choice for the diagnosis and treatment of colorectal diseases. It can not only intuitively judge the nature of the lesion, but also obtain biopsy specimens for pathological diagnosis. Colonoscopy is of great significance, especially in preventing and treating colorectal cancer, as it can be used to screen and follow up high-risk groups in patients who are asymptomatic. We can greatly reduce the incidence of colorectal cancer by adopting corresponding treatments according to the condition, and achieve the purpose of primary prevention. Even if colorectal lesions develop to the early stage of cancer, the 5-year survival rate of endoscopic treatment can still exceed 90%[6].

Studies have found that gradual expansion of colorectal cancer screening in asymptomatic populations and the early diagnosis promotion have extremely important socio-economic significance[8-10]. The popularization of colonoscopy screening among high-risk populations is restricted by the hard operation, excessive physical exertion, and limitation of technical inheritance, which has caused bottlenecks. At this time, the development and maturity of AI technology provides new ideas and possibilities for breaking through these bottlenecks.

RESEARCH ON THE MECHANISM OF COLONOSCOPY INTO LOOPS AND UNLOOPS

According to the anatomical characteristics of the intestine, the ascending colon, descending colon and upper rectum, which are straighter and smaller in extension, are generally easier to pass with colonoscopy. However, the transverse colon and sigmoid colon are in a free state, with longer mesentery and larger mobility, which can easily cause loops. Common types of loops in the sigmoid colon include N loops, α loops, reverse α loops, and atypical loops, while the common types of loops in the transverse colon include deep loops/dangling loops, deep large γ loops, and inverted splenic loops[11]. Usually, the time for a skilled endoscopist to enter the cecum is about 4-6 minutes, but someone who have difficulty in this process may not be able to reach it, even if the operation time is more than 1 h[12].

In view of the factors of patients who develop a loop during colonoscopy, experts have conducted many studies which found that factors including long-term constipation, abdominal surgery history, female, body mass index is lower or higher than normal, the volume of visceral fat tissue is low and the proficiency of colonoscopy directly affect the formation of intestinal loops[13] (Figure 1).

The successful removal of the loop is key for a colonoscopy to reach the cecum, and it is necessary for the endoscopist to be able to observe and monitor the shape of colonoscopy in order to overcome this technical difficulty. With the continuous advancement of colonoscopy accessories, magnetic endoscopic imaging (MEI), a real-time three-dimensional imaging colonoscopy-assisted positioning technology, has become an effective tool for observing the shape of the colonoscopy in human body [14]. There is a meta-analysis that summarizes 8 randomized controlled trials and contains 2967 patients which compares cecal intubation rates and times, sedation dose, abdominal pain scores and the use of ancillary maneuvers between MEI and standard colonoscopy. The conclusion is that compared with traditional technique, MEI has an advantages in cecal intubation rate, but MEI did not have any distinct advantages for cecal intubation time and lower pain scores[15] (DW1). The variable stiffness of the colonoscopy body, flexible tubing, and Responsive Insertion Technology (RIT)[16,17] make the inspection equipment more maneuverable. Prieto-de-Frías *et al*[17] and Pasternak *et al*[18] studied the application of RIT technology in reducing discomfort and pain during colonoscopy insertion. The results showed that the RIT group shortened the cecal intubation time, decrease intestinal loop formation, lower manual pressure of abdomen and decrease discomfort or pain of patients. Although RIT technology has shown good application prospects, it still relies on the experience of unwinding of endoscopists, some examinations are time-consuming and patients cannot achieve a good medical result.

MEI and RIT technology are an improvement of traditional colonoscopy in response to the actual problems in the endoscopy process. AI can explore the images of MEI technology in guiding colonoscopy. Applying deep learning to analyze a large number of unloop images, it is possible in the future to form a complete set of loop prediction and unlooping strategies system. The RIT technology can automatically adjust the bending angle of the intestinal cavity by sensing the degree of curvature of the endoscopic body, and minimize the formation of acute angles. These measures help to reduce the traction of the colonoscopy on the mesentery and the damage to the intestinal mucosa, and achieve the purpose of reducing the pain and injury of the patient during the colonoscopy. In general, MEI and RIT technologies provide useful explorations for the gradual migration of colonoscopy from artificial to intelligent (DW2).

COMBINATION OF COLONOSCOPY AND AI

Traditional research methods have limitations, such as multi-factors, complex variables, interrelationships, descriptive difficulties and quantitative mechanisms. It is urgent to introduce new ideas and methods to solve these problems. It can be described with a simplified model by demonstrating whether the colonoscopy is looped, and providing the corresponding unlooping strategy, as we mentioned above. The operation of the colonoscopy handle and insertion part by the endoscopist can be regarded as an input function. Analyze the correspondence between the data of the input function under the loop condition and the corresponding results of loop and unloop in a large number of cases, also fitting the unloop strategy function to assist the doctor in decision-making through the intelligent system. MEI and other technologies can display the posture of the colonoscopy in the intestine in real time, and wearable pressure sensor device can generate a series of mechanical data. A specific neural network model can be constructed to synthesize a loop-free strategy function by analyzing large amounts of data. We look forward to the AI-assisted system will be able to realize a loopless and painless colonoscopy in the future (DW3).

COLONOSCOPY FOR SMART MEDICINE

Smart medicine is the application of AI to improve the ability of medical services, which is the trend of future medical advancement. Smart medical care is to create a regional medical information platform for health records and use advanced Internet of Things technology to realize the interaction among patient-medical staff, institutions

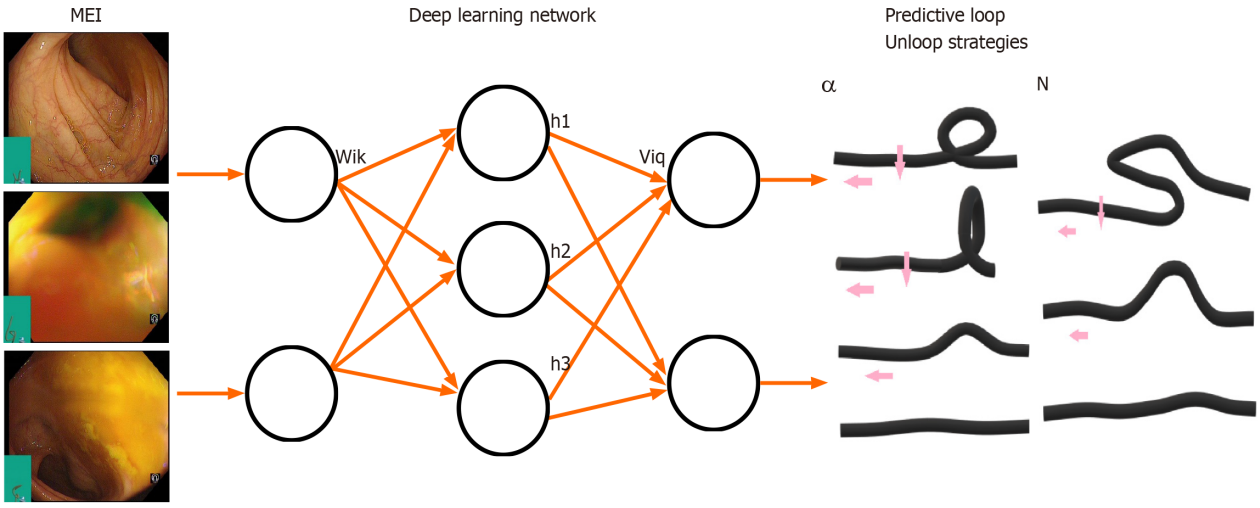

Figure 1 The idea of using magnetic endoscopic imaging to guide endoscopists in colonoscopy. MEI: Magnetic endoscopic imaging.

and equipment for achieving informatization gradually. Intelligent medicine cannot be separated from AI technology. On the basis of digital medicine, internet medicine and mobile medicine, smart medicine is gradually taking shape.

The emergence of smart medicine provides a new feasible path to solve the outstanding problems that restrict the medical development. Intelligent medical care plays an important role in science, it not only changes the traditional diagnosis and treatment methods but also improves the accuracy and efficiency, in addition, it relies on the advanced algorithms and powerful computing power of AI technology to significantly increase the success rate of medical innovation research and development and shorten time. In addition, smart medicine can also solve social problems, such as insufficient medical resources, unbalanced regional distribution, costs, personalized medical services, and respond to aging and chronic disease diagnosis and treatment needs. With the development of smart medical technology, AI can completely assist doctors in such arduous diagnoses in future, for example pathological diagnosis, laboratory test diagnosis, and imaging diagnosis.

COLONOSCOPY CONTINUUM ROBOT-ASSIST SYSTEM

Regarding the colonoscopy continuum robot-assisted system, some scholars have studied structural design, passability, compliance control based on force perception, and multi-motor control system design. Lee *et al*[19] proposed a caterpillar-like flexible self-propelled colonoscopy robot, which can effectively corner bends and conducted clinical trials, while Breedveld proposed a colonoscopy robot movement method based on a rollable doughnut[20]. Scholars research on the relevant working environment and clinical experiment results of the colonoscopy continuum robot assistance system, the flexible arbitrary bending of the colonoscopy assistance system, the exploration of the biomimetic and the continuum robot design, which are the most irreplaceable (DW4) part of the robot-assisted colonoscopy system, its structure and design provide an important reference.

FLEXIBLE ENDOSCOPY CONTROL ROBOT

In December 1998, the first Da Vinci Robot-Assisted Surgery System came out. In June 2000, the Da Vinci Robot-Assisted Surgery System became the first automatic mechanical system approved by the Food and Drug Administration for laparoscopic surgery. At present, the system is widely used. In 2017, the flexible endoscopy manipulation robot developed by the General Hospital of the Chinese People's Liberation Army successfully carried out clinical applications. It surpassed the traditional endoscopy operation method in terms of coordinated operation of multiple degrees of freedom of the endoscopy and quantitative display of operating parameters, and laid the foundation for high-quality standardized operation and

internet medical treatment.

The research on small soft robots with multi-mode motion published by Hu has attracted widespread attention[21]. The article pointed out that the soft robot has bright prospects in the fields of bioengineering and minimally invasive treatment. They have greater potential to achieve high maneuverability through multi-channel motion because small soft robots have a higher degree of freedom than rigid robots. We can expect that these small flexible robots are equipped with camera devices to produce soft motion which is similar to worms that can move in the human digestive tract and has better control and operability than the magnetic-control capsule endoscopy.

At present, there are no reports on the use of flexible endoscopic robots for endoscopic treatment, and the author believes that the reason is that endoscopic treatment is different from examination. Endoscopic treatment have higher requirements for the operation technology, including horizontal and vertical joint movement of the endoscope handle to achieve rotation, control colonoscopy and handle strength during the treatment (DW5). The grasp of the patients' breathing and coordination with its movement are relatively subtle that are difficult to achieve at this stage. However, with the accumulation of quantitatively analyzed endoscopic operation data and the construction of software endoscopic operation strategy functions, combined with powerful algorithms and machine learning, AI will continue to improve the existing colonoscopy equipment, accessories and instruments in the future. At the same time, it may partly replace manual labor, reduce medical costs and improve efficiency.

APPLICATION STATUS OF AI IN COLONOSCOPY IMAGE RECOGNITION

With the progress of colonoscopy operation technology and endoscopic imaging technology, especially magnifying endoscopy has achieved remarkable results in the detection of fine structure on the surface of colorectal tumors. It should be pointed out that the development of electronic staining endoscopy is extremely rapid, such as narrowband imaging technology (NBI), flexible spectral imaging color enhancement technology (FICE) and i-Scan digital contrast technology (iSCAN), *etc.* (DW6). These imaging technologies can highlight the mucosal surface structure or capillary morphology by switching between different wavelengths of light, clearly observe the boundary and scope of the lesion, and obtain a visual effect similar to chromoendoscopy.

Depth research for colonoscopy image recognition has already started, using specific data sets and special deep learning network structure models to establish a labeled colonic lesion image data set to provide technical support for intelligent image recognition of colonoscopy images. Computer-aided diagnosis analysis used for accurately classify neoplastic/hyperplastic, adenoma/non-adenomas colorectal polyps found that the system have a classification accuracy rate above 90%, and the diagnosis time required is decreased compared with endoscopy experts and non-experts[4,22-24].

The dynamic recognition system decomposes the real-time video of the colonoscopy into a continuous picture. The deep learning neural network is used for the recognition of the marked images, and the fine recognition of each image is carried out to realize the purpose of automatically discovering and classifying the lesions. Mori *et al*[25] used deep learning models to analyze colonoscopy videos to classify adenomatous and hyperplastic polyps in real time, the results find that the accuracy of the AI model is 94%, the sensitivity and the specificity is 98% and 83% respectively (Figure 2).

We expect that AI combined with white light, chromoendoscopy and magnifying endoscopy will greatly reduce the time spent on diagnosis and treatment in the future, thereby providing great help for the clinical and scientific research of gastrointestinal diseases.

APPLICATION OF AI IN CAPSULE ENDOSCOPY

In recent years, the rapid development of capsule endoscopy technology, especially the appearance of magnetron capsule endoscopy, which has realized the controllability of the capsule endoscopy on some extent. The emergence of capsule endoscopy has made up for the insufficiency of gastroscopy and colonoscopy, the patients acceptance is high because of the whole examination process is painless. Nowadays,

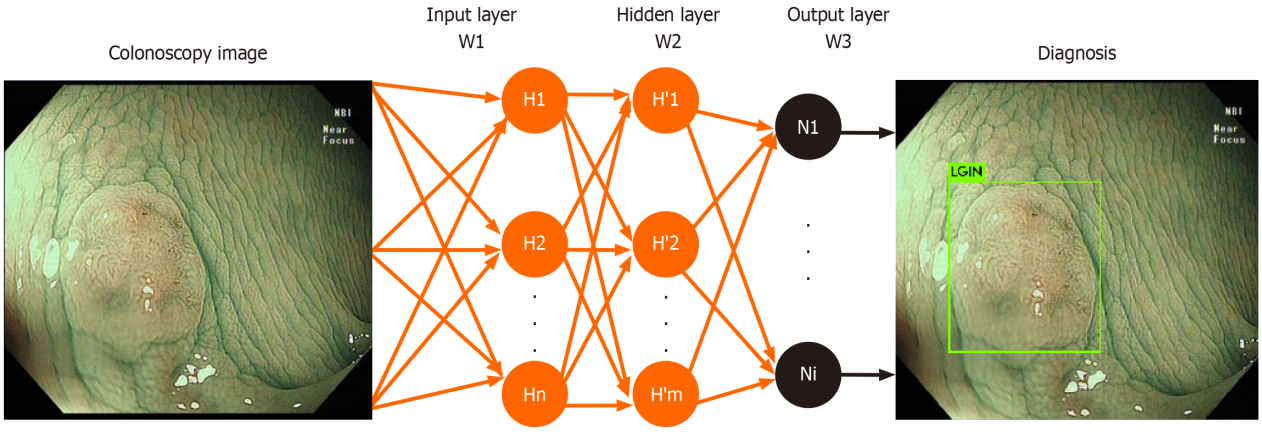

Figure 2 Artificial intelligence diagnosis system for colonoscopy lesions based on deep learning.

the application of capsule endoscopy is mostly focused on discovery of small bowel disease, for example bleeding.

AI is widely used in capsule endoscopy technology. The pixels are grouped by super pixel segmentation, the red ratio in the RGB space is used to extract the features of each super pixel, and these things are input into Support Vector Machines (SVM) for classification for intelligent recognition of capsule endoscopic bleeding. The specificity of the experimental results is 83%-98%, and the sensitivity is 94%-99%[26, 27].

In order to identify polyps in capsule endoscopy images, Yuan and Meng[28] proposed a new complex feature learning method, which is a stacked sparse autoencoder with image manifold constraint. This method introduces multiple image constraints force images in the same category to share similar learning features and keep them, so the learned features retain a large number of differences and small internal differences in the images. The results show that the average overall recognition accuracy of this method is 98%, and could be further utilized in the clinical trials to help physicians from the tedious image reading work.

THE PROBLEMS FACED BY AI IN THE APPLICATION OF COLONOSCOPY

The development of depth research has enabled AI to achieve fruitful results in many aspects. However, there is no major breakthrough in the theory that AI follows, and the methods from supervised learning to unsupervised learning are still being explored. Therefore, looking for in-depth theoretical explanations is an important issue that must be solved in the development of the studies. In addition, deep learning generally requires a large amount of data, but not all applications have the conditions for it. Therefore, how to realize traditional knowledge expression and data-driven knowledge learning is an important research direction in the future. Furthermore, the neural network model needs to be adapted to transfer the learned knowledge to new conditions and environments in order to acquire the ability to solve many practical problems from a small number of learning samples. Finally, the method of machine learning is determined according to the functional relationship between the data and the target, a "deep forest" learning method, with a comparable setting proposed by Zhou and Feng[29], achieved a considerable or even better than deep neural networks.

In the field of colonoscopy image recognition, experts and scholars have made very useful explorations on the intelligent recognition of colorectal lesions, but most of them are limited to judge colorectal polyps. To achieve the integration of doctors and patients with auxiliary examination equipment, it is necessary to further expand the colorectal lesion image data set and the types of diseases involved. It must be pointed out that the endoscopic manifestations of colorectal diseases are various, the same disease often manifests differences in different periods and different diseases have very little difference in a specific period, and pathological diagnosis is still the gold standard.

CONCLUSION

In short, AI in colonoscopy has significant social benefits and bright application prospects, and it is foreseeable that smart medicine is an inevitable trend in medical development. Based on previous research, integrating colonoscopy's loop factors, unlooping strategies, active lesion capture and recognition, and assistive robotics technology, we have reason to believe that the future smart colonoscopy system will bring a revolution, and promote the diagnosis and treatment of colorectal diseases, especially the widespread development of colorectal cancer screening for the benefit of mankind.

ACKNOWLEDGEMENTS

In the writing process of this article, I have adopted the opinions of Dr. Jiang X, the chief physician of the Department of Gastroenterology of Beijing Tsinghua Chang Gung Hospital, and the postdoctoral fellow of iCenter Liang X of Tsinghua University. I would like to express my sincere thanks!

REFERENCES

1 **Meskó B**, Drobni Z, Bényei É, Gergely B, Győrffy Z. Digital health is a cultural transformation of traditional healthcare. *Mhealth* 2017; **3**: 38 [PMID: 29184890 DOI: 10.21037/mhealth.2017.08.07]
2 **Leachman SA**, Merlino G. Medicine: The final frontier in cancer diagnosis. *Nature* 2017; **542**: 36-38 [PMID: 28150762 DOI: 10.1038/nature21492]
3 **Furiasse N**, Thomas JD. Automated Algorithmic Software in Echocardiography: Artificial Intelligence? *J Am Coll Cardiol* 2015; **66**: 1467-1469 [PMID: 26403343 DOI: 10.1016/j.jacc.2015.08.009]
4 **Zhang R**, Zheng Y, Mak TW, Yu R, Wong SH, Lau JY, Poon CC. Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE J Biomed Health Inform* 2017; **21**: 41-47 [PMID: 28114040 DOI: 10.1109/JBHI.2016.2635662]
5 **Chen W**, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; **66**: 115-132 [PMID: 26808342 DOI: 10.3322/caac.21338]
6 **DeSantis CE**, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R, Robbins AS, Jemal A. Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin* 2014; **64**: 252-271 [PMID: 24890451 DOI: 10.3322/caac.21235]
7 **Zhang YZ**, Li YY. Inflammatory bowel disease: pathogenesis. *World J Gastroenterol* 2014; **20**: 91-99 [PMID: 24415861 DOI: 10.3748/wjg.v20.i1.91]
8 **Lejeune C**, Sassi F, Ellis L, Godward S, Mak V, Day M, Rachet B. Socio-economic disparities in access to treatment and their impact on colorectal cancer survival. *Int J Epidemiol* 2010; **39**: 710-717 [PMID: 20378687 DOI: 10.1093/ije/dyq048]
9 **Borowski DW**, Cawkwell S, Zaidi SM, Toward M, Maguire N, Gill TS. Primary care referral practice, variability and socio-economic deprivation in colorectal cancer. *Colorectal Dis* 2016; **18**: 1072-1079 [PMID: 27110954 DOI: 10.1111/codi.13360]
10 **Solmi F**, Von Wagner C, Kobayashi LC, Raine R, Wardle J, Morris S. Decomposing socio-economic inequality in colorectal cancer screening uptake in England. *Soc Sci Med* 2015; **134**: 76-86 [PMID: 25917138 DOI: 10.1016/j.socscimed.2015.04.010]
11 **Chan WK**, Saravanan A, Manikam J, Goh KL, Mahadeva S. Appointment waiting times and education level influence the quality of bowel preparation in adult patients undergoing colonoscopy. *BMC Gastroenterol* 2011; **11**: 86 [PMID: 21798022 DOI: 10.1186/1471-230X-11-86]
12 **Zhao SB**, Yang X, Fang J, Wang SL, Gu L, Xia T, Su XJ, Wang D, Li ZS, Bai Y. Effect of left lateral tilt-down position on cecal intubation time: a 2-center, pragmatic, randomized controlled trial. *Gastrointest Endosc* 2018; **87**: 852-861 [PMID: 29158180 DOI: 10.1016/j.gie.2017.11.012]
13 **Moon SY**, Kim BC, Sohn DK, Han KS, Kim B, Hong CW, Park BJ, Ryu KH, Nam JH. Predictors for difficult cecal insertion in colonoscopy: The impact of obesity indices. *World J Gastroenterol* 2017; **23**: 2346-2354 [PMID: 28428714 DOI: 10.3748/wjg.v23.i13.2346]
14 **Bruce M**, Choi J. Detection of endoscopic looping during colonoscopy procedure by using embedded bending sensors. *Med Devices (Auckl)* 2018; **11**: 171-191 [PMID: 29849469 DOI: 10.2147/MDER.S146934]
15 **Chen Y**, Duan YT, Xie Q, Qin XP, Chen B, Xia L, Zhou Y, Li NN, Wu XT. Magnetic endoscopic imaging *vs* standard colonoscopy: meta-analysis of randomized controlled trials. *World J Gastroenterol* 2013; **19**: 7197-7204 [PMID: 24222966 DOI: 10.3748/wjg.v19.i41.7197]
16 **Shah SG**, Saunders BP. Aids to insertion: magnetic imaging, variable stiffness, and overtubes. *Gastrointest Endosc Clin N Am* 2005; **15**: 673-686 [PMID: 16278132 DOI: 10.1016/j.giec.2005.08.011]

17 **Prieto-de-Frías C**, Muñoz-Navas M, Carretero C, Carrascosa J, Betés MT, de-la-Riva S, Herraiz MT, Súbtil JC. Comparative study of a responsive insertion technology (RIT) colonoscope *vs* a variable-stiffness colonoscope. *Rev Esp Enferm Dig* 2013; **105**: 208-213 [PMID: 23859449 DOI: 10.4321/s1130-01082013000400005]

18 **Pasternak A**, Szura M, Solecki R, Matyja M, Szczepanik A, Matyja A. Impact of responsive insertion technology (RIT) on reducing discomfort during colonoscopy: randomized clinical trial. *Surg Endosc* 2017; **31**: 2247-2254 [PMID: 27631316 DOI: 10.1007/s00464-016-5226-x]

19 **Lee D**, Joe S, Choi J, Lee BI, Kim B. An elastic caterpillar-based self-propelled robotic colonoscope with high safety and mobility. *Mechatronics* 2016; **39**: 54-62 [DOI: 10.1016/j.mechatronics.2016.08.002]

20 **Rösch T**, Adler A, Pohl H, Wettschureck E, Koch M, Wiedenmann B, Hoepffner N. A motor-driven single-use colonoscope controlled with a hand-held device: a feasibility study in volunteers. *Gastrointest Endosc* 2008; **67**: 1139-1146 [PMID: 18355823 DOI: 10.1016/j.gie.2007.10.065]

21 **Hu W**, Lum GZ, Mastrangeli M, Sitti M. Small-scale soft-bodied robot with multimodal locomotion. *Nature* 2018; **554**: 81-85 [PMID: 29364873 DOI: 10.1038/nature25443]

22 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: 29042219 DOI: 10.1053/j.gastro.2017.10.010]

23 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez ML, Kelly R, Iqbal N, Chandelier F, Rex DK. Su1614 Artificial Intelligence (AI) in Endoscopy--Deep Learning for Optical Biopsy of Colorectal Polyps in Real-Time on Unaltered Endoscopic Videos. *Gastrointest Endosc* 2017; **85**: AB364-AB365 [DOI: 10.1016/j.gie.2017.03.843]

24 **Komeda Y**, Handa H, Watanabe T, Nomura T, Kitahashi M, Sakurai T, Okamoto A, Minami T, Kono M, Arizumi T, Takenaka M, Hagiwara S, Matsui S, Nishida N, Kashida H, Kudo M. Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience. *Oncology* 2017; **93** Suppl 1: 30-34 [PMID: 29258081 DOI: 10.1159/000481227]

25 **Mori Y**, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018; **169**: 357-366 [PMID: 30105375 DOI: 10.7326/M18-0249]

26 **Hassan AR**, Haque MA. Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos. *Comput Methods Programs Biomed* 2015; **122**: 341-353 [PMID: 26390947 DOI: 10.1016/j.cmpb.2015.09.005]

27 **Xu W**, Yan G, Wang Z, Liu G, Kuang S, Zhao S. [A method for bleeding detection in endoscopy images using SVM]. *Zhongguo Yiliao Qixie Zazhi* 2015; **39**: 9-12 [PMID: 26027285]

28 **Yuan Y**, Meng MQ. Deep learning for polyp recognition in wireless capsule endoscopy images. *Med Phys* 2017; **44**: 1379-1389 [PMID: 28160514 DOI: 10.1002/mp.12147]

29 **Zhou ZH**, Feng J. Deep forest. *Nat Sci Rev* 2019; **6**: 74-86 [DOI: 10.1093/nsr/nwy108]

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 June 28; 2(3): 71-78

DOI: 10.37126/aige.v2.i3.71

ISSN 2689-7164 (online)

MINIREVIEWS

Application of convolutional neural network in detecting and classifying gastric cancer

Xin-Yi Feng, Xi Xu, Yun Zhang, Ye-Min Xu, Qiang She, Bin Deng

ORCID number: Xin-Yi Feng 0000-0001-5960-1028; Xi Xu 0000-0003-1810-7220; Yun Zhang 0000-0002-2801-7547; Ye-Min Xu 0000-0002-9058-1966; Qiang She 0000-0001-8861-2674; Bin Deng 0000-0002-5590-3755.

Author contributions: Feng XY and Xu X contributed equally to this work; Feng XY and Xu X conceived and drafted the manuscript; Feng XY, Xu X, Zhang Y, and Xu YM collected the relevant information; She Q and Deng B revised the manuscript.

Supported by The Key Project for Social Development of Yangzhou, No. YZ2020069.

Conflict-of-interest statement: The authors report no conflicts of interest in this work.

Xin-Yi Feng, Xi Xu, Yun Zhang, Ye-Min Xu, Qiang She, Bin Deng, Department of Gastroenterology, Affiliated Hospital of Yangzhou University, Yangzhou 225000, Jiangsu Province, China

Corresponding author: Bin Deng, MD, Associate Professor, Chief Physician, Department of Gastroenterology, Affiliated Hospital of Yangzhou University, No. 368 Hanjiang Middle Road, Yangzhou 225000, Jiangsu Province, China. chinadbin@126.com

Abstract

Gastric cancer (GC) is the fifth most common cancer in the world, and at present, esophagogastroduodenoscopy is recognized as an acceptable method for the screening and monitoring of GC. Convolutional neural networks (CNNs) are a type of deep learning model and have been widely used for image analysis. This paper reviews the application and prospects of CNNs in detecting and classifying GC, aiming to introduce a computer-aided diagnosis system and to provide evidence for subsequent studies.

Key Words: Artificial intelligence; Convolutional neural network; Endoscopy; Gastric cancer; Deep learning

Core Tip: With the development of new algorithms and big data, great achievements in artificial intelligence (AI) based on deep learning have been made in diagnostic imaging, especially convolutional neural network (CNN). Esophagogastroduodenoscopy (EGD) is currently the most common method for screening and diagnosing gastric cancer (GC). When AI was combined with EGD, the diagnostic efficacy of GC could be improved. Therefore, we review the application and prospect of CNN in detecting and classifying GC, aiming to introduce a computer-aided diagnosis system and provide evidence for following studies.

Citation: Feng XY, Xu X, Zhang Y, Xu YM, She Q, Deng B. Application of convolutional neural network in detecting and classifying gastric cancer. *Artif Intell Gastrointest Endosc* 2021; 2(3): 71-78

URL: https://www.wjgnet.com/2689-7164/full/v2/i3/71.htm

DOI: https://dx.doi.org/10.37126/aige.v2.i3.71

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: April 27, 2021
Peer-review started: April 27, 2021
First decision: April 28, 2021
Revised: May 21, 2021
Accepted: June 7, 2021
Article in press: June 7, 2021
Published online: June 28, 2021

P-Reviewer: Viswanath YK
S-Editor: Gao CC
L-Editor: Wang TQ
P-Editor: Wang LYT

INTRODUCTION

Gastric cancer (GC) is a globally prevalent cancer, and its incidence and mortality rank fifth and fourth, respectively, among cancers worldwide[1]. It is estimated that in 2020 there were over 1000000 new cases and 769000 deaths of GC globally. The lack of early detection and treatment contributes to the high mortality and poor outcomes of GC [2]. Esophagogastroduodenoscopy (EGD) is currently the most common method for screening and diagnosing GC. However, the efficacy of EGD varies significantly[3]. It has been reported that the false negative rate of EGD in detecting GC ranges from 4.6%-25.8%[4-6]. GC lesions are difficult to recognize due to the subtle changes in the gastric mucosa[7]. Additionally, the quality of EGD can be heavily influenced by the subjective determination of endoscopists[8]. Therefore, it is significant to develop an objective and reliable method to recognize possible early GC (EGC) lesions and blind spots.

With the development of new algorithms and big data, great achievements in artificial intelligence (AI) based on deep learning (DL) have been made for diagnostic imaging. Meanwhile, as one of the most representative network models in DL, convolutional neural network (CNN) contributes to enhancing the accuracy of image analysis. CCN is now being successfully applied in detecting the gastrointestinal tract [9-11]. CNNs have achieved tremendous successes and wide application in image recognition and classification[12,13]. Therefore, we applied CNN in endoscopic diagnosis, aiming to improve the diagnostic efficacy of EGC. In this review, we scrupulously elucidate the application and evolution of CNN in the detection and classification of GC.

CONVOLUTIONAL NEURAL NETWORK

With the development of neuroscience, researchers have attempted to build artificial neural networks to simulate the structure of the human brain by mathematically activating neuronal activity. DL has been the mainstream machine learning method in many applications. It is a type of representation learning method in which a complex neural network architecture automatically learns representative data by transforming the input information into multiple levels of abstractions[10]. Computer-aided diagnosis requires the extraction of extensive original image data and the application of a series of complex algorithms. DL has a strong modeling and reasoning ability that is superb in realizing computer output diagnosis.

CNNs are neural networks sharing connections between hidden units that feature a shortened computational time and translational invariance properties[14]. A typical CNN framework includes three main components: A convolutional layer, an activation function, and a pooling layer. The convolutional layer is composed of several small matrices. These matrices are convolved throughout the whole input image working as filters, and then a nonlinear transformation is applied in an element-wise fashion. Finally, the pooling layer aggregates contiguous values to one scalar. The common types of pooling in popular use are either average or max[15,16].

In the early 1990s, CNNs were used in many applications, such as object detection and face recognition. With the advances of technology, CNN was first applied to the analysis of medical images in 1993. Lo *et al*[17] reported the detection of lung nodules using a CNN in 1995. However, due to the limitation of computer language, CNNs have been underestimated in their value for a long time. In 2012, Krizhevsky *et al*[18] proposed a CNN with five convolutional layers and three fully connected layers (namely, AlexNet) and achieved breakthrough performances in the ImageNet Large Scale Visual Recognition Challenge. Since then, CNNs have been of great interest and widely applied. For example, CNNs have been applied to identify diabetic retinopathy from fundus photographs and distinguish benign proliferative breast lesions from malignant[19]. In 2020, Plaksin *et al*[20] estimated the possibility of diagnosing malignant pleural effusion from facies images of pleural exudates obtained by the method of wedge-shaped dehydration using CNNs.

Compared with the general neural network, CNN is superior in the adaptation of the image structure, extraction, and classification, and as a result it presents satisfactory work efficiency.

APPLICATION OF CNN IN GC

Automatic detection

At present, CNNs have been applied to detect GC, showing distinctive improvements. Hirasawa *et al*[10] created and trained a CNN-based diagnostic system containing 13584 endoscopic images. In this study, the constructed CNN was able to detect 92.2% of GC cases, including small intramucosal GC, through a quick analysis of an independent test set involving 2296 stomach images, which is extremely difficult even by experienced endoscopists. To achieve the real-time detection of EGD, Ishioka *et al* [21] tested their CNN system for identifying video images and achieved a high detection rate (94.1%). The detection rate in video images by CNN is similar to that of still images, demonstrating the great potential of CNN in the early detection of GC.

Magnifying endoscopy with narrow band imaging (M-NBI) has been used for the differential diagnosis of various focal, superficial gastric lesions. By observing the microvasculature and fine mucosal structure, M-NBI has a better accuracy in the diagnosis of early GC than ordinary white light endoscopy[22]. Li *et al*[23] developed a novel CNN-based system for analyzing gastric mucosal lesions observed by M-NBI. The test results showed that the sensitivity, specificity, and accuracy of the CNN system in diagnosing early GC were 91.18%, 90.64%, and 90.91%, respectively. Notably, the specificity and accuracy of CNN diagnostics are comparable to those of experts with more than 10 years of clinical experience.

Ikenoyama *et al*[24] compared the diagnostic ability of CNN and 67 endoscopists, and the results showed that CNN had a faster processing speed and 25% higher sensitivity than endoscopists [95% confidence interval (CI): 14.9-32.5]. The use of CNN can effectively urge endoscopists to re-examine and evaluate ambiguous lesions, which also helps reduce false negatives and false positives (Table 1).

Histological classification

An excellent endoscopist not only detects mucosal lesions but also distinguishes benign and malignant features. Cho *et al*[25] trained three CNN models, namely, Inception-v4, Resnet-152, and Inception-Resnet-v2, to classify gastric lesions into five categories: Advanced GC, EGC, high-grade dysplasia, low-grade dysplasia, and non-neoplasm. Among these systems, the Inception-Resnet-v2 model showed the best performance; the weighted average accuracy reached 84.6%, and the mean area under the curve (AUC) of the model for differentiating GC and neoplasm was 0.877 and 0.927, respectively.

To date, pathological diagnosis is still the gold standard to assess the presence or absence of cancerous lesions, cancer types, and degree of malignancy. Nevertheless, the accuracy of diagnosis and workload alleviation of pathologists are still challenging, and advanced computer-aided technologies are expected to play a key role in assisting pathological diagnosis. By optically scanning histologic tissue slides and converting them into ultrahigh-resolution digital images called whole slide images (WSIs), digital pathology is available for further investigations[26]. With the rapid development of EGD, the combination of DL models such as CNN and digital pathology is expected to greatly reduce the increasing workload of pathologists.

Sharma *et al*[27] explored two computerized applications of CNNs in GC, cancer classification and necrosis detection, based on immunohistochemistry of human epidermal growth factor receptor 2 and hematoxylin-eosin staining of histopathological WSIs. The overall classification accuracies that they obtained were 0.6990 and 0.8144, respectively. However, their study is limited by a small sample size with only 11 WSIs involved.

Iizuka *et al*[28] collected a large dataset of 4128 WSIs of stomach samples to train CNN and a recurrent neural network, and the evaluation results of CNN showed that the AUC for detecting gastric adenocarcinoma and adenoma was up to 0.97 and 0.99, respectively. They proposed that DL models can be used as a component in an integrated workflow alongside slide scanning, thus determining the top priority of the most valuable case, enhancing the accuracy of diagnosis, and speeding up the work efficacy.

Song *et al*[29] established a multicenter massive WSI dataset and tested slides collected from different hospitals that were detected with the histopathological diagnosis system for GC detection using DL. The results showed that the AUCs of the AI assistance system developed at the Chinese PLA General Hospital, Peking Union Medical College Hospital, and Cancer Hospital, Chinese Academy of Medical Sciences, were 0.986, 0.990, and 0.996, respectively, confirming its consistent stable performance. Their model-building approach may also be applied to identify multiple

Table 1 Detailed information on studies concerning automatic detection by convolutional neural network in gastric cancer

Ref.	Endoscopic images	Training dataset	Test dataset	Resolution	Sensitivity %	Specificity %	Accuracy/AUC %	PPV %	NPV %
Hirasawa *et al* [10] (2018)	WLI/NBI/chromoendoscopy images	13584	2296	300 × 300	92.2	NA	NA	30.6	NA
Ishioka *et al* [21] (2019)	Video images	NA	68	NA	94.1	NA	NA	NA	NA
Li *et al*[23] (2020)	M-NBI images	20000	341	512 × 512	91.18	90.64	90.91	90.64	91.18
Ikenoyama *et al*[24](2021)	WLI/NBI/chromoendoscopy images	13584	2940	300 × 300	58.4	87.3	75.7	26.0	96.5

AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value; WLI: White-light imaging; NBI: Narrow-band imaging; M-NBI: Magnifying narrow-band imaging; NA: Not applicable.

cancers in different organ systems in the future (Table 2).

Prediction of depth of tumor invasion

EGC is categorized as a lesion confined to the mucosa (T1A) or the submucosa (T1B). An accurate identification of the depth of tumor invasion is the basis for determining the therapeutic schedule[30]. Endoscopic mucosal changes, such as irregular surfaces and submucosal tumors (*e.g.*, marginal elevation), have been suggested as predictors of the depth of tumor invasion[31].

Zhu *et al*[11] built a CNN computer-aided detection (CNN-CAD) system to determine the depth of tumor invasion, which is expected to avoid unnecessary gastrectomy. In this system, there was a development dataset of 790 images and a test dataset of 203 images. The final results showed that the AUC for the CNN-CAD system was 0.94 (95%CI: 0.90-0.97), and the overall accuracy was 89.16%, which was significantly higher than that determined by endoscopists (17.25%, 95%CI: 11.63-22.59). Yoon *et al*[32] proposed a novel loss function for developing an optimized EGC depth prediction model, called the lesion-based visual geometry group-16. Using this novel function, the depth prediction model is able to accurately activate the EGC regions during training and simultaneously measure classification and localization errors. After experimenting with a total of 11539 endoscopic images, including 896 images of T1A-EGC, 809 of T1B-EGC, and 9834 of non-EGC, the AUC of the EGC depth prediction model was 0.851. In this study, it was also demonstrated that histopathological differentiation significantly affects the diagnostic accuracy of AI for determining T staging.

Upper abdominal enhanced computed tomography (CT) is the main imaging examination for T staging of GC[33]. Zheng *et al*[34] retrospectively collected 3500 venous phase-enhanced CT images of the upper abdomen from 225 patients with advanced GC, aiming to predict the depth of GC invasion and extract different regions of interest. The dataset was then enhanced by cropping and flipping, and the Faster R-CNN detection model was trained using other data enhancement methods. They found that the AUC of the experimentally established CNN model was 0.93, and the recognition accuracies for T2, T3, and T4 GC were 90%, 93%, and 95%, respectively. The abovementioned findings may be helpful for radiologists to predict the progression and postoperative outcomes of advanced GC (Table 3).

CURRENT EXISTING PROBLEMS

Limitations of studies

Selection bias: In most studies, researchers tend to select clear, typical, high-quality endoscopic images for training and testing image sets[10,35]. Because low-quality images with air, postbiopsy bleeding, halation, blurs, defocusing, or mucus secretion have been excluded, the results of retrospective clinical tests are often superior to actual ones. Therefore, prospective studies that are less affected by biases should be thoroughly analyzed to improve the accuracy and specificity of clinical trials, thus ensuring the reliability of the results.

Table 2 Detailed information on studies concerning histological classification by convolutional neural network in gastric cancer

Ref.	Training dataset	Test dataset	Resolution	Group	AUC %
Cho *et al*[25] (2019)	4205	812	1280 × 640	Five-category classification	84.6
				Cancer *vs* non-cancer	87.7
				Neoplasm *vs* non-neoplasm	92.7
Sharma *et al*[27] (2017)	231000 for cancer classification	NA	512 × 512	Cancer classification	69.9
	47130 for necrosis detection			Necrosis detection	81.4
Iizuka *et al*[28] (2020)	3628	500	512 × 512	Adenocarcinoma	98
				Adenoma	93.6
Song *et al*[29] (2020)	2123	3212 from PLAGH	320 × 320	Benign and malignant cases and tumour subtypes	98.6
		595 from PUMCH			99.0
		987 from CHCAMS			99.6

PLAGH: Chinese PLA General Hospital; PUMCH: Peking Union Medical College Hospital; CHCAMS: Cancer Hospital, Chinese Academy of Medical Sciences; AUC: Area under the curve.

Table 3 Detailed information on studies concerning prediction of depth of tumor invasion by convolutional neural network in gastric cancer

Ref.	Dataset	Resolution	Sensitivity %	Specificity %	Accuracy/AUC %	PPV %	NPV %
Zhu *et al* [11] (2019)	Development datasets: 5056; Validation datasets: 1264; Test dataset: 203	299 × 299	76.47	95.56	89.16	89.66	88.97
Yoon *et al* [32] (2019)	11539 images were randomly organized into five different folds, and at each fold, the training: validation: testing dataset ratio was 3:1:1	NA	79.2	77.8	85.1	79.3	77.7
Zheng *et al* [34] (2020)	Totally 5855, training:verification dataset ratio was 4:1	512 × 557	NA	NA	T2 stage: 90; T3 stage: 93; T4 stage: 95	NA	NA

AUC: Area under the curve; PPV: Positive predictive value; NPV: Negative predictive value; NA: Not applicable.

Single-center studies: Most of the testing images are obtained from a single-center institution using the same type of endoscope and endoscopic video system, which may result in potential biases. In future studies, images obtained from multicenter institutions using different types of endoscopic devices should be collected for analysis.

Lack of endoscopic video images: Still images are used for the training and test dataset in most studies, which may limit the extensive clinical application[36]. Using video images may improve the performance of the CNN and represent real-life scenarios[21].

Limitations of CNN

False positive and false negative results: The specificity and sensitivity of automatic detection are very important to determine the choice of therapeutic schedule. False positive and false negative results directly lead to improper treatment. For example, gastritis with pathological manifestations of redness, atrophy, and intestinal metaplasia is easily confused with EGC, which increases the false positive rate[10]. In addition, early-stage cancer lesions are often too small to be found, which increases the false negative rate. The main reason for false positive and false negative results may be attributed to the limited quantity and quality of learning samples. Therefore, it is necessary to collect a large number of high-quality endoscopic images for training algorithms, thus enhancing the detection accuracy.

Ethical and moral issues: AI will not completely replace doctors. Who should be responsible for the safety of patients if misdiagnosed? Patient consent should be obtained before using AI to determine who should be responsible for misdiagnosis or incorrect treatment that can possibly occur[37].

CONCLUSION

As a classical and widely used DL model, CNN has been widely used in the medical field, especially for EGD detection. In remote or crowded areas, CNNs can be used to assist early cancer screening to prevent misdiagnosis due to a lack of experience and professional knowledge of endoscopists. Additionally, CNN is a promising method to provide online professional training for improving the professional skills of young endoscopists. Most importantly, CNN helps endoscopists detect, classify, and even predict the invasion depth of EGC.

At present, most of studies are still in the early stages of system development. More powerful, efficient, and stable algorithms, and more prospective studies are urgently required in the future to make AI more sensitive, specific, and accurate in cancer detection and classification.

REFERENCES

1 **Sung H**, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; **71**: 209-249 [PMID: 33538338 DOI: 10.3322/caac.21660]

2 **Zong L**, Abe M, Seto Y, Ji J. The challenge of screening for early gastric cancer in China. *Lancet* 2016; **388**: 2606 [PMID: 27894662 DOI: 10.1016/S0140-6736(16)32226-7]

3 **Rutter MD**, Senore C, Bisschops R, Domagk D, Valori R, Kaminski MF, Spada C, Bretthauer M, Bennett C, Bellisario C, Minozzi S, Hassan C, Rees C, Dinis-Ribeiro M, Hucl T, Ponchon T, Aabakken L, Fockens P. The European Society of Gastrointestinal Endoscopy Quality Improvement Initiative: developing performance measures. *United European Gastroenterol J* 2016; **4**: 30-41 [PMID: 26966520 DOI: 10.1177/2050640615624631]

4 **Hosokawa O**, Hattori M, Douden K, Hayashi H, Ohta K, Kaizaki Y. Difference in accuracy between gastroscopy and colonoscopy for detection of cancer. *Hepatogastroenterology* 2007; **54**: 442-444 [PMID: 17523293]

5 **Raftopoulos SC**, Segarajasingam DS, Burke V, Ee HC, Yusoff IF. A cohort study of missed and new cancers after esophagogastroduodenoscopy. *Am J Gastroenterol* 2010; **105**: 1292-1297 [PMID: 20068557 DOI: 10.1038/ajg.2009.736]

6 **Vradelis S**, Maynard N, Warren BF, Keshav S, Travis SP. Quality control in upper gastrointestinal endoscopy: detection rates of gastric cancer in Oxford 2005-2008. *Postgrad Med J* 2011; **87**: 335-339 [PMID: 21257996 DOI: 10.1136/pgmj.2010.101832]

7 **Pasechnikov V**, Chukov S, Fedorov E, Kikuste I, Leja M. Gastric cancer: prevention, screening and early diagnosis. *World J Gastroenterol* 2014; **20**: 13842-13862 [PMID: 25320521 DOI: 10.3748/wjg.v20.i38.13842]

8 **Scaffidi MA**, Grover SC, Carnahan H, Khan R, Amadio JM, Yu JJ, Dargavel C, Khanna N, Ling SC, Yong E, Nguyen GC, Walsh CM. Impact of experience on self-assessment accuracy of clinical colonoscopy competence. *Gastrointest Endosc* 2018; **87**: 827-836. e2 [PMID: 29122599 DOI: 10.1016/j.gie.2017.10.040]

9 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

10 **Hirasawa T**, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J, Tada T. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018; **21**: 653-660 [PMID: 29335825 DOI: 10.1007/s10120-018-0793-2]

11 **Zhu Y**, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, Zhang YQ, Chen WF, Yao LQ, Zhou PH, Li QL. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointest Endosc* 2019; **89**: 806-815. e1 [PMID: 30452913 DOI: 10.1016/j.gie.2018.11.011]

12 **Brinker TJ**, Hekler A, Enk AH, von Kalle C. Enhanced classifier training to improve precision of a convolutional neural network to identify images of skin lesions. *PLoS One* 2019; **14**: e0218713 [PMID: 31233565 DOI: 10.1371/journal.pone.0218713]

13 **Huang Y**, Xu J, Zhou Y, Tong T, Zhuang X; Alzheimer's Disease Neuroimaging Initiative (ADNI). Diagnosis of Alzheimer's Disease *via* Multi-Modality 3D Convolutional Neural Network. *Front Neurosci* 2019; **13**: 509 [PMID: 31213967 DOI: 10.3389/fnins.2019.00509]

14 **Ke Q**, Li Y. Is Rotation a Nuisance in Shape Recognition? *IEEE Confer Comp Vis Patt Rec* 2014;

4146-4153 [DOI: 10.1109/CVPR.2014.528]
15 **Arevalo J**, González FA, Ramos-Pollán R, Oliveira JL, Guevara Lopez MA. Representation learning for mammography mass lesion classification with convolutional neural networks. *Comput Methods Programs Biomed* 2016; **127**: 248-257 [PMID: 26826901 DOI: 10.1016/j.cmpb.2015.12.014]
16 **Song Q**, Zhao L, Luo X, Dou X. Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. *J Healthc Eng* 2017; **2017**: 8314740 [PMID: 29065651 DOI: 10.1155/2017/8314740]
17 **Lo SB**, Lou SA, Lin JS, Freedman MT, Chien MV, Mun SK. Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Trans Med Imaging* 1995; **14**: 711-718 [PMID: 18215875 DOI: 10.1109/42.476112]
18 **Krizhevsky A**, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2012; **60**: 84-90
19 **Shaban M**, Ogur Z, Mahmoud A, Switala A, Shalaby A, Abu Khalifeh H, Ghazal M, Fraiwan L, Giridharan G, Sandhu H, El-Baz AS. A convolutional neural network for the screening and staging of diabetic retinopathy. *PLoS One* 2020; **15**: e0233514 [PMID: 32569310 DOI: 10.1371/journal.pone.0233514]
20 **Plaksin SA**, Farshatova LI, Veselov IV, Zamyatina EB. [Diagnosis of malignant pleural effusions using convolutional neural networks by the morphometric image analysis of facies of pleural exudate]. *Khirurgiia (Mosk)* 2020; 42-48 [PMID: 32500688 DOI: 10.17116/hirurgia202005142]
21 **Ishioka M**, Hirasawa T, Tada T. Detecting gastric cancer from video images using convolutional neural networks. *Dig Endosc* 2019; **31**: e34-e35 [PMID: 30449050 DOI: 10.1111/den.13306]
22 **Kaise M**, Kato M, Urashima M, Arai Y, Kaneyama H, Kanzazawa Y, Yonezawa J, Yoshida Y, Yoshimura N, Yamasaki T, Goda K, Imazu H, Arakawa H, Mochizuki K, Tajiri H. Magnifying endoscopy combined with narrow-band imaging for differential diagnosis of superficial depressed gastric lesions. *Endoscopy* 2009; **41**: 310-315 [PMID: 19340733 DOI: 10.1055/s-0028-1119639]
23 **Li L**, Chen Y, Shen Z, Zhang X, Sang J, Ding Y, Yang X, Li J, Chen M, Jin C, Chen C, Yu C. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. *Gastric Cancer* 2020; **23**: 126-132 [PMID: 31332619 DOI: 10.1007/s10120-019-00992-2]
24 **Ikenoyama Y**, Hirasawa T, Ishioka M, Namikawa K, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Takeuchi Y, Shichijo S, Katayama N, Fujisaki J, Tada T. Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. *Dig Endosc* 2021; **33**: 141-150 [PMID: 32282110 DOI: 10.1111/den.13688]
25 **Cho BJ**, Bang CS, Park SW, Yang YJ, Seo SI, Lim H, Shin WG, Hong JT, Yoo YT, Hong SH, Choi JH, Lee JJ, Baik GH. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019; **51**: 1121-1129 [PMID: 31443108 DOI: 10.1055/a-0981-6133]
26 **Jansen I**, Lucas M, Savci-Heijink CD, Meijer SL, Marquering HA, de Bruin DM, Zondervan PJ. Histopathology: ditch the slides, because digital and 3D are on show. *World J Urol* 2018; **36**: 549-555 [PMID: 29396786 DOI: 10.1007/s00345-018-2202-1]
27 **Sharma H**, Zerbe N, Klempert I, Hellwich O, Hufnagl P. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Comput Med Imaging Graph* 2017; **61**: 2-13 [PMID: 28676295 DOI: 10.1016/j.compmedimag.2017.06.001]
28 **Iizuka O**, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep Learning Models for Histopathological Classification of Gastric and Colonic Epithelial Tumours. *Sci Rep* 2020; **10**: 1504 [PMID: 32001752 DOI: 10.1038/s41598-020-58467-9]
29 **Song Z**, Zou S, Zhou W, Huang Y, Shao L, Yuan J, Gou X, Jin W, Wang Z, Chen X, Ding X, Liu J, Yu C, Ku C, Liu C, Sun Z, Xu G, Wang Y, Zhang X, Wang D, Wang S, Xu W, Davis RC, Shi H. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 2020; **11**: 4294 [PMID: 32855423 DOI: 10.1038/s41467-020-18147-8]
30 **Wang J**, Yu JC, Kang WM, Ma ZQ. Treatment strategy for early gastric cancer. *Surg Oncol* 2012; **21**: 119-123 [PMID: 21256735 DOI: 10.1016/j.suronc.2010.12.004]
31 **Tsujii Y**, Kato M, Inoue T, Yoshii S, Nagai K, Fujinaga T, Maekawa A, Hayashi Y, Akasaka T, Shinzaki S, Watabe K, Nishida T, Iijima H, Tsujii M, Takehara T. Integrated diagnostic strategy for the invasion depth of early gastric cancer by conventional endoscopy and EUS. *Gastrointest Endosc* 2015; **82**: 452-459 [PMID: 25841580 DOI: 10.1016/j.gie.2015.01.022]
32 **Yoon HJ**, Kim S, Kim JH, Keum JS, Oh SI, Jo J, Chun J, Youn YH, Park H, Kwon IG, Choi SH, Noh SH. A Lesion-Based Convolutional Neural Network Improves Endoscopic Detection and Depth Prediction of Early Gastric Cancer. *J Clin Med* 2019; **8** [PMID: 31454949 DOI: 10.3390/jcm8091310]
33 **Ajani JA**, D'Amico TA, Almhanna K, Bentrem DJ, Chao J, Das P, Denlinger CS, Fanta P, Farjah F, Fuchs CS, Gerdes H, Gibson M, Glasgow RE, Hayman JA, Hochwald S, Hofstetter WL, Ilson DH, Jaroszewski D, Johung KL, Keswani RN, Kleinberg LR, Korn WM, Leong S, Linn C, Lockhart AC, Ly QP, Mulcahy MF, Orringer MB, Perry KA, Poultsides GA, Scott WJ, Strong VE, Washington MK, Weksler B, Willett CG, Wright CD, Zelman D, McMillian N, Sundar H. Gastric Cancer, Version 3.2016, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2016; **14**: 1286-1312 [PMID: 27697982 DOI: 10.6004/jnccn.2016.0137]
34 **Zheng L**, Zhang X, Hu J, Gao Y, Zhang M, Li S, Zhou X, Niu T, Lu Y, Wang D. Establishment and

Applicability of a Diagnostic System for Advanced Gastric Cancer T Staging Based on a Faster Region-Based Convolutional Neural Network. *Front Oncol* 2020; **10**: 1238 [PMID: 32850373 DOI: 10.3389/fonc.2020.01238]

35 **Gotoda T**, Uedo N, Yoshinaga S, Tanuma T, Morita Y, Doyama H, Aso A, Hirasawa T, Yano T, Uchita K, Ho SH, Hsieh PH. Basic principles and practice of gastric cancer screening using high-definition white-light gastroscopy: Eyes can only see what the brain knows. *Dig Endosc* 2016; **28** Suppl 1: 2-15 [PMID: 26836611 DOI: 10.1111/den.12623]

36 **England JR**, Cheng PM. Artificial Intelligence for Medical Image Analysis: A Guide for Authors and Reviewers. *AJR Am J Roentgenol* 2019; **212**: 513-519 [PMID: 30557049 DOI: 10.2214/AJR.18.20490]

37 **Jin P**, Ji X, Kang W, Li Y, Liu H, Ma F, Ma S, Hu H, Li W, Tian Y. Artificial intelligence in gastric cancer: a systematic review. *J Cancer Res Clin Oncol* 2020; **146**: 2339-2350 [PMID: 32613386 DOI: 10.1007/s00432-020-03304-9]

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 June 28; 2(3): 79-88

DOI: 10.37126/aige.v2.i3.79

ISSN 2689-7164 (online)

MINIREVIEWS

Utility of artificial intelligence in colonoscopy

Niel Shah, Abhilasha Jyala, Harish Patel, Jasbir Makker

ORCID number: Niel Shah 0000-0001-7459-0099; Abhilasha Jyala 0000-0002-3685-1697; Harish Patel 0000-0003-3638-9495; Jasbir Makker 0000-0002-1673-8515.

Author contributions: Patel H conceived with an idea about this work and approved the final manuscript; Shah N, Jyala M and Makker J were responsible for reviewing of the literatures and studies, drafting the outline and main manuscript; Patel H and Makker J were responsible for critically reviewing the manuscript for intellectual content, reviewing of the literatures and editing of the main manuscript.

Conflict-of-interest statement: Authors have nothing to disclose.

Manuscript source: Invited

Niel Shah, Abhilasha Jyala, Department of Internal Medicine, BronxCare Hospital Center, Bronx, NY 10457, United States

Harish Patel, Jasbir Makker, Department of Internal Medicine, Gastroenterology, BronxCare Hospital Center, Bronx, NY 10457, United States

Corresponding author: Harish Patel, MD, Attending Doctor, Department of Internal Medicine, Gastroenterology, BronxCare Hospital Center, 1650 Grand Concourse , Bronx, NY 10457, United States. hpatel@bronxcare.org

Abstract

Colorectal cancer is one of the major causes of death worldwide. Colonoscopy is the most important tool that can identify neoplastic lesion in early stages and resect it in a timely manner which helps in reducing mortality related to colorectal cancer. However, the quality of colonoscopy findings depends on the expertise of the endoscopist and thus the rate of missed adenoma or polyp cannot be controlled. It is desirable to standardize the quality of colonoscopy by reducing the number of missed adenoma/polyps. Introduction of artificial intelligence (AI) in the field of medicine has become popular among physicians nowadays. The application of AI in colonoscopy can help in reducing miss rate and increasing colorectal cancer detection rate as per recent studies. Moreover, AI assistance during colonoscopy has also been utilized in patients with inflammatory bowel disease to improve diagnostic accuracy, assessing disease severity and predicting clinical outcomes. We conducted a literature review on the available evidence on use of AI in colonoscopy. In this review article, we discuss about the principles, application, limitations, and future aspects of AI in colonoscopy.

Key Words: Artificial intelligence; Colonoscopy; Colorectal cancer; Inflammatory bowel disease; Adenoma detection rate; Adenoma

Core Tip: Artificial intelligence (AI) pertains to performance of intelligent tasks like human beings by computer-controlled machines. Machine learning, one of the most important and fundamental principles of AI, essentially means automatically using the available data to learn and make decisions without human intervention. AI based detection models have been developed for polyp detection and to differentiate

manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: United States

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: June 2, 2021
Peer-review started: June 2, 2021
First decision: June 18, 2021
Revised: June 20, 2021
Accepted: June 28, 2021
Article in press: June 28, 2021
Published online: June 28, 2021

P-Reviewer: Guerra I
S-Editor: Liu M
L-Editor: A
P-Editor: Li JH

malignant from nonmalignant lesions. It has been also utilized to analyze endoscopic images for inflammatory bowel disease diagnosis, grading its severity and predicting treatment response.

Citation: Shah N, Jyala A, Patel H, Makker J. Utility of artificial intelligence in colonoscopy. *Artif Intell Gastrointest Endosc* 2021; 2(3): 79-88
URL: https://www.wjgnet.com/2689-7164/full/v2/i3/79.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i3.79

INTRODUCTION

What is artificial intelligence

The capability of human brain to perceive, analyze and react is defined as intelligence. Gottfredson[1] described it as ability of a human beings to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It has been a long desire of human beings to build machines which can think and act autonomously to ease human work. Several complex computer algorithms and models have been developed to provide automation to these machines. The famous Turing test invented by Alan Turing in 1950 demonstrated that it may be difficult for a blinded investigator to distinguish humans from intelligent machines[2]. However, intelligence of these machines is still way below human intelligence which is based on logic, reasoning, and adaptive learning. In 1997 International Business Machines's artificial intelligence (AI) driven chess playing system defeated world chess champion Garry Kasparov. Although this victory of computer programs over human beings in chess was criticized by many, and it was argued that machines can only be as good as the programs developed for them by human beings, nevertheless it remains an important landmark in the history of AI.

There is no one formal definition of AI. It is vaguely defined as ability of computer-controlled machines to perform intelligent tasks like human beings. There are two basic subtypes of AI- weak or soft and strong or hard AI[3]. Weak or soft AI is also called as narrow AI and as the name suggests it specializes in a very specific task like face recognition, voice recognition capabilities. On the other hand, strong or hard AI which is also known as general AI has more broad application due to its capability to understand, think and act like human beings. It is at the core of advanced robotic systems.

Machine learning (ML) is one of the most important and fundamental principles of AI. ML is at the heart of any AI system and essentially means automatically using the available data to learn and make decisions without human intervention. It is an adaptive technology which is continuously learning and hence gets better with each use. ML utilizes three fundamental methods which include supervised learning, unsupervised learning, and reinforcement learning. Artificial neural network (ANN) is a ML algorithm adapted from model of biological neurons in humans. ANN is an information processing technology, also considered as mathematical models utilized to analyze data.

AI IN THE FIELD OF GASTROENTEROLOGY

In the last two decades, substantial progress has been made in the use of AI driven algorithms in the field of medical science. Use of AI in the field of medical practice can be categorized in two broad categories-virtual and physical[4]. The virtual category of AI pertains to its use in electronic health record. It is based on ML and deep learning *via* mathematical algorithms to identify individuals at risk of some specific disease and help in clinical decision making. The physical category of AI includes use of medical devices and robotics for delivering medical care.

AI operated systems have been utilized to monitor patient's medical conditions remotely. More specifically in gastroenterology, AI based detection models have been developed to differentiate malignant from nonmalignant lesions, detect gastrointestinal bleeding using wireless video capsule endoscopy, detecting pancreatic cancer,

and detecting liver fibrosis. In the subsequent sections, we have detailed progress of AI and its application during colonoscopy.

AI AND COLON POLYPS

Colorectal cancer (CRC) is the third most common form of the cancer worldwide and is the 2nd most common cause of cancer related mortality globally[5]. Colonoscopy is the primary method for detection and removal of polyps and thus for prevention of CRC. It has been shown in the study that with every 1% increase in the adenoma detection rate (ADR), the risk of CRC decreases by 3%[6]. However, colonoscopy is not the perfect tool as polyps can be missed during colonoscopy mainly because of two factors: Blind spot and proceduralist error. The error due to blind spot can be overcome by using wide-angle camera, but the error due to proceduralist cannot be overcome easily. Small polyps (1-5 mm) are prone to be missed regardless of experience of proceduralist. Some studies have shown improvement in the rate of polyp detection with the help of second observer[7,8]. The factors responsible for proceduralist error could be fatigue, distraction, visual perception, impaired level of alertness, recognition error and poor bowel preparation. The application of AI in endoscopic field has shown improvement in ADR in recent studies and it helps in overcoming proceduralist error. Computer-aided detection and characterization of colorectal polyps is now getting popular among endoscopists.

PRINCIPLES AND APPLICATION OF AI IN COLONOSCOPY FOR POLYP DETECTION

AI has been a part of medical field since early 1950s. The concept and use of basic technology of computer-aided diagnosis (CAD) for colonoscopy has been explored since past one decade[9]. Use of CAD system in detection of colon polyps was first demonstrated by Karkanis *et al*[10]. Although the sensitivity of detecting adenomatous polyps demonstrated by these authors was 90%, this system was not used in clinical practice as it relied on static images rather than live endoscopic videos. In 2011, Bernal *et al*[11] introduced how intelligent systems can help in colonoscopy. Bernal *et al*[12] later introduced window median depth of valley accumulation (WM- DOVA) energy maps as a tool for automatic polyp detection in colonoscopy images. Fernández-Esparrach *et al*[13] for the first-time reported use of CAD system based on WM-DOVA maps and utilized colonoscopy videos in assisting colon polyp detection. With significant advancements in computer power and emergence of deep learning algorithms over past decade, it is being realized that CAD assistance during colonoscopy can be used in real time[14]. The inclusion of CAD for colonoscopy can help by automatic detection of polyps in real time which could be easily overlooked by endoscopists visually, thus resulting in higher ADR. Additionally, it helps in characterization of polyps in real time that in turn would help in reducing unnecessary biopsies of non-neoplastic polyps significantly[15].

There have been multiple studies to prove the advantage of inclusion of AI in the field of colonoscopy (Table 1). Most of these studies are of retrospective design, however few of them done recently were conducted prospectively. Luo *et al*[16] conducted a prospective, randomized cohort study using 150 participants to explore whether a high-performance, real-time automatic polyp detection system could improve the polyp detection rate in the actual clinical environment. The results showed that a real-time automatic polyp detection system can increase the ADR, especially for small polyps which are usually easily missed by conventional colonoscopy technique. Furthermore, Misawa *et al*[17] developed a 3-D convolutional network model for automated polyp detection which worked nearly in real time. They demonstrated sensitivity of 90% and a specificity 63% using 50 polyp videos and 85 non-polyp videos as test sets. Subsequently, Urban *et al*[18] developed a CAD model to improve polyp detection rate and they tested the model for its diagnostic capability on 8641 hand-labeled colonoscopy images collected from more than 2000 patients and on 20 colonoscopy videos. The results showed diagnostic accuracy of 96.4% and an area under the receiver operating characteristic curve of 0.991. However, the false positive rate was 7%. Additionally, Wang *et al*[19] developed the deep-learning algorithm which provided > 90% sensitivity and specificity for video-based analysis after testing their model on many polyp images and colonoscopy video recordings from patients.

Table 1 List of studies evaluating role of artificial intelligence in the detection of colon polyps during the colonoscopy

Ref.	Country of origin	Study design	Results
Fernandez-Esparrach *et al* [13], 2016	Spain	Retrospective	Sensitivity 70%, Specificity 72 %
Geetha *et al*[36], 2016	India	Retrospective	Sensitivity 95%, Specificity 97%
Misawa *et al*[37], 2017	Japan	Retrospective	Accuracy higher than trainees (87.8 *vs* 63.4%; $P = 0.01$), but similar to experts (87.8 *vs* 84.2%; $P = 0.76$)
Zhang *et al*[38], 2017	China	Retrospective	Accuracy 86%
Yu *et al*[39], 2017	China	Retrospective	Sensitivity 71%, PPV 88%
Billah *et al*[40], 2017	Bangladesh	Retrospective	Sensitivity 99%, Specificity 98.5%, Accuracy 99%
Chen *et al*[23], 2018	Taiwan	Retrospective	Sensitivity 96.3%, Specificity 78.1%
Urban *et al*[18], 2018	United States	Retrospective	Accuracy 96.4%
Misawa *et al*[17], 2018	Japan	Retrospective	Sensitivity, Specificity, and Accuracy were 90%, 63%, and 76%, respectively
Wang *et al*[19], 2018	China	Retrospective	Sensitivity 94.38%, Specificity 95.92%
Su *et al*[41], 2019	China	Prospective	Polyp detection rate was 38.3% as compared to 25.4% in control group ($P < 0.001$)
Wang *et al*[42], 2019	China	Prospective	Polyp detection rate was 45% as compared to 29% in the control group ($P < 0.001$)
Klare *et al*[43], 2019	Germany	Prospective	Larger polyp detection, Odds ration 2.71, $P = 0.042$
Figueiredo *et al*[44], 2019	Portugal	Retrospective	Sensitivity 99.7%, Specificity 84.9%, Accuracy 91.1%
Yamada *et al*[45], 2019	Japan	Retrospective	Sensitivity 97.3%, Specificity: 99%
Lee[46], 2020	South Korea	Retrospective	Accuracy 93.4%, Sensitivity 89.9%, Specificity 93.7%
Luo *et al*[16], 2020	China	Prospective	Polyp detection rate for diminutive polyps increased (38.7% *vs* 34%, $P < 0.001$). No difference was found for larger polyps
Gong[47], 2020	China	Prospective	Polyp detection rate was 47% as compared to 34% in control group ($P = 0.0016$)
Liu *et al*[48], 2020	China	Prospective	Polyp detection rate was 44% as compared to 28% in control group ($P < 0.001$)
Ozawa *et al*[49], 2020	Japan	Retrospective	Sensitivity 92%, PPV 86%, Accuracy 83%
Wang *et al*[50], 2020	China	Prospective	Polyp detection rate was 52% as compared to 37% in control group ($P < 0.0001$)
Hasssan *et al*[51], 2020	Italy	Retrospective	Sensitivity 99.7%
Repici *et al*[52], 2020	Italy	Prospective	Adenoma detection rate was 54.8% as compared to 40.4% in control group ($P < 0.001$)

PPV: Positive predictive value.

In a recent meta-analysis[20] from the researchers in Norway, who included five randomized control trials, AI aided colonoscopy had a ADR of 29.6% as compared to 19.3% without AI. In another recent meta-analysis involving 5 randomized control trials including 4354 patients, ADR was 36.6% with AI aided colonoscopy as compared to 25.2% in the standard control group ($P < 0.01$)[21].

In addition to improvement in colorectal polyp detection, AI has also been shown accuracy in polyp characterization in several studies. Byrne *et al*[22] developed an AI model for real-time characterization of colorectal polyps. They assessed their model using 125 unaltered endoscopic videos containing diminutive polyps. The AI model did not generate sufficient confidence to predict the histology of 19 out of 125 diminutive polyps which was about 15% of the polyps. For the remaining 106 diminutive polyps, the accuracy of the model was 94%, the sensitivity for identification of adenomas was 98%, specificity was 83%, negative predictive value (NPV) was 97%, and positive predictive value (PPV) was 90%. On the other hand, Chen *et al* [23] assessed their model using 284 diminutive polyps. The model identified neoplastic or hyperplastic polyps with 96.3% sensitivity, 78.1% specificity, NPV of 91.5% and PPV of 89.6%. There have been several other studies from across the world analyzing capacity of AI to characterize colon polyps (Table 2).

Table 2 List of studies evaluating role of artificial intelligence in characterization of colon polyps during the colonoscopy

Ref.	Country of origin	Study design	Results
Misawa *et al*[53], 2016	Japan	Retrospective	Sensitivity 84.5%, Specificity 98%
Mori *et al*[54], 2016	Japan	Retrospective	Accuracy 89%
Kominami *et al*[55], 2016	Japan	Prospective	Sensitivity 93%, Specificity 93.3%
Komeda *et al*[56], 2017	Japan	Retrospective	Accuracy 75%
Takeda *et al*[57], 2017	Japan	Retrospective	Sensitivity 89.4%, Specificity 98.9%, Accuracy 94.1 %
Chen *et al*[23], 2018	Taiwan	Retrospective	PPV of 89.6%, and a NPV of 91.5%
Renner[58], 2018	Germany	Retrospective	Sensitivity 92.3% and NPV 88.2%
Mori *et al*[59], 2018	Japan	Prospective	Accuracy 98.1%
Blanes-Vidal *et al*[60], 2019	Denmark	Retrospective	Accuracy 96.4%
Min *et al*[61], 2019	China	Prospective	Sensitivity 83.3%, Specificity 70.1%
Byrne [22], 2019	Canada	Retrospective	Accuracy 94%
Sánchez-Monteset al[62], 2019	Spain	Retrospective	Sensitivity 92.3%, Specificity 89.2%
Horiuchi *et al*[63], 2019	Japan	Prospective	Sensitivity 80%, Specificity 95.3%
Lui *et al*[64], 2019	China	Retrospective	Sensitivity 88.2%, Specificity 77.9%
Ozawa *et al*[49], 2020	Japan	Retrospective	Sensitivity 97%, PPV 84%, NPV 88%
Jin *et al*[65], 2020	South Korea	Prospective	Sensitivity 83.3%, Specificity 91.7%
Rodriguez-Diazet al[66], 2020	United States	Prospective	Sensitivity 96%, Specificity 84%
Kudo *et al*[67], 2020	Japan	Retrospective	Sensitivity 96.9%, Specificity 100%

NPV: Negative predictive value; PPV: Positive predictive value.

AI AND INFLAMMATORY BOWEL DISEASE

Inflammatory bowel disease (IBD) comprises of mainly ulcerative colitis and crohn's disease. It results from complex interplay of environmental, immunological, microbial, and genomic factors[24]. The prevalence of IBD has exceeded 0.3% in the Western countries, and its incidence is rising in newly industrialized countries all over the world[25].

Over the last decade, role of AI has been explored in the field of inflammatory bowel disease (IBD). It has been utilized to analyze endoscopic images for disease diagnosis, grading of severity of disease and predicting treatment response. It has been also utilized to build risk prediction models based on integration of clinical, laboratory as well as gene expression data[26]. There are limited studies exploring the utility of AI aided colonoscopy in the field of IBD. Mosotto *et al* employed machine learning mathematical model of endoscopic and histologic data to distinguish different types of pediatric IBD and found 83.3% accuracy[27]. Similarly, a study from China found AI through machine learning model to be a promising approach specially for unexperienced endoscopists for subtyping of IBD[28].

There are clinical scores available for grading the severity of IBD. AI assisted models have been applied to improve accuracy and precision in assessing the disease severity. In a prospective study from Japan, deep neural network was utilized for evaluating endoscopic images from patients with ulcerative colitis and it showed 90.1% accuracy for endoscopic remission and 92.9% accuracy for histologic remission[29]. In another study from Belgium, computer algorithm for pattern recognition from endoscopic images had significantly better accuracy in determining endoscopic and histologic inflammation in patients with ulcerative colitis[30]. In a retrospective study involving 777 patients with ulcerative colitis, deep learning aided assessment of Mayo endoscopic sub-score for the automated grading of disease yielded 72.4% sensitivity, 85.7% specificity, 77.7% PPV, 87% NPV[31]. Ozawa *et al*[32] constructed a CAD system using convolutional neural network and the results showed better performance for identification of normal mucosa in patients with ulcerative colitis. In a prospective trial, Gottlieb *et al* showed that deep learning algorithm can be used effectively in

predicting ulcerative colitis disease severity[33].

Currently, these AI aided algorithms are mainly used in research setting. Further studies are needed to explore their utility in clinical practice and management of patients with IBD.

LIMITATIONS

One of the possible limitations for the use of CAD could be significantly large number of false positive results[34]. Sometimes CAD system may flag frames which usually endoscopists may never have considered as suspicious area. Thus, the endoscopists may have to spend some extra time to go through all those flagged frames to differentiate between actual false positives and possible false negatives[35]. Additionally, false positive results may lead to unnecessary biopsy and thus related complications which could have been avoided. Hassen *et al*[34] conducted a post hoc analysis of randomized trial comparing colonoscopy with and without CAD to assess relative distribution of false positives in real life setting. During this analysis, two main reasons were found as causes of false positive results, such as artifacts from either mucosal wall or bowel content. Out of total false positives, 88% were due to artifacts from bowel wall, while 12% were due to artifacts from bowel content. However, most of the false positives were rejected by endoscopists right away and there was only 1% increase in the total withdrawal time due to false positives. Another limiting factor is cost effectiveness of the use of AI in colonoscopy, and it needs to be established. Also, the impact of the use of AI in colonoscopy on long-term clinical outcomes, such as decrease in CRC rate or increase in surveillance interval for colonoscopy is not known [35]. We require long-term prospective cohort studies to address these issues.

FUTURE DIRECTIONS

Food and Drug Administration has recently approved the first real-time CAD system for colonoscopy in April 2021, known as gastrointestinal (GI) Genius. It can identify the regions of the colon within the endoscope's field of view where a colorectal polyp might be located, allowing for a more extended examination in real time during colonoscopy. After getting the alert from the device, it is up to the clinician to decide whether the identified region contains a suspected lesion, and how the lesion should be managed and processed per standard clinical practice and guidelines. However, GI Genius is not intended to characterize or classify a lesion, nor to replace lab sampling as a means of diagnosis. The device does not provide any diagnostic assessments of colorectal polyp pathology, nor does it suggest to the clinician how to manage suspicious polyps.

Although many studies have shown good results but most of these studies were retrospective studies which could be subject to considerable selection bias. On the other hand, only few prospective studies are available till date which are more statistically significant than retrospective studies. Thus, we need to design more prospective studies and should be directed towards polyp characterization during real-time colonoscopy. Additionally, future studies can explore AI assisted identification of polyps with submucosal invasion. The prospect of a fully automated independent colonoscopy system is still too premature at this stage. Furthermore, trials to build more cost-effective models should be conducted in near future before considering use of CAD assisted colonoscopy widespread in daily practice.

CONCLUSION

In conclusion, utility of AI methods and algorithms have significantly evolved over the last decade. AI technology provides us a very robust tool to improve the accuracy and precision during the colonoscopy. ML models of AI technology provide us a valuable tool to transform the healthcare. Further larger and prospective studies are needed to see if these positive outcomes can be replicated in a cost-effective manner in clinical practice.

REFERENCES

1 **Gottfredson LS**. Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence* 1997; **24**: 13-23 [DOI: 10.1016/S0160-2896(97)90011-8]
2 **Conrad M**. The brain-machine disanalogy. *Biosystems* 1989; **22**: 197-213 [PMID: 2650754 DOI: 10.1016/0303-2647(89)90061-0]
3 **Lee RST**. Artificial Intelligence in Daily Life. 1st ed. Singapore: Springer, 2020: 41-70 [DOI: 10.1007/978-981-15-7695-9]
4 **Hamet P**, Tremblay J. Artificial intelligence in medicine. *Metabolism* 2017; **69S**: S36-S40 [PMID: 28126242 DOI: 10.1016/j.metabol.2017.01.011]
5 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]
6 **Corley DA**, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 1298-1306 [PMID: 24693890 DOI: 10.1056/NEJMoa1309086]
7 **Aslanian HR**, Shieh FK, Chan FW, Ciarleglio MM, Deng Y, Rogart JN, Jamidar PA, Siddiqui UD. Nurse observation during colonoscopy increases polyp detection: a randomized prospective study. *Am J Gastroenterol* 2013; **108**: 166-172 [PMID: 23381064 DOI: 10.1038/ajg.2012.237]
8 **Lee CK**, Park DI, Lee SH, Hwangbo Y, Eun CS, Han DS, Cha JM, Lee BI, Shin JE. Participation by experienced endoscopy nurses increases the detection rate of colon polyps during a screening colonoscopy: a multicenter, prospective, randomized study. *Gastrointest Endosc* 2011; **74**: 1094-1102 [PMID: 21889137 DOI: 10.1016/j.gie.2011.06.033]
9 **Liedlgruber M**, Uhl A. Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review. *IEEE Rev Biomed Eng* 2011; **4**: 73-88 [PMID: 22273792 DOI: 10.1109/RBME.2011.2175445]
10 **Karkanis SA**, Iakovidis DK, Maroulis DE, Karras DA, Tzivras M. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE Trans Inf Technol Biomed* 2003; **7**: 141-152 [PMID: 14518727 DOI: 10.1109/titb.2003.813794]
11 **Bernal J**, Vilarino F, Sánchez J. Towards intelligent systems for colonoscopy In: Miskovitz P. Colonoscopy. InTech, 2011: 245-270
12 **Bernal J**, Sánchez FJ, Fernández-Esparrach G, Gil D, Rodríguez C, Vilariño F. WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation *vs* saliency maps from physicians. *Comput Med Imaging Graph* 2015; **43**: 99-111 [PMID: 25863519 DOI: 10.1016/j.compmedimag.2015.02.007]
13 **Fernández-Esparrach G**, Bernal J, López-Cerón M, Córdova H, Sánchez-Montes C, Rodríguez de Miguel C, Sánchez FJ. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy* 2016; **48**: 837-842 [PMID: 27285900 DOI: 10.1055/s-0042-108434]
14 **Mori Y**, Kudo SE, Berzin TM, Misawa M, Takeda K. Computer-aided diagnosis for colonoscopy. *Endoscopy* 2017; **49**: 813-819 [PMID: 28561195 DOI: 10.1055/s-0043-109430]
15 **Kudo SE**, Mori Y, Misawa M, Takeda K, Kudo T, Itoh H, Oda M, Mori K. Artificial intelligence and colonoscopy: Current status and future perspectives. *Dig Endosc* 2019; **31**: 363-371 [PMID: 30624835 DOI: 10.1111/den.13340]
16 **Luo Y**, Zhang Y, Liu M, Lai Y, Liu P, Wang Z, Xing T, Huang Y, Li Y, Li A, Wang Y, Luo X, Liu S, Han Z. Artificial Intelligence-Assisted Colonoscopy for Detection of Colon Polyps: a Prospective, Randomized Cohort Study. *J Gastrointest Surg* 2020; epub ahead of print [PMID: 32968933 DOI: 10.1007/s11605-020-04802-4]
17 **Misawa M**, Kudo SE, Mori Y, Cho T, Kataoka S, Yamauchi A, Ogawa Y, Maeda Y, Takeda K, Ichimasa K, Nakamura H, Yagawa Y, Toyoshima N, Ogata N, Kudo T, Hisayuki T, Hayashi T, Wakamura K, Baba T, Ishida F, Itoh H, Roth H, Oda M, Mori K. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. *Gastroenterology* 2018; **154**: 2027-2029. e3 [PMID: 29653147 DOI: 10.1053/j.gastro.2018.04.003]
18 **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078. e8 [PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037]
19 **Wang P**, Xiao X, Glissen Brown JR, Berzin TM, Tu M, Xiong F, Hu X, Liu P, Song Y, Zhang D, Yang X, Li L, He J, Yi X, Liu J, Liu X. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2018; **2**: 741-748 [PMID: 31015647 DOI: 10.1038/s41551-018-0301-3]
20 **Barua I**, Vinsard DG, Jodal HC, Løberg M, Kalager M, Holme Ø, Misawa M, Bretthauer M, Mori Y. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. *Endoscopy* 2021; **53**: 277-284 [PMID: 32557490 DOI: 10.1055/a-1201-7165]
21 **Hassan C**, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, Antonelli G, Yu H, Areia M, Dinis-Ribeiro M, Bhandari P, Sharma P, Rex DK, Rösch T, Wallace M, Repici A. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021; **93**: 77-85. e6 [PMID: 32598963 DOI: 10.1016/j.gie.2020.06.059]

22 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. 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 [PMID: 29066576 DOI: 10.1136/gutjnl-2017-314547]

23 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: 29042219 DOI: 10.1053/j.gastro.2017.10.010]

24 **de Souza HSP**, Fiocchi C, Iliopoulos D. The IBD interactome: an integrated view of aetiology, pathogenesis and therapy. *Nat Rev Gastroenterol Hepatol* 2017; **14**: 739-749 [PMID: 28831186 DOI: 10.1038/nrgastro.2017.110]

25 **Ng SC**, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, Panaccione R, Ghosh S, Wu JCY, Chan FKL, Sung JJY, Kaplan GG. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2017; **390**: 2769-2778 [PMID: 29050646 DOI: 10.1016/S0140-6736(17)32448-0]

26 **Gubatan J**, Levitte S, Patel A, Balabanis T, Wei MT, Sinha SR. Artificial intelligence applications in inflammatory bowel disease: Emerging technologies and future directions. *World J Gastroenterol* 2021; **27**: 1920-1935 [PMID: 34007130 DOI: 10.3748/wjg.v27.i17.1920]

27 **Mossotto E**, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep* 2017; **7**: 2427 [PMID: 28546534 DOI: 10.1038/s41598-017-02606-2]

28 **Tong Y**, Lu K, Yang Y, Li J, Lin Y, Wu D, Yang A, Li Y, Yu S, Qian J. Can natural language processing help differentiate inflammatory intestinal diseases in China? *BMC Med Inform Decis Mak* 2020; **20**: 248 [PMID: 32993636 DOI: 10.1186/s12911-020-01277-w]

29 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: 32060000 DOI: 10.1053/j.gastro.2020.02.012]

30 **Bossuyt P**, Nakase H, Vermeire S, de Hertogh G, Eelbode T, Ferrante M, Hasegawa T, Willekens H, Ikemoto Y, Makino T, Bisschops R. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut* 2020; **69**: 1778-1786 [PMID: 31915237 DOI: 10.1136/gutjnl-2019-320056]

31 **Bhambhvani HP**, Zamora A. Deep learning enabled classification of Mayo endoscopic subscore in patients with ulcerative colitis. *Eur J Gastroenterol Hepatol* 2021; **33**: 645-649 [PMID: 33079775 DOI: 10.1097/MEG.0000000000001952]

32 **Ozawa T**, Ishihara S, Fujishiro M, Saito H, Kumagai Y, Shichijo S, Aoyama K, Tada T. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc* 2019; **89**: 416-421. e1 [PMID: 30367878 DOI: 10.1016/j.gie.2018.10.020]

33 **Gottlieb K**, Requa J, Karnes W, Chandra Gudivada R, Shen J, Rael E, Arora V, Dao T, Ninh A, McGill J. Central Reading of Ulcerative Colitis Clinical Trial Videos Using Neural Networks. *Gastroenterology* 2021; **160**: 710-719. e2 [PMID: 33098883 DOI: 10.1053/j.gastro.2020.10.024]

34 **Hassan C**, Badalamenti M, Maselli R, Correale L, Iannone A, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rösch T, Repici A. Computer-aided detection-assisted colonoscopy: classification and relevance of false positives. *Gastrointest Endosc* 2020; **92**: 900-904. e4 [PMID: 32561410 DOI: 10.1016/j.gie.2020.06.021]

35 **Antonelli G**, Gkolfakis P, Tziatzios G, Papanikolaou IS, Triantafyllou K, Hassan C. Artificial intelligence-aided colonoscopy: Recent developments and future perspectives. *World J Gastroenterol* 2020; **26**: 7436-7443 [PMID: 33384546 DOI: 10.3748/wjg.v26.i47.7436]

36 **k G**, c R. Automatic Colorectal Polyp Detection in Colonoscopy Video Frames Asian Pac J Cancer Prev 2016; 17: 4869-4873 [PMID: 28030914 DOI: 10.22034/APJCP.2016.17.11.4869]

37 **Misawa M**, Kudo SE, Mori Y, Takeda K, Maeda Y, Kataoka S, Nakamura H, Kudo T, Wakamura K, Hayashi T, Katagiri A, Baba T, Ishida F, Inoue H, Nimura Y, Oda M, Mori K. Accuracy of computer-aided diagnosis based on narrow-band imaging endocytoscopy for diagnosing colorectal lesions: comparison with experts. *Int J Comput Assist Radiol Surg* 2017; **12**: 757-766 [PMID: 28247214 DOI: 10.1007/s11548-017-1542-4]

38 **Zhang R**, Zheng Y, Mak TW, Yu R, Wong SH, Lau JY, Poon CC. Automatic Detection and Classification of Colorectal Polyps by Transferring Low-Level CNN Features From Nonmedical Domain. *IEEE J Biomed Health Inform* 2017; **21**: 41-47 [PMID: 28114040 DOI: 10.1109/JBHI.2016.2635662]

39 **Lequan Yu**, Hao Chen, Qi Dou, Jing Qin, Pheng Ann Heng. Integrating Online and Offline Three-Dimensional Deep Learning for Automated Polyp Detection in Colonoscopy Videos. *IEEE J Biomed Health Inform* 2017; **21**: 65-75 [PMID: 28114049 DOI: 10.1109/JBHI.2016.2637004]

40 **Billah M**, Waheed S, Rahman MM. An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features. *Int J Biomed Imaging* 2017; **2017**: 9545920 [PMID: 28894460 DOI: 10.1155/2017/9545920]

41 **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective

randomized controlled study (with videos). *Gastrointest Endosc* 2020; **91**: 415-424. e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

42 **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

43 **Klare P**, Sander C, Prinzen M, Haller B, Nowack S, Abdelhafez M, Poszler A, Brown H, Wilhelm D, Schmid RM, von Delius S, Wittenberg T. Automated polyp detection in the colorectum: a prospective study (with videos). *Gastrointest Endosc* 2019; **89**: 576-582. e1 [PMID: 30342029 DOI: 10.1016/j.gie.2018.09.042]

44 **Figueiredo PN**, Figueiredo IN, Pinto L, Kumar S, Tsai YR, Mamonov AV. Polyp detection with computer-aided diagnosis in white light colonoscopy: comparison of three different methods. *Endosc Int Open* 2019; **7**: E209-E215 [PMID: 30705955 DOI: 10.1055/a-0808-4456]

45 **Yamada M**, Saito Y, Imaoka H, Saiko M, Yamada S, Kondo H, Takamaru H, Sakamoto T, Sese J, Kuchiba A, Shibata T, Hamamoto R. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Sci Rep* 2019; **9**: 14465 [PMID: 31594962 DOI: 10.1038/s41598-019-50567-5]

46 **Lee JY**, Jeong J, Song EM, Ha C, Lee HJ, Koo JE, Yang DH, Kim N, Byeon JS. Real-time detection of colon polyps during colonoscopy using deep learning: systematic validation with four independent datasets. *Sci Rep* 2020; **10**: 8379 [PMID: 32433506 DOI: 10.1038/s41598-020-65387-1]

47 **Gong D**, Wu L, Zhang J, Mu G, Shen L, Liu J, Wang Z, Zhou W, An P, Huang X, Jiang X, Li Y, Wan X, Hu S, Chen Y, Hu X, Xu Y, Zhu X, Li S, Yao L, He X, Chen D, Huang L, Wei X, Wang X, Yu H. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020; **5**: 352-361 [PMID: 31981518 DOI: 10.1016/S2468-1253(19)30413-3]

48 **Liu WN**, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, Huang J. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; **26**: 13-19 [PMID: 31898644 DOI: 10.4103/sjg.SJG_377_19]

49 **Ozawa T**, Ishihara S, Fujishiro M, Kumagai Y, Shichijo S, Tada T. Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. *Therap Adv Gastroenterol* 2020; **13**: 1756284820910659 [PMID: 32231710 DOI: 10.1177/1756284820910659]

50 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

51 **Hassan C**, Wallace MB, Sharma P, Maselli R, Craviotto V, Spadaccini M, Repici A. New artificial intelligence system: first validation study versus experienced endoscopists for colorectal polyp detection. *Gut* 2020; **69**: 799-800 [PMID: 31615835 DOI: 10.1136/gutjnl-2019-319914]

52 **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology* 2020; **159**: 512-520. e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

53 **Misawa M**, Kudo SE, Mori Y, Nakamura H, Kataoka S, Maeda Y, Kudo T, Hayashi T, Wakamura K, Miyachi H, Katagiri A, Baba T, Ishida F, Inoue H, Nimura Y, Mori K. Characterization of Colorectal Lesions Using a Computer-Aided Diagnostic System for Narrow-Band Imaging Endocytoscopy. *Gastroenterology* 2016; **150**: 1531-1532. e3 [PMID: 27072671 DOI: 10.1053/j.gastro.2016.04.004]

54 **Mori Y**, Kudo SE, Chiu PW, Singh R, Misawa M, Wakamura K, Kudo T, Hayashi T, Katagiri A, Miyachi H, Ishida F, Maeda Y, Inoue H, Nimura Y, Oda M, Mori K. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. *Endoscopy* 2016; **48**: 1110-1118 [PMID: 27494455 DOI: 10.1055/s-0042-113609]

55 **Kominami Y**, Yoshida S, Tanaka S, Sanomura Y, Hirakawa T, Raytchev B, Tamaki T, Koide T, Kaneda K, Chayama K. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; **83**: 643-649 [PMID: 26264431 DOI: 10.1016/j.gie.2015.08.004]

56 **Komeda Y**, Handa H, Watanabe T, Nomura T, Kitahashi M, Sakurai T, Okamoto A, Minami T, Kono M, Arizumi T, Takenaka M, Hagiwara S, Matsui S, Nishida N, Kashida H, Kudo M. Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience. *Oncology* 2017; **93** Suppl 1: 30-34 [PMID: 29258081 DOI: 10.1159/000481227]

57 **Takeda K**, Kudo SE, Mori Y, Misawa M, Kudo T, Wakamura K, Katagiri A, Baba T, Hidaka E, Ishida F, Inoue H, Oda M, Mori K. Accuracy of diagnosing invasive colorectal cancer using computer-aided endocytoscopy. *Endoscopy* 2017; **49**: 798-802 [PMID: 28472832 DOI: 10.1055/s-0043-105486]

58 **Renner J**, Phlipsen H, Haller B, Navarro-Avila F, Saint-Hill-Febles Y, Mateus D, Ponchon T, Poszler A, Abdelhafez M, Schmid RM, von Delius S, Klare P. Optical classification of neoplastic colorectal polyps - a computer-assisted approach (the COACH study). *Scand J Gastroenterol* 2018; **53**: 1100-1106 [PMID: 30270677 DOI: 10.1080/00365521.2018.1501092]

59 **Mori Y**, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S,

Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018; **169**: 357-366 [PMID: 30105375 DOI: 10.7326/M18-0249]

60 **Blanes-Vidal V**, Baatrup G, Nadimi ES. Addressing priority challenges in the detection and assessment of colorectal polyps from capsule endoscopy and colonoscopy in colorectal cancer screening using machine learning. *Acta Oncol* 2019; **58**: S29-S36 [PMID: 30836800 DOI: 10.1080/0284186X.2019.1584404]

61 **Min M**, Su S, He W, Bi Y, Ma Z, Liu Y. Computer-aided diagnosis of colorectal polyps using linked color imaging colonoscopy to predict histology. *Sci Rep* 2019; **9**: 2881 [PMID: 30814661 DOI: 10.1038/s41598-019-39416-7]

62 **Sánchez-Montes C**, Sánchez FJ, Bernal J, Córdova H, López-Cerón M, Cuatrecasas M, Rodríguez de Miguel C, García-Rodríguez A, Garcés-Durán R, Pellisé M, Llach J, Fernández-Esparrach G. Computer-aided prediction of polyp histology on white light colonoscopy using surface pattern analysis. *Endoscopy* 2019; **51**: 261-265 [PMID: 30360010 DOI: 10.1055/a-0732-5250]

63 **Horiuchi H**, Tamai N, Kamba S, Inomata H, Ohya TR, Sumiyama K. Real-time computer-aided diagnosis of diminutive rectosigmoid polyps using an auto-fluorescence imaging system and novel color intensity analysis software. *Scand J Gastroenterol* 2019; **54**: 800-805 [PMID: 31195905 DOI: 10.1080/00365521.2019.1627407]

64 **Lui TKL**, Wong KKY, Mak LLY, Ko MKL, Tsao SKK, Leung WK. Endoscopic prediction of deeply submucosal invasive carcinoma with use of artificial intelligence. *Endosc Int Open* 2019; **7**: E514-E520 [PMID: 31041367 DOI: 10.1055/a-0849-9548]

65 **Jin EH**, Lee D, Bae JH, Kang HY, Kwak MS, Seo JY, Yang JI, Yang SY, Lim SH, Yim JY, Lim JH, Chung GE, Chung SJ, Choi JM, Han YM, Kang SJ, Lee J, Chan Kim H, Kim JS. Improved Accuracy in Optical Diagnosis of Colorectal Polyps Using Convolutional Neural Networks with Visual Explanations. *Gastroenterology* 2020; **158**: 2169-2179. e8 [PMID: 32119927 DOI: 10.1053/j.gastro.2020.02.036]

66 **Rodriguez-Diaz E**, Baffy G, Lo WK, Mashimo H, Vidyarthi G, Mohapatra SS, Singh SK. Real-time artificial intelligence-based histologic classification of colorectal polyps with augmented visualization. *Gastrointest Endosc* 2021; **93**: 662-670 [PMID: 32949567 DOI: 10.1016/j.gie.2020.09.018]

67 **Kudo SE**, Misawa M, Mori Y, Hotta K, Ohtsuka K, Ikematsu H, Saito Y, Takeda K, Nakamura H, Ichimasa K, Ishigaki T, Toyoshima N, Kudo T, Hayashi T, Wakamura K, Baba T, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Artificial Intelligence-assisted System Improves Endoscopic Identification of Colorectal Neoplasms. *Clin Gastroenterol Hepatol* 2020; **18**: 1874-1881. e2 [PMID: 31525512 DOI: 10.1016/j.cgh.2019.09.009]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 June 28; 2(3): 89-94

DOI: 10.37126/aige.v2.i3.89

ISSN 2689-7164 (online)

MINIREVIEWS

Use of artificial intelligence in endoscopic ultrasound evaluation of pancreatic pathologies

Ravinder Mankoo, Ahmad H Ali, Ghassan M Hammoud

ORCID number: Ravinder Mankoo 0000-0001-7581-2167; Ahmad H Ali 0000-0001-8553-7744; Ghassan M Hammoud 0000-0003-3551-0652.

Author contributions: Mankoo R performed the majority of the writing, gathered and reviewed references and prepared the figure; Ali AH performed writing and revision of the manuscript; Hammoud GM performed writing, reviewed references and performed a final revision of the manuscript.

Conflict-of-interest statement: There is no conflict of interest associated with any of the authors of this manuscript.

Manuscript source: Invited

Ravinder Mankoo, Ahmad H Ali, Ghassan M Hammoud, Division of Gastroenterology and Hepatology, University of Missouri School of Medicine, Columbia, MO 65212, United States

Corresponding author: Ghassan M Hammoud, MD, MPH, Professor, Division of Gastroenterology and Hepatology, University of Missouri School of Medicine, One hospital Dr, Columbia, MO 65212, United States. hammoudg@health.missouri.edu

Abstract

The application of artificial intelligence (AI) using deep learning and machine learning approaches in modern medicine is rapidly expanding. Within the field of Gastroenterology, AI is being evaluated across a breadth of clinical and diagnostic applications including identification of pathology, differentiation of disease processes, and even automated procedure report generation. Many pancreatic pathologies can have overlapping features creating a diagnostic dilemma that provides a window for AI-assisted improvement in current evaluation and diagnosis, particularly using endoscopic ultrasound. This topic highlight will review the basics of AI, history of AI in gastrointestinal endoscopy, and prospects for AI in the evaluation of autoimmune pancreatitis, pancreatic ductal adenocarcinoma, chronic pancreatitis and intraductal papillary mucinous neoplasm.

Key Words: Artificial intelligence; Deep learning; Machine learning; Convolutional neural network; Endoscopic ultrasound; Autoimmune pancreatitis; Pancreatic ductal adenocarcinoma; Chronic pancreatitis; Intraductal papillary mucinous neoplasm

Core Tip: Artificial intelligence is an emerging diagnostic tool that may further aid clinicians in the current evaluation of diseases of the pancreas.

Citation: Mankoo R, Ali AH, Hammoud GM. Use of artificial intelligence in endoscopic ultrasound evaluation of pancreatic pathologies. *Artif Intell Gastrointest Endosc* 2021; 2(3): 89-94

URL: https://www.wjgnet.com/2689-7164/full/v2/i3/89.htm

DOI: https://dx.doi.org/10.37126/aige.v2.i3.89

manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: United States

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: June 2, 2021
Peer-review started: June 2, 2021
First decision: June 18, 2021
Revised: June 20, 2021
Accepted: June 28, 2021
Article in press: June 28, 2021
Published online: June 28, 2021

P-Reviewer: Byeon H, Imai Y
S-Editor: Liu M
L-Editor: A
P-Editor: Wang LYT

INTRODUCTION

Artificial intelligence (AI) has emerged as a mechanism to assist clinicians, particularly in the analysis and interpretation of clinical data such as radiologic images and pathology. In general, AI encompasses the use of computer algorithms and learning models designed to complete undertakings that typically require conscious human processing[1]. For pattern recognition in images, a deep neural network learns multiple representations of the input images at different levels of abstractions. Subsets of AI include machine learning (support vector machine algorithms, artificial neural networks) and direct learning (convolutional neural networks, recurrent neural networks[1-3]. Deep learning has shown great promise in healthcare applications ranging from early detection of cancers to predicting disease survivability. The overarching goal of AI in medicine has been to decrease inter-operator variability while improving diagnostic accuracy and real-time decision making[4]. The application of AI in Gastroenterology has largely been focused on endoscopy, ranging from the detection and classification of colon polyps, to the diagnosis of esophageal and gastric cancer[1,3]. However, more recently there has been further evaluation of the role of AI in biliopancreatic endoscopy, including improved endoscopic ultrasound (EUS) differentiation between pancreatic ductal adenocarcinoma (PDAC) and other pancreatic pathologies such as autoimmune pancreatitis (AIP), chronic pancreatitis (CP) and cystic pancreas lesions such as intraductal papillary mucinous neoplasm (IPMN). This "topic highlight" will focus on the potential use of AI in the EUS evaluation of pancreatic conditions.

HISTORY OF AI IN GASTROINTESTINAL ENDOSCOPY

Early studies on the application of AI in GI endoscopy dating back to the 1990s-2000s were focused on aiding the detection and classification of colorectal polyps to improve adenoma detection rates and decrease interval colon cancers[5-8]. Additional studies have used AI to help diagnose inflammatory bowel disease and predict histologic inflammation during colonoscopy evaluation[9,10], as well as grade bowel preparation [11]. The use of AI in upper endoscopy has been assessed in the identification and labeling of basic anatomic structures with automatic image capture[12], diagnosis of Helicobacter pylori infection[13], identification of gastric and esophageal cancer[14], as well as diagnosis of dysplasia in Barrett's esophagus[15]. With regards to capsule endoscopy, existing technology within current software platforms allows for removal of redundant or uninformative images and identifies potential images of bleeding through color detection, while more recent studies are looking into the use of AI to identify other small bowel pathologies[16]. PDAC and AIP are diseases with a highly analogous visual presentation that are difficult to distinguish by imaging. AI systems have been developed to aid EUS evaluation of pancreatic lesions with the particular goal of distinguishing pancreatic cancer from other pancreatic pathologies including CP and AIP[17-19].

AI IN PANCREATICOBILIARY ENDOSCOPY

The use of AI in pancreaticobiliary endoscopy is still in its infancy, therefore there is a paucity of literature related to EUS evaluation of pancreatic conditions using AI-based systems. However, the need for improved diagnostic evaluation of pancreatic conditions including AIP, PDAC, CP and pancreatic cystic lesions, provides an exciting niche for further research. AI has previously been applied in EUS differentiation of pancreatic cystic lesions and pancreatic tumors, thereby offering the capability of earlier and more accurate diagnosis. Both conventional machine learning and deep learning architectures have been used. A convolutional neural network (CNN) is a deep learning algorithm developed based on the concepts of visual tasks and signaling. In building a CNN for EUS, initial image data is collected and labeled based on the findings, these images are then entered as input and filtered through a multi-layer deep learning program which allows the system to learn key features of the provided EUS images. Multiple rounds of this process allow for the formation of a neural network where the system can then apply the previously learned features in analyzing novel images (Figure 1).

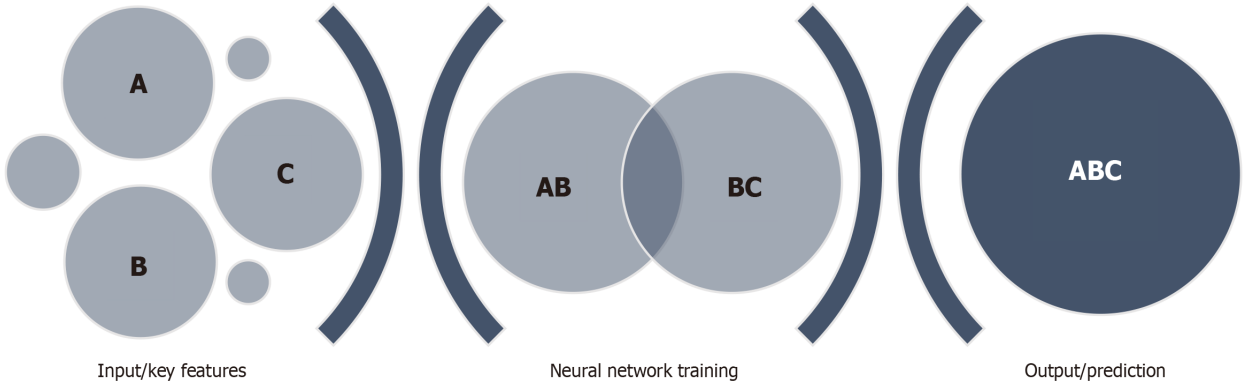

Figure 1 Example of neural network design.

LITERATURE SEARCH

To identify relevant literature on this topic, we searched the PubMed database through our institution's library for articles combining the terms "autoimmune pancreatitis", "pancreatic adenocarcinoma", "chronic pancreatitis", "intraductal papillary mucinous neoplasm", "artificial intelligence", and "endoscopic ultrasound".

AI IN THE EVALUATION OF AUTOIMMUNE PANCREATITIS

Autoimmune pancreatitis is an inflammatory condition of the pancreas commonly associated with a constellation of findings referred to as immunoglobulin G4-related disease. AIP is characterized radiologically/endoscopically by diffuse or focal enlargement of the pancreas parenchyma and diffuse irregular narrowing of the main pancreatic duct, histologically by pancreatic fibrosis and lymphoplasmacytic infiltration, and serologically by increased levels of serum gamma globulin, including immunoglobulin G4 (IgG4)[20,21]. The diagnosis of AIP can be challenging due to the overlap of clinical, laboratory and imaging findings with those of PDAC[22-24]. Studies have shown that 2%-5% of patients who undergo pancreatic resection of suspected cancer are found to have AIP on histopathologic evaluation, and instead of receiving highly effective immunosuppressive therapy such as corticosteroids, these patients are left to manage the morbidity associated with an invasive surgery[25,26]. While EUS remains the preeminent diagnostic tool in evaluating pancreatic diseases, the yield of needle aspiration/biopsy techniques can be inconclusive or non-specific, creating a diagnostic dilemma that may ultimately delay or compromise patient care [25-28].

In late 2020, Marya *et al*[22] published novel research on the development of EUS-based AI to improve the diagnosis of AIP. Using a CNN built from a large collection of EUS images and videos (583 patients: 146 AIP, 292 PDAC, 72 CP, 73 normal pancreas), their team sought to develop a reliable, real-time method of distinguishing AIP from PDAC on EUS evaluation. Going one step further, they also used occlusion heatmapping to identify key sonographic features of AIP compared to PDAC, further strengthening the utility of their model. On combined still image and continuous video image analysis, the developed CNN was able to distinguish AIP from PDAC with 90% sensitivity and 87% specificity; and distinguish AIP from all other studied diagnoses (PDAC, CP, normal pancreas) with 90% sensitivity and 78% specificity. On continuous video image analysis, the developed CNN was able to successfully differentiate AIP from PDAC with a sensitivity of 90% and specificity of 93%; and differentiate AIP from all other studied diagnoses with a sensitivity of 90% and specificity of 85%. Furthermore, occlusion heatmap evaluation showed that "enhanced hyperechoic interfaces between pancreas parenchyma and pancreas duct/vessels" were predictive of AIP, and "post-acoustic enhancement deep to a dilated pancreas duct" was more commonly associated with PDAC. In addition, the study evaluated the accuracy of diagnosis between the CNN and a group of expert endosonographers, showing that the CNN correctly diagnosed AIP with a sensitivity of 88.2% and specificity of 82.5%, while expert endosonographers correctly diagnosed AIP with a sensitivity of 53.8% and specificity of 86.7%. Overall, this study serves as a model for the application of AI in the EUS evaluation of pancreatic pathologies including AIP.

AI IN THE EVALUATION OF CHRONIC PANCREATITIS

CP is an irreversible fibro-inflammatory condition caused by recurrent or persistent pancreatic parenchymal injury[29]. The diagnosis of CP is often made by analyzing a patient's risk factors, radiographic imaging results and direct/indirect pancreatic function laboratory tests. EUS-guided tissue acquisition still serves as the gold standard for CP diagnosis when less invasive tools are inconclusive, however, studies have found similar sensitivities and specificities in the diagnosis of CP using EUS, MRI or CT[30]. This again identifies another diagnostic dilemma for which AI may serve a role to improve diagnostic accuracy, thereby improving patient care and outcomes.

Computer aided diagnosis based on digital image analysis (DIA) was initially utilized in a small study attempting to differentiate between focal, pseudotumorous pancreatitis and pancreatic malignancy with an overall diagnostic accuracy of 89% [31]. In 2008, Săftoiu *et al* developed a neural network to differentiate between CP and pancreatic malignancy through imaging features of EUS-elastography, further expanding to include the evaluation of contrast-enhanced EUS images in 2015[19]. Their initial system was able to differentiate between malignant and benign pancreatic masses with a sensitivity of 91.4%, specificity of 87.9% and accuracy of 89.7%. Das *et al* [32] used DIA of the spatial distribution of pixels on EUS images to create a neural network that could differentiate PDAC and CP with a 93% accuracy. In 2013, Zhu *et al* [33] published data on the use of a support vector machine predictive model to differentiate PDAC and CP based on EUS images which achieved a diagnostic accuracy of 94%. Overall, these studies provide positive reinforcement to the notion that AI can improve EUS differentiation of pancreatic malignancy from other pathologies including CP.

AI IN THE EVALUATION OF INTRADUCTAL PAPILLARY MUCINOUS NEOPLASMS

With the increasing detection of pancreatic cystic lesions on cross-sectional imaging, IPMNs have become an important pancreatic pathology given their potential for malignant transformation[34]. Early resection of IPMNs, particularly those with high grade dysplasia limit the progression to PDAC. International consensus guidelines for IPMN management have identified high risk stigmata (*i.e.*, obstructive jaundice) and worrisome features (size > 3 cm, enhancing mural nodule < 5 mm, thickened cyst wall, MPD > 5-9 mm, abrupt change in MPD diameter) of malignancy associated with IPMN[34]. However, the use of these features alone to differentiate benign *vs* malignant IPMN leaves room for improvement, particularly through the use of AI-assisted EUS evaluation. In 2019, Kuwahara *et al*[35] performed a retrospective single-center study that developed an EUS-based CNN to differentiate benign *vs* malignant IPMNs. Their model identified malignant IPMNs with a diagnostic accuracy of 94%, compared to the human pre-operative diagnosis control group based on consensus guidelines which had an accuracy of 56%. While further research in this area is needed, the overarching theme of improved diagnostic accuracy when AI is applied to EUS evaluation of pancreatic disease appears to be evident.

CONCLUSION

The diagnosis of pancreatic lesions can be difficult, often stemming from the overlap of features found in benign lesions with those found in PDAC. The development of improved diagnostic tools to differentiate PDAC from other pancreatic lesions presents an opportunity for significant impact on the overall care of patients with pancreatic disease. More robust studies are needed to validate the current available research, namely in the form of prospective, multicenter studies which may further determine the generalizability of current models and the overall, real-time clinical application of these AI systems. It should be noted that standardization of endoscopic image capture and reporting may better help facilitate future interdisciplinary work in this field[36,37]. While the use of AI to evaluate the pancreas appears to be in its early stages, the potential for AI-assisted EUS assessment provides an exciting and promising future for the diagnosis and management of pancreatic lesions.

REFERENCES

1 **Pannala R**, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, Trikudanathan G, Trindade AJ, Watson RR, Maple JT, Lichtenstein DR. Artificial intelligence in gastrointestinal endoscopy. *VideoGIE* 2020; **5**: 598-613 [PMID: 33319126 DOI: 10.1016/j.vgie.2020.08.013]

2 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

3 **Abadir AP**, Ali MF, Karnes W, Samarasena JB. Artificial Intelligence in Gastrointestinal Endoscopy. *Clin Endosc* 2020; 132-141 [PMID: 32252506 DOI: 10.5946/ce.2020.038]

4 **Akshintala VS**, Khashab MA. Artificial intelligence in pancreaticobiliary endoscopy. *J Gastroenterol Hepatol* 2021; **36**: 25-30 [PMID: 33448514 DOI: 10.1111/jgh.15343]

5 **Fernández-Esparrach G**, Bernal J, López-Cerón M, Córdova H, Sánchez-Montes C, Rodríguez de Miguel C, Sánchez FJ. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy* 2016; **48**: 837-842 [PMID: 27285900 DOI: 10.1055/s-0042-108434]

6 **Tajbakhsh N**, Gurudu SR, Liang J. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information. *IEEE Trans Med Imaging* 2016; **35**: 630-644 [PMID: 26462083 DOI: 10.1109/TMI.2015.2487997]

7 **Takemura Y**, Yoshida S, Tanaka S, Onji K, Oka S, Tamaki T, Kaneda K, Yoshihara M, Chayama K. Quantitative analysis and development of a computer-aided system for identification of regular pit patterns of colorectal lesions. *Gastrointest Endosc* 2010; **72**: 1047-1051 [PMID: 21034905 DOI: 10.1016/j.gie.2010.07.037]

8 **Tischendorf JJ**, Gross S, Winograd R, Hecker H, Auer R, Behrens A, Trautwein C, Aach T, Stehle T. Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. *Endoscopy* 2010; **42**: 203-207 [PMID: 20101564 DOI: 10.1055/s-0029-1243861]

9 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

10 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: 32060000 DOI: 10.1053/j.gastro.2020.02.012]

11 **Zhou J**, Wu L, Wan X, Shen L, Liu J, Zhang J, Jiang X, Wang Z, Yu S, Kang J, Li M, Hu S, Hu X, Gong D, Chen D, Yao L, Zhu Y, Yu H. A novel artificial intelligence system for the assessment of bowel preparation (with video). *Gastrointest Endosc* 2020; **91**: 428-435. e2 [PMID: 31783029 DOI: 10.1016/j.gie.2019.11.026]

12 **Wu L**, Zhang J, Zhou W, An P, Shen L, Liu J, Jiang X, Huang X, Mu G, Wan X, Lv X, Gao J, Cui N, Hu S, Chen Y, Hu X, Li J, Chen D, Gong D, He X, Ding Q, Zhu X, Li S, Wei X, Li X, Wang X, Zhou J, Zhang M, Yu HG. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut* 2019; **68**: 2161-2169 [PMID: 30858305 DOI: 10.1136/gutjnl-2018-317366]

13 **Zheng W**, Zhang X, Kim JJ, Zhu X, Ye G, Ye B, Wang J, Luo S, Li J, Yu T, Liu J, Hu W, Si J. High Accuracy of Convolutional Neural Network for Evaluation of Helicobacter pylori Infection Based on Endoscopic Images: Preliminary Experience. *Clin Transl Gastroenterol* 2019; **10**: e00109 [PMID: 31833862 DOI: 10.14309/ctg.0000000000000109]

14 **Luo H**, Xu G, Li C, He L, Luo L, Wang Z, Jing B, Deng Y, Jin Y, Li Y, Li B, Tan W, He C, Seeruttun SR, Wu Q, Huang J, Huang DW, Chen B, Lin SB, Chen QM, Yuan CM, Chen HX, Pu HY, Zhou F, He Y, Xu RH. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol* 2019; **20**: 1645-1654 [PMID: 31591062 DOI: 10.1016/S1470-2045(19)30637-0]

15 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929. e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

16 **Ding Z**, Shi H, Zhang H, Meng L, Fan M, Han C, Zhang K, Ming F, Xie X, Liu H, Liu J, Lin R, Hou X. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* 2019; **157**: 1044-1054. e5 [PMID: 31251929 DOI: 10.1053/j.gastro.2019.06.025]

17 **Zhang MM**, Yang H, Jin ZD, Yu JG, Cai ZY, Li ZS. Differential diagnosis of pancreatic cancer from normal tissue with digital imaging processing and pattern recognition based on a support vector machine of EUS images. *Gastrointest Endosc* 2010; **72**: 978-985 [PMID: 20855062 DOI: 10.1016/j.gie.2010.06.042]

18 **Zhu J**, Wang L, Chu Y, Hou X, Xing L, Kong F, Zhou Y, Wang Y, Jin Z, Li Z. A new descriptor for computer-aided diagnosis of EUS imaging to distinguish autoimmune pancreatitis from chronic pancreatitis. *Gastrointest Endosc* 2015; **82**: 831-836. e1 [PMID: 25952089 DOI:

10.1016/j.gie.2015.02.043]

19 **Săftoiu A**, Vilmann P, Gorunescu F, Gheonea DI, Gorunescu M, Ciurea T, Popescu GL, Iordache A, Hassan H, Iordache S. Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. *Gastrointest Endosc* 2008; **68**: 1086-1094 [PMID: 18656186 DOI: 10.1016/j.gie.2008.04.031]

20 **Chari ST**, Smyrk TC, Levy MJ, Topazian MD, Takahashi N, Zhang L, Clain JE, Pearson RK, Petersen BT, Vege SS, Farnell MB. Diagnosis of autoimmune pancreatitis: the Mayo Clinic experience. *Clin Gastroenterol Hepatol* 2006; **4**: 1010-1016 [PMID: 16843735 DOI: 10.1016/j.cgh.2006.05.017]

21 **Chari ST**, Takahashi N, Levy MJ, Smyrk TC, Clain JE, Pearson RK, Petersen BT, Topazian MA, Vege SS. A diagnostic strategy to distinguish autoimmune pancreatitis from pancreatic cancer. *Clin Gastroenterol Hepatol* 2009; **7**: 1097-1103 [PMID: 19410017 DOI: 10.1016/j.cgh.2009.04.020]

22 **SARLES H**, SARLES JC, MURATORE R, GUIEN C. Chronic inflammatory sclerosis of the pancreas--an autonomous pancreatic disease? *Am J Dig Dis* 1961; **6**: 688-698 [PMID: 13746542 DOI: 10.1007/BF02232341]

23 **Comings DE**, Skubi KB, Van Eyes J, Motulsky AG. Familial multifocal fibrosclerosis. Findings suggesting that retroperitoneal fibrosis, mediastinal fibrosis, sclerosing cholangitis, Riedel's thyroiditis, and pseudotumor of the orbit may be different manifestations of a single disease. *Ann Intern Med* 1967; **66**: 884-892 [PMID: 6025229 DOI: 10.7326/0003-4819-66-5-884]

24 **Gardner TB**, Levy MJ, Takahashi N, Smyrk TC, Chari ST. Misdiagnosis of autoimmune pancreatitis: a caution to clinicians. *Am J Gastroenterol* 2009; **104**: 1620-1623 [PMID: 19574965 DOI: 10.1038/ajg.2008.89]

25 **Marya NB**, Powers PD, Chari ST, Gleeson FC, Leggett CL, Abu Dayyeh BK, Chandrasekhara V, Iyer PG, Majumder S, Pearson RK, Petersen BT, Rajan E, Sawas T, Storm AC, Vege SS, Chen S, Long Z, Hough DM, Mara K, Levy MJ. Utilisation of artificial intelligence for the development of an EUS-convolutional neural network model trained to enhance the diagnosis of autoimmune pancreatitis. *Gut* 2021; **70**: 1335-1344 [PMID: 33028668 DOI: 10.1136/gutjnl-2020-322821]

26 **Hewitt MJ**, McPhail MJ, Possamai L, Dhar A, Vlavianos P, Monahan KJ. EUS-guided FNA for diagnosis of solid pancreatic neoplasms: a meta-analysis. *Gastrointest Endosc* 2012; **75**: 319-331 [PMID: 22248600 DOI: 10.1016/j.gie.2011.08.049]

27 **Ishikawa T**, Kawashima H, Ohno E, Suhara H, Hayashi D, Hiramatsu T, Matsubara H, Suzuki T, Kuwahara T, Ishikawa E, Shimoyama Y, Kinoshita F, Hirooka Y, Fujishiro M. Usefulness of endoscopic ultrasound-guided fine-needle biopsy for the diagnosis of autoimmune pancreatitis using a 22-gauge Franseen needle: a prospective multicenter study. *Endoscopy* 2020; **52**: 978-985 [PMID: 32583394 DOI: 10.1055/a-1183-3583]

28 **Nishimori I**, Tamakoshi A, Otsuki M; Research Committee on Intractable Diseases of the Pancreas, Ministry of Health, Labour, and Welfare of Japan. Prevalence of autoimmune pancreatitis in Japan from a nationwide survey in 2002. *J Gastroenterol* 2007; **42** Suppl 18: 6-8 [PMID: 17520216 DOI: 10.1007/s00535-007-2043-y]

29 **Gardner TB**, Adler DG, Forsmark CE, Sauer BG, Taylor JR, Whitcomb DC. ACG Clinical Guideline: Chronic Pancreatitis. *Am J Gastroenterol* 2020; **115**: 322-339 [PMID: 32022720 DOI: 10.14309/ajg.0000000000000535]

30 **Issa Y**, Kempeneers MA, van Santvoort HC, Bollen TL, Bipat S, Boermeester MA. Diagnostic performance of imaging modalities in chronic pancreatitis: a systematic review and meta-analysis. *Eur Radiol* 2017; **27**: 3820-3844 [PMID: 28130609 DOI: 10.1007/s00330-016-4720-9]

31 **Norton ID**, Zheng Y, Wiersema MS, Greenleaf J, Clain JE, Dimagno EP. Neural network analysis of EUS images to differentiate between pancreatic malignancy and pancreatitis. *Gastrointest Endosc* 2001; **54**: 625-629 [PMID: 11677484 DOI: 10.1067/mge.2001.118644]

32 **Das A**, Nguyen CC, Li F, Li B. Digital image analysis of EUS images accurately differentiates pancreatic cancer from chronic pancreatitis and normal tissue. *Gastrointest Endosc* 2008; **67**: 861-867 [PMID: 18179797 DOI: 10.1016/j.gie.2007.08.036]

33 **Zhu M**, Xu C, Yu J, Wu Y, Li C, Zhang M, Jin Z, Li Z. Differentiation of pancreatic cancer and chronic pancreatitis using computer-aided diagnosis of endoscopic ultrasound (EUS) images: a diagnostic test. *PLoS One* 2013; **8**: e63820 [PMID: 23704940 DOI: 10.1371/journal.pone.0063820]

34 **Tanaka M**, Fernández-Del Castillo C, Kamisawa T, Jang JY, Levy P, Ohtsuka T, Salvia R, Shimizu Y, Tada M, Wolfgang CL. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. *Pancreatology* 2017; **17**: 738-753 [PMID: 28735806 DOI: 10.1016/j.pan.2017.07.007]

35 **Kuwahara T**, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, Kurita Y, Koda H, Toriyama K, Onishi S, Ishihara M, Tanaka T, Tajika M, Niwa Y. Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Clin Transl Gastroenterol* 2019; **10**: 1-8 [PMID: 31117111 DOI: 10.14309/ctg.0000000000000045]

36 **Kuwahara T**, Hara K, Mizuno N, Haba S, Okuno N, Koda H, Miyano A, Fumihara D. Current status of artificial intelligence analysis for endoscopic ultrasonography. *Dig Endosc* 2021; **33**: 298-305 [PMID: 33098123 DOI: 10.1111/den.13880]

37 **Ahmad OF**, Stassen P, Webster GJ. Artificial intelligence in biliopancreatic endoscopy: Is there any role? *Best Pract Res Clin Gastroenterol* 2021; **52-53**: 101724 [PMID: 34172251 DOI: 10.1016/J.BPG.2020.101724]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com

ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 95-197

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 4 August 28, 2021

OPINION REVIEW

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 4 August 28, 2021

Contents

ABOUT COVER

Associate Editor of *Artificial Intelligence in Gastrointestinal Endoscopy*, Peter Bauerfeind, MD, Professor, Department of Gastroenterology and Hepatology, University hospital Zurich, Zurich 8091, Switzerland.
peter.bauerfeind@usz.ch

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Lin-YuTong Wang*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jin-Lei Wang*.

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Fatih Altintoprak, Sahin Coban, Krish Ragunath

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
August 28, 2021

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 95-102

DOI: 10.37126/aige.v2.i4.95

ISSN 2689-7164 (online)

OPINION REVIEW

Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease

Bobby Lo, Johan Burisch

ORCID number: Bobby Lo 0000-0002-0252-9341; Johan Burisch 0000-0002-3312-5139.

Author contributions: Lo B and Burisch J authors have made a significant contribution to the research described in this manuscript; all authors approved the final manuscript as well as the authorship list.

Conflict-of-interest statement: Lo B has received a lecture fee from Janssen-Cilag; Burisch J has received consulting fees from Celgene, Janssen-Cilag, AbbVie, Vifor Pharma, Jansen and Ferring; lecture fees from Abbvie, Pfizer, MSD, Pharmacosmos and Takeda Pharma, and unrestricted grant support from Takeda Pharma and Tillotts Pharma.

Bobby Lo, Johan Burisch, Gastrounit, Medical Section, Copenhagen University Hospital Hvidovre, Hvidovre 2650, Denmark

Bobby Lo, Johan Burisch, Copenhagen Centre for Inflammatory Bowel Disease in Children, Adolescents and Adults, Copenhagen University Hospital Hvidovre, Hvidovre 2650, Denmark

Corresponding author: Bobby Lo, MD, Doctor, Gastrounit, Medical Section, Copenhagen University Hospital Hvidovre, Kettegård Alle 30, Hvidovre 2650, Denmark.
bobby_lo@protonmail.com

Abstract

Assessment of endoscopic disease activity can be difficult in patients with inflammatory bowel disease (IBD) [comprises Crohn's disease (CD) and ulcerative colitis (UC)]. Endoscopic assessment is currently the foundation of disease evaluation and the grading is pivotal for the initiation of certain treatments. Yet, disharmony is found among experts; even when reassessed by the same expert. Some studies have demonstrated that the evaluation is no better than flipping a coin. In UC, the greatest achieved consensus between physicians when assessing endoscopic disease activity only reached a Kappa value of 0.77 (or 77% agreement adjustment for chance/accident). This is unsatisfactory when dealing with patients at risk of surgery or disease progression without proper care. Lately, across all medical specialities, computer assistance has become increasingly interesting. Especially after the emanation of machine learning – colloquially referred to as artificial intelligence (AI). Compared to other data analysis methods, the strengths of AI lie in its capability to derive complex models from a relatively small dataset and its ability to learn and optimise its predictions from new inputs. It is therefore evident that with such a model, one hopes to be able to remove inconsistency among humans and standardise the results across educational levels, nationalities and resources. This has manifested in a handful of studies where AI is mainly applied to capsule endoscopy in CD and colonoscopy in UC. However, due to its recent place in IBD, there is a great inconsistency between the results, as well as the reporting of the same. In this opinion review, we will explore and evaluate the method and results of the published studies utilising AI within IBD (with examples), and discuss the future possibilities AI can offer within IBD.

Key Words: Inflammatory bowel disease; Artificial intelligence; Deep learning; Endoscopy; Disease severity; Machine learning

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: Denmark

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): C, C, C, C
Grade D (Fair): 0
Grade E (Poor): 0

Received: April 28, 2021
Peer-review started: April 28, 2021
First decision: June 13, 2021
Revised: June 27, 2021
Accepted: August 16, 2021
Article in press: August 16, 2021
Published online: August 28, 2021

P-Reviewer: Chabrun F, El-Shabrawi MH, Kamran M, Li A
S-Editor: Liu M
L-Editor: Webster JR
P-Editor: Xing YX

Core Tip: Artificial intelligence (AI) is on the rise in inflammatory bowel diseases (IBD). Endoscopic evaluation is so far the most studied modality with promising results. Studies with others or the combination of several modalities have been carried out with moderate results leaving room for future research. Data availability and standardisation of the reporting of these new models seem to be the biggest challenges for AI's breakthrough within IBD. International consensus in the field is required to optimise research in AI.

Citation: Lo B, Burisch J. Artificial intelligence assisted assessment of endoscopic disease activity in inflammatory bowel disease. *Artif Intell Gastrointest Endosc* 2021; 2(4): 95-102
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/95.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.95

INTRODUCTION

The inflammatory bowel diseases (IBD), which mainly consist of Crohn's disease (CD) and ulcerative colitis (UC), are idiopathic immune-mediated diseases usually affecting young adults[1,2].

Currently, colonoscopy is considered the gold standard in the disease assessment of patients with UC as well as CD located in the terminal ileum and/or colon[3,4]. Disease activity of UC is assessed using scoring systems such as the Mayo Endoscopic Subscore (MES) or UC Endoscopic Index of Severity[5]. Despite their widespread use and being easy to use, both indices suffer from moderate to high inter-observer variation which reduces the credibility of the scores[6]. This has been demonstrated in clinical trials where up to one-third of patients deemed eligible for inclusion based on the MES did not live up to the inclusion criteria after reassessment[7]. Even central reading is associated with noteworthy inter-observer variation[7,8].

In CD, the CD Endoscopic Index of Severity and Simple Endoscopic Score for CD are currently the most used indices[4]. Both have demonstrated varying observer variance with central reading improving the inter-observer variation[9-11]. Capsule endoscopy (CE) for evaluating the small bowel can be scored using the Lewis score [12]. While widely used, the interobserver agreement between parameters in the index fluctuates widely (kappa 0.37-0.83)[13,14].

These interobserver variations and the risk of misclassification has led to the exploration of artificial intelligence (AI) assisted endoscopic assessment[15], especially in the field of colon cancer detection[16,17]. AI, depending on which method is used, mimics the human brain by having interconnected neurons that process the information given; however, in contrast to the human brain, AI can theoretically process an unlimited number of variables. In the field of IBD, the use of AI remains limited although it has received increasing attention. In the following review, we will discuss the use of AI-assisted assessment of endoscopic disease activity among CD and UC patients from a clinical perspective, the challenges the model faces and unexplored areas where AI has the potential to help patients and physicians.

CROHN'S DISEASE

CD can be examined using many modalities. Imaging has been an area of interest in terms of AI - especially CE[18]. A CE camera takes between 2-4 frames per second and has an approximate transit time of 250 min which can result in a total of approximately 60000 images[18]. One of the challenges CE entails is that it is a time-consuming process whereby a trained person must subsequently review all images. New AI has since assisted physicians and endoscopists in filtering out non-informative images, thereby leaving an image series where the computer believes there is an area of interest. Since the year 2000, AI has been used to identify polyps/tumours, ulcers, celiac disease, hookworms, angioectasia, and bleeding[18]. Among CD patients, special focus has been on small bowel lesions, erosions and ulceration[19]. The

majority of recent studies that have examined the listed parameters use a convolutional neural network - a deep learning method that has been shown to be effective in image recognition[18,20]. Overall, these studies have shown an accuracy of > 90% which must be considered close to perfect. However, the majority of these studies are conducted retrospectively and prospective results are wanted to demonstrate the models potential in clinical practice.

ULCERATIVE COLITIS

Due to UC only involving the colon it has been easier to categorise these patients than CD according to the extent and severity of inflammation[21]. Accordingly, most advances regarding AI in IBD has been done in UC and several clinical tools have been developed to assess the endoscopic disease severity. Such models have achieved an accuracy of 56%–77% in assessing the disease severity according to the MES or UC Endoscopic Index of Severity which was comparable to IBD experts[22-27]. The majority of studies have used methods such as the convolutional neural network to categorize images taken during a colonoscopy or sigmoidoscopy according to the MES. Recently, studies have also investigated the applicability of AI on videos; demonstrating a promising area under the receiver operating characteristic curve (AUROC or AUC) > 90%[24,26,27].

Currently, the available models are unable to distinguish between the different levels of the MES with sufficient accuracy. However, this is an area under great development and it is expected that within the coming years a model will be able to distinguish between the different MES levels with a satisfactory result and thereby eliminate the inter-observer variance, and standardize the clinical and academic evaluation of the endoscopic disease severity[28].

Few studies have further examined their model's MES score in relation to histological findings[29,30]. One study used endocytoscopy with a support vector machine and achieved an accuracy of approximately 90% in predicting histological findings which must be considered excellent results[29]. Endocytoscopy is, however, not an integral method in most clinics. Furthermore, although the study group utilized both a training and a test set, the training and optimizing process of the models is not described, leaving the reader with uncertainty with regard to *e.g.*, model selection and tuning of. Finally, samples were divided into active inflammation *vs* remission which might be too simplified a way of considering both the endoscopic and histological findings. Similar results were demonstrated by Takenaka *et al*[30] with white-light endoscopy, but with the same challenges. Ultimately, none of these studies validated the results on an independent cohort analyzed by independent experts, in order to test the performance of their model when compared to another population or to the point of view of different experts.

POTENTIAL AND DIFFICULTIES

As previously mentioned, AI has been shown to have great potential in the evaluation of endoscopic severity among patients with CD and UC. The models have shown to be at a level with or better than physicians to classify endoscopic disease severity; especially among UC patients[25]. Uniformity in the approach to the endoscopic procedure will make new clinical tools more credible and hopefully lead to less discrepancy between clinical and observational studies[31]. However, it is crucial that new models are developed for clinical purposes, which can be implemented more quickly, thereby reducing the gap between research and clinical practice.

Besides endoscopic evaluation, disease prediction in IBD has also been investigated using AI models. Waljee *et al*[32,33] used two clinical trial databases to predict C-reactive protein < 5 mg/L after 42 wk treatment with ustekinumab and steroid-free remission after 52 wk treatment with vedolizumab among CD patients, respectively. These studies used a combination of demographic, clinical, and biochemical data in a random forest model to predict patients' course after initiation of treatment. The models achieved an accuracy of 42% and 69%, respectively. Furthermore, the same study group investigated the treatment effect of vedolizumab in UC patients[34]. Using a random forest model, the model achieved an accuracy of 58% in predicting corticosteroid-free remission after 52 wk. When grouping UC and CD together, Biasci *et al*[35] used transcriptomics to identify a blood sample panel of 17 genes with sensitivity and specificity of approximately 73% to predict patients' risk of treatment

escalating within 1 year. A 5-year prediction study from Choi *et al*[36] demonstrated a sensitivity and specificity of 71% for predicting the risk of the use of biologics. In contrast to Biasci *et al*[35], this study utilized only demographical, clinical and common laboratory markers. Furthermore, Waljee *et al*[37,38] attempted twice to predict the treatment effect within 1 year, resulted in an AUC of 79% and 87% and accuracy of 72% and 80%, respectively. A limitation of these studies is that findings are only presented for IBD patients in total and not stratified according to the type of IBD. Despite these efforts, accuracies below at least 80% must be considered insufficient. Furthermore, even with accuracies above 80%, the results must be taken into perspective with the sensitivity, specificity and AUC to achieve an overall picture of the model's performance. Unfortunately, the majority of the studies have only reported some but not all measures of validity of which AUC is most commonly reported.

OTHER AREAS

It is not uncommon for some patients to undergo a lengthy diagnostic process before a definite diagnosis of CD or UC can be made[39]. This can be a challenge for both physicians and patients, and result in over or under treatments with major consequences for the patient. Recent studies using AI have attempted to use several modalities to better distinguish between these patients: endoscopy, histology, genetic markers, biochemical markers, clinical factors, omics, or a combination of one or more of these modalities[40-43]. These have shown acceptable results with AUC and accuracy of > 80%. It should be emphasized that these studies do not always report all results and many of the results are from validation data and not necessarily test data (unseen data) exposing the models to overfitting. However, to our knowledge, none of these models has been applied in clinical practice and real-life data are warranted to evaluate their efficacy.

To our knowledge, no other modalities explored in connection with AI have been published to date. In particular, the complexity of CD results in several challenges when developing new AI models. One area that remains untouched is the use of AI during colonoscopy in CD patients. This could be due to challenges in the endoscopic disease assessment of CD as the disease can be patchy and the severity varies between patches. Besides, indices for CD are difficult or time-consuming to use in clinical practice[4]. This could be accommodated by developing new scoring indices based on an evaluation from an AI model, allowing the possibility of assessing the gut as a whole rather than the segmented method currently being used.

In addition to endoscopies for both UC and CD, modalities such as ultrasound, magnetic resonance imaging, colon CE and computed tomography are obvious opportunities for the development of new clinical tools[44].

Unfortunately, this field is also challenged by several issues. First and foremost, a paradigm shift is needed; from a medical professional to a computer-aided assessment. This will first and foremost require doctors to accept the new technology [45] which can be difficult to understand as the latest AI architectures use deep learning where a black-box appears (the process between input and output)[46]. As it is not 100% possible to account for what happens in this black-box, mistrust might arise among the clinicians toward the models. Despite different ways of explaining the black-box, mathematically and illustratively, it is only possible to give an estimate of its process[46].

Secondly, medical education may need to be reorganized in the future to have more focus on interpretation and critical evaluation of the results of these new models. The medical field has experienced a similar paradigm shift before with the introduction of the World Wide Web[47]. This gave patients equal access to knowledge that physicians had and doctors went from being the ultimate definitive truth to now having to explain how the symptoms and the disease are connected and which diagnosis and disease courses are most likely[47]. However, a new organization of the medical education in connection with AI may require interdisciplinary involvement with, among others, bioinformatics and computer scientists to better equip doctors to interpret and critically evaluate the models' output.

Thirdly, larger amounts of data are needed – more than previously accustomed to developing these new models. However, the amount of data needed varies significantly in relation to the outcome and the methods used and no specific number of required data exists. As data is resource demanding, the estimate must be adjusted to what is clinically possible. In recent years, cross-border collaborations have been

Table 1 Recommendations for reporting of studies regarding artificial intelligence

Section	Requirements
Method	
	Origin of dataset and description of the acquisition process
	Pre-processing methods
	Definition of ground truth
	Split of data set and should include a training, validation and test set. A clear statement that the test set is not used to tune hyperparameters or in the selection of the model
	Method and architecture used, whether it is pretrained or not, and what dataset it is pretrained on
	Full technical detail should be included in supplementary files
	Statement of post-selection analyses and why these are conducted
Results	
	A complete report of all results including but not restricted to AUC, sensitivity, specificity, accuracy and kappa value for the overall model's performance and not for selected tasks
Discussion	
	Risks of overfitting and bias
	Generalisability and cautions to take
	Clinical implementation

AUC: Area under the receiver operating characteristic curve.

formed to make large amounts of data available. However, these are rarely freely available and the quality must also be critically evaluated when the workflow and equipment vary markedly between nations. We, therefore, encourage everyone to make their data at least partially accessible - a good example is The HyperKvasir dataset[48].

Finally, international reporting standards must be set within the field of IBD regarding AI studies. AI is still a relatively unexploited territory within IBD. This has led to great variation in the way the studies report both their methods and results, despite several calls for uniformity[49]. A good example is the endoscopic evaluation of disease severity in UC patients. Often, only AUC is reported, which can be misleading as sensitivity, specificity and accuracy may be only modest[25]. This is due to the fact that the AUC is a measure of how well the true positive can be separated from the rest, while measures of *e.g.*, accuracy hint at the actual performance of the models. Even when the studies report the wanted parameters, the reporting method can vary. For example, calculating the sensitivity, specificity and accuracy for each class rather than reporting the overall sensitivity, specificity and accuracy for the entire index. We, therefore, encourage that future articles as a minimum must report the information and parameters described in Table 1.

In addition, international journals should set standards for what is required of future AI studies within the field. The use of previous reporting methods, *e.g.*, STARD guidelines, seems outdated and should be updated to the new technological reality [50].

CONCLUSION

AI is on the rise in IBD. Endoscopic evaluation is so far the most studied modality with promising results. Studies with others or the combination of several modalities have been carried out with moderate results leaving room for future research. Data availability and standardization of the reporting of these new models seem to be the biggest challenges for the AI's breakthrough within IBD. International consensus in the field is required to optimize research in AI.

REFERENCES

1 **Torres J**, Mehandru S, Colombel JF, Peyrin-Biroulet L. Crohn's disease. *Lancet* 2017; **389**: 1741-1755 [PMID: 27914655 DOI: 10.1016/S0140-6736(16)31711-1]

2 **Ungaro R**, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel JF. Ulcerative colitis. *Lancet* 2017; **389**: 1756-1770 [PMID: 27914657 DOI: 10.1016/S0140-6736(16)32126-2]

3 **Magro F**, Gionchetti P, Eliakim R, Ardizzone S, Armuzzi A, Barreiro-de Acosta M, Burisch J, Gecse KB, Hart AL, Hindryckx P, Langner C, Limdi JK, Pellino G, Zagórowicz E, Raine T, Harbord M, Rieder F; European Crohn's and Colitis Organisation [ECCO]. Third European Evidence-based Consensus on Diagnosis and Management of Ulcerative Colitis. Part 1: Definitions, Diagnosis, Extra-intestinal Manifestations, Pregnancy, Cancer Surveillance, Surgery, and Ileo-anal Pouch Disorders. *J Crohns Colitis* 2017; **11**: 649-670 [PMID: 28158501 DOI: 10.1093/ecco-jcc/jjx008]

4 **Peyrin-Biroulet L**, Sandborn W, Sands BE, Reinisch W, Bemelman W, Bryant RV, D'Haens G, Dotan I, Dubinsky M, Feagan B, Fiorino G, Gearry R, Krishnareddy S, Lakatos PL, Loftus EV Jr, Marteau P, Munkholm P, Murdoch TB, Ordás I, Panaccione R, Riddell RH, Ruel J, Rubin DT, Samaan M, Siegel CA, Silverberg MS, Stoker J, Schreiber S, Travis S, Van Assche G, Danese S, Panes J, Bouguen G, O'Donnell S, Pariente B, Winer S, Hanauer S, Colombel JF. Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): Determining Therapeutic Goals for Treat-to-Target. *Am J Gastroenterol* 2015; **110**: 1324-1338 [PMID: 26303131 DOI: 10.1038/ajg.2015.233]

5 **Peyrin-Biroulet L**, Panés J, Sandborn WJ, Vermeire S, Danese S, Feagan BG, Colombel JF, Hanauer SB, Rycroft B. Defining Disease Severity in Inflammatory Bowel Diseases: Current and Future Directions. *Clin Gastroenterol Hepatol* 2016; **14**: 348-354.e17 [PMID: 26071941 DOI: 10.1016/j.cgh.2015.06.001]

6 **Mohammed Vashist N**, Samaan M, Mosli MH, Parker CE, MacDonald JK, Nelson SA, Zou GY, Feagan BG, Khanna R, Jairath V. Endoscopic scoring indices for evaluation of disease activity in ulcerative colitis. *Cochrane Database Syst Rev* 2018; **1**: CD011450 [PMID: 29338066 DOI: 10.1002/14651858.CD011450.pub2]

7 **Samaan MA**, Mosli MH, Sandborn WJ, Feagan BG, D'Haens GR, Dubcenco E, Baker KA, Levesque BG. A systematic review of the measurement of endoscopic healing in ulcerative colitis clinical trials: recommendations and implications for future research. *Inflamm Bowel Dis* 2014; **20**: 1465-1471 [PMID: 24831558 DOI: 10.1097/MIB.0000000000000046]

8 **Feagan BG**, Sandborn WJ, D'Haens G, Pola S, McDonald JWD, Rutgeerts P, Munkholm P, Mittmann U, King D, Wong CJ, Zou G, Donner A, Shackelton LM, Gilgen D, Nelson S, Vandervoort MK, Fahmy M, Loftus EV Jr, Panaccione R, Travis SP, Van Assche GA, Vermeire S, Levesque BG. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013; **145**: 149-157.e2 [PMID: 23528626 DOI: 10.1053/j.gastro.2013.03.025]

9 **Daperno M**, Comberlato M, Bossa F, Biancone L, Bonanomi AG, Cassinotti A, Cosintino R, Lombardi G, Mangiarotti R, Papa A, Pica R, Rizzello F, D'Incà R, Orlando A. Inter-observer agreement in endoscopic scoring systems: preliminary report of an ongoing study from the Italian Group for Inflammatory Bowel Disease (IG-IBD). *Dig Liver Dis* 2014; **46**: 969-973 [PMID: 25154049 DOI: 10.1016/j.dld.2014.07.010]

10 **Daperno M**, Comberlato M, Bossa F, Armuzzi A, Biancone L, Bonanomi AG, Cassinotti A, Cosintino R, Lombardi G, Mangiarotti R, Papa A, Pica R, Grassano L, Pagana G, D'Incà R, Orlando A, Rizzello F; IGIBDEndo Group. Training Programs on Endoscopic Scoring Systems for Inflammatory Bowel Disease Lead to a Significant Increase in Interobserver Agreement Among Community Gastroenterologists. *J Crohns Colitis* 2017; **11**: 556-561 [PMID: 28453758 DOI: 10.1093/ecco-jcc/jjw181]

11 **Rutgeerts P**, Reinisch W, Colombel JF, Sandborn WJ, D'Haens G, Petersson J, Zhou Q, Iezzi A, Thakkar RB. Agreement of site and central readings of ileocolonoscopic scores in Crohn's disease: comparison using data from the EXTEND trial. *Gastrointest Endosc* 2016; **83**: 188-197 [PMID: 26234693 DOI: 10.1016/j.gie.2015.06.018]

12 **Gralnek IM**, Defranchis R, Seidman E, Leighton JA, Legnani P, Lewis BS. Development of a capsule endoscopy scoring index for small bowel mucosal inflammatory change. *Aliment Pharmacol Ther* 2008; **27**: 146-154 [PMID: 17956598 DOI: 10.1111/j.1365-2036.2007.03556.x]

13 **Cotter J**, Dias de Castro F, Magalhães J, Moreira MJ, Rosa B. Validation of the Lewis score for the evaluation of small-bowel Crohn's disease activity. *Endoscopy* 2015; **47**: 330-335 [PMID: 25412092 DOI: 10.1055/s-0034-1390894]

14 **Esaki M**, Washio E, Morishita T, Sakamoto K, Fuyuno Y, Hirano A, Umeno J, Kitazono T, Matsumoto T, Suzuki Y. Su1262 inter- and intra-observer variation of capsule endoscopic findings for the diagnosis of crohn's disease: A case-control study. *Gastrointest Endosc* 2018; **87**: AB302 [DOI: 10.1016/j.gie.2018.04.1649]

15 **Buch VH**, Ahmed I, Maruthappu M. Artificial intelligence in medicine: current trends and future possibilities. *Br J Gen Pract* 2018; **68**: 143-144 [PMID: 29472224 DOI: 10.3399/bjgp18X695213]

16 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

17 **Wang Y**, He X, Nie H, Zhou J, Cao P, Ou C. Application of artificial intelligence to the diagnosis and therapy of colorectal cancer. *Am J Cancer Res* 2020; **10**: 3575-3598 [PMID: 33294256]

18 **Yang YJ**. The Future of Capsule Endoscopy: The Role of Artificial Intelligence and Other Technical Advancements. *Clin Endosc* 2020; **53**: 387-394 [PMID: 32668529 DOI: 10.5946/ce.2020.133]

19 **Klang E**, Barash Y, Margalit RY, Soffer S, Shimon O, Albshesh A, Ben-Horin S, Amitai MM, Eliakim R, Kopylov U. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc* 2020; **91**: 606-613 [PMID: 31743689 DOI: 10.1016/j.gie.2019.11.012]

20 **Soffer S**, Klang E, Shimon O, Nachmias N, Eliakim R, Ben-Horin S, Kopylov U, Barash Y. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. *Gastrointest Endosc* 2020; **92**: 831-839.e8 [PMID: 32334015 DOI: 10.1016/j.gie.2020.04.039]

21 **Satsangi J**, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 2006; **55**: 749-753 [PMID: 16698746 DOI: 10.1136/gut.2005.082909]

22 **Stidham RW**, Liu W, Bishu S, Rice MD, Higgins PDR, Zhu J, Nallamothu BK, Waljee AK. Performance of a Deep Learning Model vs Human Reviewers in Grading Endoscopic Disease Severity of Patients With Ulcerative Colitis. *JAMA Netw Open* 2019; **2**: e193963 [PMID: 31099869 DOI: 10.1001/jamanetworkopen.2019.3963]

23 **Ozawa T**, Ishihara S, Fujishiro M, Saito H, Kumagai Y, Shichijo S, Aoyama K, Tada T. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc* 2019; **89**: 416-421.e1 [PMID: 30367878 DOI: 10.1016/j.gie.2018.10.020]

24 **Gottlieb K**, Requa J, Karnes W, Chandra Gudivada R, Shen J, Rael E, Arora V, Dao T, Ninh A, McGill J. Central Reading of Ulcerative Colitis Clinical Trial Videos Using Neural Networks. *Gastroenterology* 2021; **160**: 710-719 [PMID: 33098883 DOI: 10.1053/j.gastro.2020.10.024]

25 **Gutierrez Becker B**, Arcadu F, Thalhammer A, Gamez Serna C, Feehan O, Drawnel F, Oh YS, Prunotto M. Training and deploying a deep learning model for endoscopic severity grading in ulcerative colitis using multicenter clinical trial data. *Ther Adv Gastrointest Endosc* 2021; **14**: 2631774521990623 [PMID: 33718871 DOI: 10.1177/2631774521990623]

26 **Yao H**, Najarian K, Gryak J, Bishu S, Rice MD, Waljee AK, Wilkins HJ, Stidham RW. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2021; **93**: 728-736 [PMID: 32810479 DOI: 10.1016/j.gie.2020.08.011]

27 **Alammari A**, Islam ABMR, Oh JH, Tavanapong W, Wong J, De Groen PC. Classification of ulcerative colitis severity in colonoscopy videos using CNN. ICIME 2017: Proceedings of the 9th International Conference on Information Management and Engineering; 2019 Oct 9-11; Barcelona, Spain. Barcelona Spain, 2017: 139-44 [DOI: 10.1145/3149572.3149613]

28 **Lo B**, Liu Z, Bendtsen F, Igel C, Vind I, Burisch J. Deep Learning Surpasses Gastrointestinal Experts at Classifying Endoscopic Severity in Patients with Ulcerative Colitis. 2021 Preprint. Available from: SSRN Electron Journal [DOI: 10.2139/ssrn.3824683]

29 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

30 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: 32060000 DOI: 10.1053/j.gastro.2020.02.012]

31 **Salleron J**, Danese S, D'Agay L, Peyrin-Biroulet L. Effectiveness Research in Inflammatory Bowel Disease: A Necessity and a Methodological Challenge. *J Crohns Colitis* 2016; **10**: 1096-1102 [PMID: 26944416 DOI: 10.1093/ecco-jcc/jjw068]

32 **Waljee AK**, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, Stidham RW, Zhu J, Higgins PDR. Development and Validation of Machine Learning Models in Prediction of Remission in Patients With Moderate to Severe Crohn Disease. *JAMA Netw Open* 2019; **2**: e193721 [PMID: 31074823 DOI: 10.1001/jamanetworkopen.2019.3721]

33 **Waljee AK**, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, Higgins PDR. Predicting Corticosteroid-Free Biologic Remission with Vedolizumab in Crohn's Disease. *Inflamm Bowel Dis* 2018; **24**: 1185-1192 [PMID: 29668915 DOI: 10.1093/ibd/izy031]

34 **Waljee AK**, Liu B, Sauder K, Zhu J, Govani SM, Stidham RW, Higgins PDR. Predicting corticosteroid-free endoscopic remission with vedolizumab in ulcerative colitis. *Aliment Pharmacol Ther* 2018; **47**: 763-772 [PMID: 29359519 DOI: 10.1111/apt.14510]

35 **Biasci D**, Lee JC, Noor NM, Pombal DR, Hou M, Lewis N, Ahmad T, Hart A, Parkes M, McKinney EF, Lyons PA, Smith KGC. A blood-based prognostic biomarker in IBD. *Gut* 2019; **68**: 1386-1395 [PMID: 31030191 DOI: 10.1136/gutjnl-2019-318343]

36 **Choi YI**, Park SJ, Chung JW, Kim KO, Cho JH, Kim YJ, Lee KY, Kim KG, Park DK. Development of Machine Learning Model to Predict the 5-Year Risk of Starting Biologic Agents in Patients with Inflammatory Bowel Disease (IBD): K-CDM Network Study. *J Clin Med* 2020; **9**: 3427 [PMID: 33114505 DOI: 10.3390/jcm9113427]

37 **Waljee AK**, Lipson R, Wiitala WL, Zhang Y, Liu B, Zhu J, Wallace B, Govani SM, Stidham RW, Hayward R, Higgins PDR. Predicting Hospitalization and Outpatient Corticosteroid Use in

Inflammatory Bowel Disease Patients Using Machine Learning. *Inflamm Bowel Dis* 2017; **24**: 45-53 [PMID: 29272474 DOI: 10.1093/ibd/izx007]

38 **Waljee AK**, Joyce JC, Wang S, Saxena A, Hart M, Zhu J, Higgins PD. Algorithms outperform metabolite tests in predicting response of patients with inflammatory bowel disease to thiopurines. *Clin Gastroenterol Hepatol* 2010; **8**: 143-150 [PMID: 19835986 DOI: 10.1016/j.cgh.2009.09.031]

39 **Vind I**, Riis L, Jess T, Knudsen E, Pedersen N, Elkjaer M, Bak Andersen I, Wewer V, Nørregaard P, Moesgaard F, Bendtsen F, Munkholm P; DCCD study group. Increasing incidences of inflammatory bowel disease and decreasing surgery rates in Copenhagen City and County, 2003-2005: a population-based study from the Danish Crohn colitis database. *Am J Gastroenterol* 2006; **101**: 1274-1282 [PMID: 16771949 DOI: 10.1111/j.1572-0241.2006.00552.x]

40 **Li J**, Qian JM. Artificial intelligence in inflammatory bowel disease: current status and opportunities. *Chin Med J (Engl)* 2020; **133**: 757-759 [PMID: 32132365 DOI: 10.1097/CM9.0000000000000714]

41 **Mossotto E**, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep* 2017; **7**: 2427 [PMID: 28546534 DOI: 10.1038/s41598-017-02606-2]

42 **Douglas GM**, Hansen R, Jones CMA, Dunn KA, Comeau AM, Bielawski JP, Tayler R, El-Omar EM, Russell RK, Hold GL, Langille MGI, Van Limbergen J. Multi-omics differentially classify disease state and treatment outcome in pediatric Crohn's disease. *Microbiome* 2018; **6**: 13 [PMID: 29335008 DOI: 10.1186/s40168-018-0398-3]

43 **Duttagupta R**, DiRienzo S, Jiang R, Bowers J, Gollub J, Kao J, Kearney K, Rudolph D, Dawany NB, Showe MK, Stamato T, Getts RC, Jones KW. Genome-wide maps of circulating miRNA biomarkers for ulcerative colitis. *PLoS One* 2012; **7**: e31241 [PMID: 22359580 DOI: 10.1371/journal.pone.0031241]

44 **Sandberg K**, Yarger E, Saeed S. Updates in diagnosis and management of inflammatory bowel disease. *Curr Probl Pediatr Adolesc Health Care* 2020; **50**: 100785 [PMID: 32402535 DOI: 10.1016/j.cppeds.2020.100785]

45 **Oppenheim M**. Stephen Hawking: Artificial intelligence could be the greatest disaster in human history. [cited 2021 Apr 11]. Available from: https://www.independent.co.uk/news/people/stephen-hawking-artificial-intelligence-diaster-human-history-leverhulme-centre-cambridge-a7371106.html

46 **Buhrmester V**, Münch D, Arens M. Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey. 2019 Preprint. Available from: arXiv:1911.12116

47 **Hatcher M**, Heetebry I. Information technology in the future of health care. *J Med Syst* 2004; **28**: 673-688 [PMID: 15615295 DOI: 10.1023/b:joms.0000044969.66510.d5]

48 **Borgli H**, Thambawita V, Smedsrud PH, Hicks S, Jha D, Eskeland SL, Randel KR, Pogorelov K, Lux M, Nguyen DTD, Johansen D, Griwodz C, Stensland HK, Garcia-Ceja E, Schmidt PT, Hammer HL, Riegler MA, Halvorsen P, de Lange T. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Sci Data* 2020; **7**: 283 [PMID: 32859981 DOI: 10.1038/s41597-020-00622-y]

49 **van der Sommen F**, de Groof J, Struyvenberg M, van der Putten J, Boers T, Fockens K, Schoon EJ, Curvers W, de With P, Mori Y, Byrne M, Bergman JJGHM. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. *Gut* 2020; **69**: 2035-2045 [PMID: 32393540 DOI: 10.1136/gutjnl-2019-320466]

50 **Cohen JF**, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HC, Bossuyt PM. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; **6**: e012799 [PMID: 28137831 DOI: 10.1136/bmjopen-2016-012799]

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 103-109

DOI: 10.37126/aige.v2.i4.103

ISSN 2689-7164 (online)

OPINION REVIEW

Robotic pancreaticoduodenectomy: Where do we stand?

Hussein H Khachfe, Joseph R Habib, Mohamad A Chahrour, Ibrahim Nassour

ORCID number: Hussein H Khachfe 0000-0001-7537-9033; Joseph R Habib 0000-0003-3445-2278; Mohamad A Chahrour 0000-0002-8088-9801; Ibrahim Nassour 0000-0002-9845-1074.

Author contributions: Khachfe HH conceived the idea for the manuscript; Khachfe HH, Habib JR, Chahrour MA, and Nassour I reviewed the literature and drafted the manuscript; Nassour I critically reviewed the manuscript.

Conflict-of-interest statement: The authors report no conflict of interest.

Manuscript source: Invited manuscript

Specialty type: Surgery

Hussein H Khachfe, Ibrahim Nassour, Surgery Department, University of Pittsburgh Medical Center, Pittsburgh, PA 15261, United States

Joseph R Habib, Surgery Department, Johns Hopkins University, Balitmore, MD 21287, United States

Mohamad A Chahrour, Surgery Department, Henry Ford Health System, Detroit, MI 48202, United States

Corresponding author: Hussein H Khachfe, MD, Doctor, Research Fellow, Surgery Department, University of Pittsburgh Medical Center, 3550 Terrace St, Pittsburgh, PA 15261, United States. khachfehh@upmc.edu

Abstract

Pancreaticoduodenectomy (PD) is a complex operation accompanied by significant morbidity rates. Due to this complexity, the transition to minimally invasive PD has lagged behind other abdominal surgical operations. The safety, feasibility, favorable post-operative outcomes of robotic PD have been suggested by multiple studies. Compared to open surgery and other minimally invasive techniques such as laparoscopy, robotic PD offers satisfactory outcomes, with a non-inferior risk of adverse events. Trends of robotic PD have been on rise with centers substantially increasing the number the operation performed. Although promising, findings on robotic PD need to be corroborated in prospective trials.

Key Words: Pancreaticoduodenectomy; Whipple Procedure; Pancreas; Robotic; Surgery

Core Tip: The robotic Whipple procedure is a safe and technically feasible surgical operation. Robotic pancreaticoduodenectomy has shown favorable outcomes and is currently increasing in widespread implementation. Prospective trials are needed before this relatively new approach can be fully adopted as a standard of care in patients with pancreatic neoplasms.

Citation: Khachfe HH, Habib JR, Chahrour MA, Nassour I. Robotic pancreaticoduodenectomy: Where do we stand? *Artif Intell Gastrointest Endosc* 2021; 2(4): 103-109

URL: https://www.wjgnet.com/2689-7164/full/v2/i4/103.htm

DOI: https://dx.doi.org/10.37126/aige.v2.i4.103

Country/Territory of origin: United States

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): 0
Grade D (Fair): D
Grade E (Poor): 0

Received: May 18, 2021
Peer-review started: May 18, 2021
First decision: June 22, 2021
Revised: June 24, 2021
Accepted: August 19, 2021
Article in press: August 19, 2021
Published online: August 28, 2021

P-Reviewer: Cioffi U, Sakamoto Y
S-Editor: Liu M
L-Editor: A
P-Editor: Wang LYT

INTRODUCTION

Pancreaticoduodenectomy (PD) or Whipple surgery, is a complex procedure associated with significant morbidity rates[1]. Due to the complexity of this operation, PD's move to a more minimally invasive approach has lagged behind other general surgery procedures[2]. Gagner and Pomp[3], pioneered the laparoscopic PD (LPD) back in 1994, but LPD has not successfully transitioned into routine surgical care[3]. This is partly due to the difficulty associated with LPD in terms of expertise needed to perform the operation and the complexity of teaching the approach. In addition, the LEOPARD-2 trial demonstrated that LPD has a higher 90-d mortality as compared to the open PD (OPD). This eventually led to the discontinuation of the trial[4].

Robotic PD (RPD), which was first performed by Giulianotti *et al*[5], was originally described in 2001. Later in 2003, the same team published a series of 8 robotic-assisted cases[6]. The preliminary results established that RPD is both safe and feasible. Their reported mean operative time was around 8 h (490 min) in this case series.

Following these promising results, an increasing number of surgeons started utilizing the RPD approach. Different than initial reports of LPD, where some showed that LPD does not provide benefit as compared to the open approach, RPD benefits and advantages have been reported with increasing rate since its launch[7,8]. However, the "Miami International Guideline on Minimally Invasive Pancreas Resection" still does not assume minimally invasive PD is equal to OPD due to insufficient data[9].

WHAT IS THE ROBOTIC SURGICAL TECHNIQUE AND ITS CHALLENGES?

Robotic surgery is considered a direct advancement of laparoscopy. The most widely utilized surgical system to perform RPD in specific, as well as in other operations, is the DaVinci system developed by Intuitive Surgical Incorporated[10]. The robotic system provides surgeons increased dexterity employing endo-wristed instruction, three-dimensional stereoscopic views of the surgical field, filtering of user tremors, and it provides pancreatic surgeons the capability to perform extremely complex dissections, sutures, knots and reconstructions with unparalleled precision, magnification and accuracy[11,12].

Variations in robotic Whipple techniques exist between pancreatic surgeons. While some groups undergo the operation completely robotically, other choose to use a cross laparoscopic/robotic approach. Giulianotti *et al*[5] support a performing the operation entirely using the robotic approach, while other groups advocate the "hybrid" approach. The hybrid or cross method entails dissecting first using laparoscopy and then performing the reconstruction part using the robot[13,14]. At the University of Pittsburgh Medical Center, the surgeons employ a robotic exclusive approach, using four robotic ports, two assistant and one retractor port as shown in Figure 1. RPD follows the same steps as Whipple's 1935 description[15]. The gastrocolic ligament is first dissected to gain access to the lesser sac. Then, the ascending and transverse colon are mobilized. This is followed by a complete Kocher maneuver. Transection of the jejunum and the stomach (in classic Whipple) are then performed using stapling devices. Then, the porta is approached to transect the gastroduodenal artery and the hepatic duct. This is followed by transection of the pancreas at the neck and finally dissecting the uncinate of the mesenteric vessels. The reconstruction phase includes the creation of a pancreaticojejunostomy, followed by hepaticojejunostomy and finally a gastrojejunostomy. Finally, a drain is left behind and the port and extraction sites are closed.

The challenges facing the introduction of RPD are numerous. First, robotic operations are known to still have long operating time as compared to open ones. Second, due to the complexity of the robotic approach, there an increased need of training (higher learning curve) than the open and other minimally invasive techniques (laparoscopic). Third, robotic surgeries carry a high financial burden to patients, covering bodies and hospitals. This helps favor the open or laparoscopic approach for PD by insuring bodies and patients paying out-of-pocket. Fourth, RPDs require high-end infrastructure, which includes larger operating rooms, more technical staff present (in case any issues arise), and robotic certification by faculty and trainees. Finally, there is an increased difficulty in making prospective randomized trials in robotic operations. This issue arises with the decreased apparel/enrollment into robotic trials due to patient preference of open or laparoscopic approaches.

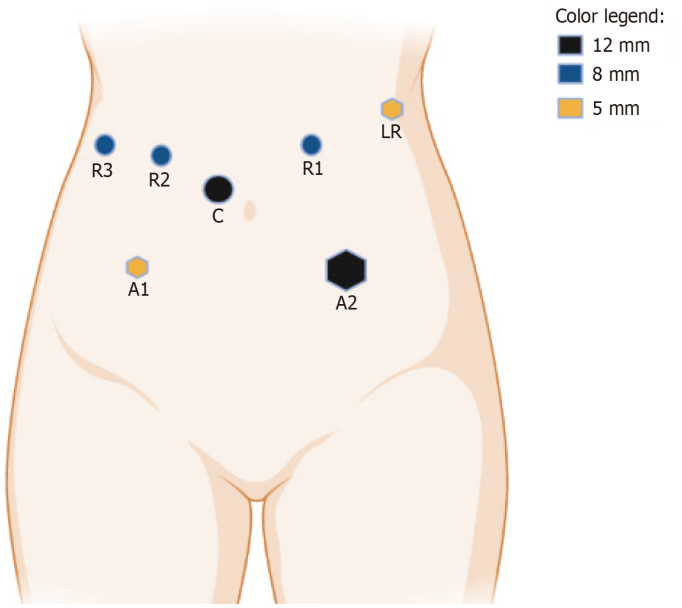

Figure 1 Port placement for robotic pancreaticoduodenectomy. R1: Robotic arm 1; R2: Robotic arm 2; R3: Robotic arm 3; C: Camera; A1: Assistant arm 1; A2: Assistant arm 2. Camera may be inserted through an 8 mm port in the Xi System. It may be inserted through a 12 mm port in the Si System.

WHAT ARE THE TRENDS AND OUTCOMES OF THE ROBOTIC WHIPPLE PROCEDURE?

A recent study exploring the trends of the RPD for pancreatic cancers demonstrated an increasing number of RPDs over the past decade. This was accompanied by a greater reach of RPD where it may be found in community centers across the US, after being present only in a few number of academic medical facilities[16]. Robotic procedures increased from 150 operations/year to around 450 operation/year from 2010-2016[16]. This is likely owing to an increase in the number of graduates from fellowship programs that include robotic pancreas surgery as part of their curriculum, as well as greater experience and "retraining" of experienced pancreatic surgeons in the robotic approach[17-20].

Overall, the robotic method appears to enhance short-term outcomes over time. Between 2010 and 2016, there was a substantial rise in the number of lymph nodes harvested (from 18 to 21), as well as a drop in postoperative mortality (from 6.7 percent to 1.8 percent)[16]. Yan *et al*[21] found that as compared to open PD, RPD had considerably longer operating time, less blood loss, shorter length of stay, and reduced infection rates in a recent meta-analysis comprising 2403 patients (788 robotic and 1615 open). There was no discernible change in lymph node harvesting, reoperation, readmission rate, or death rate[21]. Another meta-analysis by Kamarajah *et al*[22] found that RPD had substantially lower conversion and transfusion rates than LPD, with 3462 participants (1025 robotic and 2437 Laparoscopic]. RPD had a substantially shorter hospital stay after surgery, but there was no significant difference in postoperative outcomes or R0 resection rates. Zureikat *et al*[23] demonstrated that RPD was linked with decreased operating time, perioperative blood loss, and postoperative pancreatic fistula development in the largest series of RPD comprising 500 robot-assisted PD. These findings were described early in the group's experience and remained low despite growing complexity of cases. Less frequent conversion to open was also noted. As for long term outcomes, Nassour *et al*[24] identified 17831 PD from the National Cancer Database, of which 626 were RPDs. The median overall survival did not differ between the robotic (22 mo) and open (21.8 mo) approaches. Table 1 highlights RPD findings from a variety of research. In the hands of skilled surgeons, RPD is a relatively safe procedure with excellent perioperative and postoperative results.

Table 1 Outcomes of robotic pancreaticoduodenectomy in selected studies

Ref.	*n*	OR time (mean in min)	EBL (mean in mL)	Conversion (%)	R0 (%)	LN harvest (mean)	Fistula (%)	Morbidity (%)	Mortality (%)	LOS (mean in days)
Giulianotti *et al* [28], 2010	60	421	394	18.3	82	18	31.6	NR	3.3	22
Narula *et al*[29], 2010	5	420	NR	37.5	100	16	0	0	0	9.6
Zhou *et al*[30], 2011	8	718	153	0	100	NR	25	NR	0	16.4
Lai *et al*[31], 2012	20	491.5	247	5	73.3	10	35	50	0	13.7
Chalikonda *et al* [32], 2012	30	476	485	10	100	13.2	6.6	30	3	9.8
Bao *et al*[33], 2014	28	431	100	14	88	15	29	NR	2	7.4
Boone *et al*[34], 2015	200	483	250	6.5	92	22	17	67.5	3.3	9
Chen *et al*[26], 2015	60	410	400	1.7	97.8	13.6	13.3	35	1.7	20
Boggi *et al*[35], 2016	83	527	NR	1.5	NR	37	33.7	73.5	3	17
Nassour *et al*[36], 2017	193	399	NR	11.4	NR	NR	20.8	54.9	1	8
Jin *et al*[37], 2020	17	240	100	0	NR	4	59	66.4	NR	15
Mejia *et al*[38], 2020	102	352	321	12.7	73	24.2	3.9	31.3	2.9	7
Shi *et al*[39], 2020	187	279	297	3.7	94	16.6	10.2	35.6	2.1	22.4
Zureikat *et al*[23], 2021	500	415	250	5.2	85	28	20.2	68.8	1.8	8

EBL: Estimated blood loss; LN: Lymph node; LOS: Length of stay; NR: Not reported; OR: Operation; R0: Margin negative resection.

WHAT IS THE LEARNING CURVE AND FUTURE OF ROBOTIC WHIPPLE PROCEDURE?

The reported learning curves for RPD are currently variable among different institutions. The University of Pittsburgh Medical center reported that 80 RPDs would be required to optimize operative time, 40 cases for an optimal pancreatic fistula rate and 20 cases to improved blood loss and conversion[25]. This was due to the that fact the surgeons at the center had no prior training, mentorship, or guidance in the technique as the robotics program was implemented in 2008. According to Chen *et al* [26], a comparable result can be reached after 40 RPDs. At 40 patients, Zhang *et al*[27] found a comparable learning curve for RPD. The learning curve may be short if adequate training and guidance is performed in surgical formative years. A formal mastery-based curriculum which integrates complex robotic procedures into practice may help in shortening the learning curve.

The future directions of RPD will likely involve the use of robotics in borderline resectable or locally advanced pancreatic lesion cases i.e. more surgically complex cases. This also includes performing complex vasculature reconstructions using the robotic approach. However, in order to develop these surgical techniques, better infrastructure, increased training, and more prospective randomized clinical trials are required. The first step needed is to prove that RPD is noninferior to the open technique in PD with level 1 evidence. This entails increasing the number of prospective trials in order to perform meta-analyses and systematic reviews. Afterwards, increased funding and training can follow, which will allow for further developments of the RPD technique discussed. Additionally, robotic training will need to be introduced and integrated early into residency programs (possibly using simulation labs) to help with the learning curve of future robotic surgeons.

CONCLUSION

Current evidence indicates that RPD is a safe and feasible procedure. The robotic approach overcomes many of technical challenges associated with the laparoscopic Whipple procedure. RPD, in the proper hands, can help patients and surgeons with periampullary lesions achieve good results. More prospective clinical trials are still needed to verify previously published retrospective research on RPD.

REFERENCES

1 **Cameron JL**, He J. Two thousand consecutive pancreaticoduodenectomies. *J Am Coll Surg* 2015; **220**: 530-536 [PMID: 25724606 DOI: 10.1016/j.jamcollsurg.2014.12.031]

2 **van Hilst J**, de Rooij T, Klompmaker S, Rawashdeh M, Aleotti F, Al-Sarireh B, Alseidi A, Ateeb Z, Balzano G, Berrevoet F, Björnsson B, Boggi U, Busch OR, Butturini G, Casadei R, Del Chiaro M, Chikhladze S, Cipriani F, van Dam R, Damoli I, van Dieren S, Dokmak S, Edwin B, van Eijck C, Fabre JM, Falconi M, Farges O, Fernández-Cruz L, Forgione A, Frigerio I, Fuks D, Gavazzi F, Gayet B, Giardino A, Groot Koerkamp B, Hackert T, Hassenpflug M, Kabir I, Keck T, Khatkov I, Kusar M, Lombardo C, Marchegiani G, Marshall R, Menon KV, Montorsi M, Orville M, de Pastena M, Pietrabissa A, Poves I, Primrose J, Pugliese R, Ricci C, Roberts K, Røsok B, Sahakyan MA, Sánchez-Cabús S, Sandström P, Scovel L, Solaini L, Soonawalla Z, Souche FR, Sutcliffe RP, Tiberio GA, Tomazic A, Troisi R, Wellner U, White S, Wittel UA, Zerbi A, Bassi C, Besselink MG, Abu Hilal M; European Consortium on Minimally Invasive Pancreatic Surgery (E-MIPS). Minimally Invasive *vs* Open Distal Pancreatectomy for Ductal Adenocarcinoma (DIPLOMA): A Pan-European Propensity Score Matched Study. *Ann Surg* 2019; **269**: 10-17 [PMID: 29099399 DOI: 10.1097/SLA.0000000000002561]

3 **Gagner M**, Pomp A. Laparoscopic pylorus-preserving pancreatoduodenectomy. *Surg Endosc* 1994; **8**: 408-410 [PMID: 7915434 DOI: 10.1007/BF00642443]

4 **van Hilst J**, de Rooij T, Bosscha K, Brinkman DJ, van Dieren S, Dijkgraaf MG, Gerhards MF, de Hingh IH, Karsten TM, Lips DJ, Luyer MD, Busch OR, Festen S, Besselink MG; Dutch Pancreatic Cancer Group. Laparoscopic *vs* open pancreatoduodenectomy for pancreatic or periampullary tumours (LEOPARD-2): a multicentre, patient-blinded, randomised controlled phase 2/3 trial. *Lancet Gastroenterol Hepatol* 2019; **4**: 199-207 [PMID: 30685489 DOI: 10.1016/S2468-1253(19)30004-4]

5 **Giulianotti PC**, Mangano A, Bustos RE, Fernandes E, Masrur MA, Valle V, Gangemi A, Bianco FM. Educational step-by-step surgical video about operative technique in robotic pancreaticoduodenectomy (RPD) at University of Illinois at Chicago (UIC): 17 steps standardized technique-Lessons learned since the first worldwide RPD performed in the year 2001. *Surg Endosc* 2020; **34**: 2758-2762 [PMID: 31953732 DOI: 10.1007/s00464-020-07383-0]

6 **Giulianotti PC**, Coratti A, Angelini M, Sbrana F, Cecconi S, Balestracci T, Caravaglios G. Robotics in general surgery: personal experience in a large community hospital. *Arch Surg* 2003; **138**: 777-784 [PMID: 12860761 DOI: 10.1001/archsurg.138.7.777]

7 **Mesleh MG**, Stauffer JA, Asbun HJ. Minimally invasive surgical techniques for pancreatic cancer: ready for prime time? *J Hepatobiliary Pancreat Sci* 2013; **20**: 578-582 [PMID: 23591745 DOI: 10.1007/s00534-013-0614-2]

8 **Gagner M**, Pomp A. Laparoscopic pancreatic resection: Is it worthwhile? *J Gastrointest Surg* 1997; **1**: 20-25 [PMID: 9834326 DOI: 10.1007/s11605-006-0005-y]

9 **Asbun HJ**, Moekotte AL, Vissers FL, Kunzler F, Cipriani F, Alseidi A, D'Angelica MI, Balduzzi A, Bassi C, Björnsson B, Boggi U, Callery MP, Del Chiaro M, Coimbra FJ, Conrad C, Cook A, Coppola A, Dervenis C, Dokmak S, Edil BH, Edwin B, Giulianotti PC, Han HS, Hansen PD, van der Heijde N, van Hilst J, Hester CA, Hogg ME, Jarufe N, Jeyarajah DR, Keck T, Kim SC, Khatkov IE, Kokudo N, Kooby DA, Korrel M, de Leon FJ, Lluis N, Lof S, Machado MA, Demartines N, Martinie JB, Merchant NB, Molenaar IQ, Moravek C, Mou YP, Nakamura M, Nealon WH, Palanivelu C, Pessaux P, Pitt HA, Polanco PM, Primrose JN, Rawashdeh A, Sanford DE, Senthilnathan P, Shrikhande SV, Stauffer JA, Takaori K, Talamonti MS, Tang CN, Vollmer CM, Wakabayashi G, Walsh RM, Wang SE, Zinner MJ, Wolfgang CL, Zureikat AH, Zwart MJ, Conlon KC, Kendrick ML, Zeh HJ, Hilal MA, Besselink MG; International Study Group on Minimally Invasive Pancreas Surgery (I-MIPS). The Miami International Evidence-based Guidelines on Minimally Invasive Pancreas Resection. *Ann Surg* 2020; **271**: 1-14 [PMID: 31567509 DOI: 10.1097/SLA.0000000000003590]

10 **Kim K**, Hagen ME, Buffington C. Robotics in advanced gastrointestinal surgery: the bariatric experience. *Cancer J* 2013; **19**: 177-182 [PMID: 23528727 DOI: 10.1097/PPO.0b013e318289dd15]

11 **Zeh HJ 3rd**, Bartlett DL, Moser AJ. Robotic-assisted major pancreatic resection. *Adv Surg* 2011; **45**: 323-340 [PMID: 21954697 DOI: 10.1016/j.yasu.2011.04.001]

12 **Lanfranco AR**, Castellanos AE, Desai JP, Meyers WC. Robotic surgery: a current perspective. *Ann Surg* 2004; **239**: 14-21 [PMID: 14685095 DOI: 10.1097/01.sla.0000103020.19595.7d]

13 **Kim BJ**, Prakash L, Narula N, Davis CH, Kim MP, Aloia TA, Vauthey JN, Lee JE, Katz MH, Tzeng CD. Contemporary analysis of complications associated with biliary stents during neoadjuvant therapy for pancreatic adenocarcinoma. *HPB (Oxford)* 2019; **21**: 662-668 [PMID: 30522947 DOI: 10.1016/j.hpb.2018.10.009]

14 **Miller-Ocuin JL**, Hogg ME, Zureikat AH, Zeh Iii HJ. Robotic Approaches to the Patient with Pancreatic Adenocarcinoma. In: Case-Based Lessons in the Management of Complex Hepato-Pancreato-Biliary Surgery. Pawlik TM, Weber S, Gamblin TC, editors. Cham: Springer International Publishing, 2017: 323-337

15 **Whipple AO**, Parsons WB, Mullins CR. TREATMENT OF CARCINOMA OF THE AMPULLA OF VATER. *Ann Surg* 1935; **102**: 763-779 [PMID: 17856666 DOI: 10.1097/00000658-193510000-00023]

16 **Hoehn RS**, Nassour I, Adam MA, Winters S, Paniccia A, Zureikat AH. National Trends in Robotic Pancreas Surgery. *J Gastrointest Surg* 2021; **25**: 983-990 [PMID: 32314230 DOI: 10.1007/s11605-020-04591-w]

17 **Moekotte AL**, Rawashdeh A, Asbun HJ, Coimbra FJ, Edil BH, Jarufe N, Jeyarajah DR, Kendrick ML, Pessaux P, Zeh HJ, Besselink MG, Abu Hilal M, Hogg ME; International Evidence-Based Guidelines on Minimally Invasive Pancreas Resection (IG-MIPR). Safe implementation of minimally invasive pancreas resection: a systematic review. *HPB (Oxford)* 2020; **22**: 637-648 [PMID: 31836284 DOI: 10.1016/j.hpb.2019.11.005]

18 **Tam V**, Zenati M, Novak S, Chen Y, Zureikat AH, Zeh HJ 3rd, Hogg ME. Robotic Pancreatoduodenectomy Biotissue Curriculum has Validity and Improves Technical Performance for Surgical Oncology Fellows. *J Surg Educ* 2017; **74**: 1057-1065 [PMID: 28578981 DOI: 10.1016/j.jsurg.2017.05.016]

19 **Nota CL**, Zwart MJ, Fong Y, Hagendoorn J, Hogg ME, Koerkamp BG, Besselink MG, Molenaar IQ; Dutch Pancreatic Cancer Group. Developing a robotic pancreas program: the Dutch experience. *J Vis Surg* 2017; **3**: 106 [PMID: 29078666 DOI: 10.21037/jovs.2017.07.02]

20 **Mark Knab L**, Zenati MS, Khodakov A, Rice M, Al-Abbas A, Bartlett DL, Zureikat AH, Zeh HJ, Hogg ME. Evolution of a Novel Robotic Training Curriculum in a Complex General Surgical Oncology Fellowship. *Ann Surg Oncol* 2018; **25**: 3445-3452 [PMID: 30073601 DOI: 10.1245/s10434-018-6686-0]

21 **Yan Q**, Xu LB, Ren ZF, Liu C. Robotic *vs* open pancreaticoduodenectomy: a meta-analysis of short-term outcomes. *Surg Endosc* 2020; **34**: 501-509 [PMID: 31848756 DOI: 10.1007/s00464-019-07084-3]

22 **Kamarajah SK**, Bundred J, Marc OS, Jiao LR, Manas D, Abu Hilal M, White SA. Robotic *vs* conventional laparoscopic pancreaticoduodenectomy a systematic review and meta-analysis. *Eur J Surg Oncol* 2020; **46**: 6-14 [PMID: 31409513 DOI: 10.1016/j.ejso.2019.08.007]

23 **Zureikat AH**, Beane JD, Zenati MS, Al Abbas AI, Boone BA, Moser AJ, Bartlett DL, Hogg ME, Zeh HJ 3rd. 500 Minimally Invasive Robotic Pancreatoduodenectomies: One Decade of Optimizing Performance. *Ann Surg* 2021; **273**: 966-972 [PMID: 31851003 DOI: 10.1097/SLA.0000000000003550]

24 **Nassour I**, Winters SB, Hoehn R, Tohme S, Adam MA, Bartlett DL, Lee KK, Paniccia A, Zureikat AH. Long-term oncologic outcomes of robotic and open pancreatectomy in a national cohort of pancreatic adenocarcinoma. *J Surg Oncol* 2020; **122**: 234-242 [PMID: 32350882 DOI: 10.1002/jso.25958]

25 **Zureikat AH**, Postlewait LM, Liu Y, Gillespie TW, Weber SM, Abbott DE, Ahmad SA, Maithel SK, Hogg ME, Zenati M, Cho CS, Salem A, Xia B, Steve J, Nguyen TK, Keshava HB, Chalikonda S, Walsh RM, Talamonti MS, Stocker SJ, Bentrem DJ, Lumpkin S, Kim HJ, Zeh HJ 3rd, Kooby DA. A Multi-institutional Comparison of Perioperative Outcomes of Robotic and Open Pancreaticoduodenectomy. *Ann Surg* 2016; **264**: 640-649 [PMID: 27433907 DOI: 10.1097/SLA.0000000000001869]

26 **Chen S**, Chen JZ, Zhan Q, Deng XX, Shen BY, Peng CH, Li HW. Robot-assisted laparoscopic *vs* open pancreaticoduodenectomy: a prospective, matched, mid-term follow-up study. *Surg Endosc* 2015; **29**: 3698-3711 [PMID: 25761559 DOI: 10.1007/s00464-015-4140-y]

27 **Zhang T**, Zhao ZM, Gao YX, Lau WY, Liu R. The learning curve for a surgeon in robot-assisted laparoscopic pancreaticoduodenectomy: a retrospective study in a high-volume pancreatic center. *Surg Endosc* 2019; **33**: 2927-2933 [PMID: 30483970 DOI: 10.1007/s00464-018-6595-0]

28 **Giulianotti PC**, Sbrana F, Bianco FM, Elli EF, Shah G, Addeo P, Caravaglios G, Coratti A. Robot-assisted laparoscopic pancreatic surgery: single-surgeon experience. *Surg Endosc* 2010; **24**: 1646-1657 [PMID: 20063016 DOI: 10.1007/s00464-009-0825-4]

29 **Narula VK**, Mikami DJ, Melvin WS. Robotic and laparoscopic pancreaticoduodenectomy: a hybrid approach. *Pancreas* 2010; **39**: 160-164 [PMID: 19910835 DOI: 10.1097/MPA.0b013e3181bd604e]

30 **Zhou NX**, Chen JZ, Liu Q, Zhang X, Wang Z, Ren S, Chen XF. Outcomes of pancreatoduodenectomy with robotic surgery *vs* open surgery. *Int J Med Robot* 2011; **7**: 131-137 [PMID: 21412963 DOI: 10.1002/rcs.380]

31 **Lai EC**, Yang GP, Tang CN. Robot-assisted laparoscopic pancreaticoduodenectomy *vs* open pancreaticoduodenectomy--a comparative study. *Int J Surg* 2012; **10**: 475-479 [PMID: 22732431 DOI: 10.1016/j.ijsu.2012.06.003]

32 **Chalikonda S**, Aguilar-Saavedra JR, Walsh RM. Laparoscopic robotic-assisted pancreaticoduodenectomy: a case-matched comparison with open resection. *Surg Endosc* 2012; **26**: 2397-2402 [PMID: 22437947 DOI: 10.1007/s00464-012-2207-6]

33 **Bao PQ**, Mazirka PO, Watkins KT. Retrospective comparison of robot-assisted minimally invasive *vs* open pancreaticoduodenectomy for periampullary neoplasms. *J Gastrointest Surg* 2014; **18**: 682-689 [PMID: 24234245 DOI: 10.1007/s11605-013-2410-3]

34 **Boone BA**, Zenati M, Hogg ME, Steve J, Moser AJ, Bartlett DL, Zeh HJ, Zureikat AH. Assessment of quality outcomes for robotic pancreaticoduodenectomy: identification of the learning curve. *JAMA Surg* 2015; **150**: 416-422 [PMID: 25761143 DOI: 10.1001/jamasurg.2015.17]

35 **Boggi U**, Napoli N, Costa F, Kauffmann EF, Menonna F, Iacopi S, Vistoli F, Amorese G. Robotic-Assisted Pancreatic Resections. *World J Surg* 2016; **40**: 2497-2506 [PMID: 27206401 DOI: 10.1007/s00268-016-3565-3]

36 **Nassour I**, Wang SC, Porembka MR, Yopp AC, Choti MA, Augustine MM, Polanco PM, Mansour JC, Minter RM. Robotic Versus Laparoscopic Pancreaticoduodenectomy: a NSQIP Analysis. *J Gastrointest Surg* 2017; **21**: 1784-1792 [PMID: 28819886 DOI: 10.1007/s11605-017-3543-6]

37 **Jin JB**, Qin K, Yang Y, Shi YS, Wu ZC, Deng XX, Chen H, Cheng DF, Shen BY, Peng CH. Robotic pancreatectomy for solid pseudopapillary tumors in the pancreatic head: A propensity score-matched comparison and analysis from a single center. *Asian J Surg* 2020; **43**: 354-361 [PMID: 31327550 DOI: 10.1016/j.asjsur.2019.05.016]

38 **Mejia A**, Shah J, Vivian E, Acharya P. Analysis of 102 Fully Robotic Pancreaticoduodenectomies: Clinical and Financial Outcomes. *Pancreas* 2020; **49**: 668-674 [PMID: 32433405 DOI: 10.1097/MPA.0000000000001545]

39 **Shi Y**, Jin J, Qiu W, Weng Y, Wang J, Zhao S, Huo Z, Qin K, Wang Y, Chen H, Deng X, Peng C, Shen B. Short-term Outcomes After Robot-Assisted *vs* Open Pancreaticoduodenectomy After the Learning Curve. *JAMA Surg* 2020; **155**: 389-394 [PMID: 32129815 DOI: 10.1001/jamasurg.2020.0021]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 110-116

DOI: 10.37126/aige.v2.i4.110

ISSN 2689-7164 (online)

MINIREVIEWS

Robotic surgery in colon cancer: current evidence and future perspectives – narrative review

Fulvio Tagliabue, Morena Burati, Marco Chiarelli, Ugo Cioffi, Mauro Zago

ORCID number: Fulvio Tagliabue 0000-0002-4095-017X; Morena Burati 0000-0003-2562-4760; Marco Chiarelli 0000-0003-1729-4925; Ugo Cioffi 0000-0002-5321-5828; Mauro Zago 0000-0001-9322-0798.

Author contributions: All authors participated equally in the manuscript.

Conflict-of-interest statement: The authors report no conflicts of interest.

Manuscript source: Invited manuscript

Specialty type: Surgery

Country/Territory of origin: Italy

Peer-review report's scientific

Fulvio Tagliabue, Morena Burati, Marco Chiarelli, Mauro Zago, Department of Emergency and Robotic Surgery, A. Manzoni Hospital–ASST Lecco, Lecco 23900, Italy

Ugo Cioffi, Department of Surgery, University of Milan, Milano 20122, Italy

Corresponding author: Ugo Cioffi, MD, PhD, Full Professor, Professor, Surgeon, Department of Surgery, University of Milan, Via F. Sforza 35, Milano 20122, Italy. ugocioffi5@gmail.com

Abstract

In the last 10 years, surgery has been developing towards minimal invasiveness; therefore, robotic surgery represents the consequent evolution of laparoscopic surgery. Worldwide, surgeons' performances have been upgraded by the ergonomic developments of robotic systems, leading to several benefits for patients. The introduction into the market of the new Da Vinci Xi system has made it possible to perform all types of surgery on the colon, an in selected cases, to combine interventions in other organs or viscera at the same time. Optimization of the suprapubic surgical approach may shorten the length of hospital stay for patients who undergo robotic colonic resection. From this perspective, single-port robotic colectomy, has reduced the number of robotic ports needed, allowing a better anesthetic outcome and faster recovery. The introduction on the market of new surgical robotic systems from multiple manufacturers is bound to change the landscape of robotic surgery and yield high-quality surgical outcomes.

Key Words: Colon cancer; Robotic surgery; Colectomy; Laparoscopy; Surgical outcomes: Robot system

Core Tip: Robotic surgery represents the natural evolution of laparoscopic surgery in the way to perform less-invasive operations. The robotic system Da Vinci Xi® with its technological innovations has made it possible to perform all types of interventions on the colon and has yielded large benefits to patients.

Citation: Tagliabue F, Burati M, Chiarelli M, Cioffi U, Zago M. Robotic surgery in colon cancer: current evidence and future perspectives – narrative review. *Artif Intell Gastrointest*

quality classification
Grade A (Excellent): 0
Grade B (Very good): B, B
Grade C (Good): 0
Grade D (Fair): D
Grade E (Poor): 0

Received: March 10, 2021
Peer-review started: March 10, 2021
First decision: May 3, 2021
Revised: May 14, 2021
Accepted: August 19, 2021
Article in press: August 19, 2021
Published online: August 28, 2021

P-Reviewer: Bustamante-Lopez LA, Mankaney G, Tsimogiannis K
S-Editor: Fan JR
L-Editor: Kerr C
P-Editor: Wang LYT

Endosc 2021; 2(4): 110-116
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/110.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.110

INTRODUCTION

Cancer of the colon and rectum is one of the most common neoplastic diseases worldwide and is associated with high mortality rate[1]. Just as laparoscopic surgery has progressively replaced laparotomy, robotic surgery is becoming increasingly important in the treatment of this type of cancer. The advantages of robotic systems have been well known for years. Wrist flexibility, 3D vision and prevention of hand tremor enable surgeons to operate in reduced operative fields.

Many technological innovations have been introduced in recent years, such as a suprapubic approach, single port techniques and the use of tracers such as indocyanine green (used for the research of the sentinel lymph node and to verify tissues' vascularization).

The efficiency and effectiveness of robotic colonic resection have drawn the attention of many surgeons. Just as laparoscopic surgery in the late 1990s was compared to open surgery in terms of safety and effectiveness, nowadays robot-assisted surgery is often compared to the laparoscopic approach. From this point of view, robotic surgery seems to overcome the limits of laparoscopy. In fact, the proper value of the robot can be clearly appreciated in challenging tasks, such as performing intra-abdominal anastomoses in a restricted space, or in low pelvic dissection[2].

Although early results seem to encourage robot-assisted surgery, comparative studies investigating the effects of laparoscopic *versus* robotic colonic surgery are still ongoing and have not yet provided definitive data[3,4].

ROBOTIC VERSUS LAPAROSCOPY

The indications for robot-assisted and laparoscopic colorectal surgery are the same. Relative contraindications are emergency procedures, pneumoperitoneum intolerance and massive bleeding.

Comparison between robotic and laparoscopic surgery in terms of advantages and disadvantages has been considered a "hot topic" lately. Detractors of robotic surgery doubt its effective usefulness, citing the lack of definitive data demonstrating its superiority compared to the traditional laparoscopic approach[5] (many have stated that it is an "expensive toy" built to entertain surgeons). Nevertheless, increasing data about the effectiveness of robot-assisted surgery, in addition to its well-described technical advantages, have drawn the attention of surgeons all over the world.

Since the da Vinci System has been approved, an increasing number of robotic procedures has been registered worldwide. As a consequence, available data on robotics in colorectal surgery have increased greatly. In the international scientific literature, single- and multicenter studies, systemic reviews and meta-analyses can be easily found, focusing on the evaluation of robotic outcomes[6]. Two National Impatient Sample databases of laparoscopic and robotic colectomies[7,8] found no significant differences in overall complication rates and length of stay, while conversion rates were significantly lower in patients who underwent robotic resection (6.3% *vs* 10.5 %). One large study, based on the American College of Surgeons National Surgical Quality Improvement Program database, compared robotic and laparoscopic colorectal surgery in more than 11000 patients[9]. Focusing on pelvic surgery, the rate of conversion to open approach was lower in the robotic surgery group, while no significant differences in conversion rates were found in abdominal surgery. No differences were found in rates of wound infection, anastomotic leak, 30-day reoperation and 30-day readmission. When robot-assisted surgery was performed, mean hospital stay was significantly shorter but operating times were significantly longer. The reason for longer operating time is easily imagined. Robotic surgery needs longer preparation in terms of patient and arm positioning, moreover, being a new technique, the learning curve of the performing surgeon strongly affects the overall operating time. In our opinion, this highlights the importance of continued evaluation of the advances in robot-assisted surgery compared to more traditional minimally invasive techniques.

A retrospective cohort study of the Michigan Surgical Quality Collaborative registry compared robotic *versus* 2735 laparoscopy-assisted colorectal procedures in 2012–2014 [10]. Conversion rates were lower in robotic surgery, and this was significant for rectal resection. Also, hospital stay was significantly shorter in those operated upon with the robotic technique. No significant difference in rates of complications were found.

In our opinion, the most meaningful, largest and better-designed study was the Robotic Versus Laparoscopic Resection for Rectal Cancer (ROLARR) Trial[11] published in 2017. It was an international, multicenter, randomized controlled trial (RCT), involving 10 countries and 29 centers. Primary outcome was conversion to open procedure when performing total mesorectal excision (TME). Intra- and postoperative complications, circumferential resection margin, quality of life, bladder and sexual dysfunction and oncological outcomes were considered secondary outcomes. The results showed no differences in conversion rates or other secondary endpoints, demonstrating that, in expert hands, robotic colonic resection is safe and feasible. What deserves to be highlighted is that, once again, robotic surgery did result in longer operating time. Only experienced surgeons were included in the study (surgeons who performed at least 90 laparoscopic or at least 50 robotic procedures), excluding the influence of the learning curve on operating time. Therefore, we can conclude that, more likely, robotic operating time is more affected by its longer patient preparation, and instrument placement and changing. In our opinion, it is important to highlight that conversion rates were lower in the robotic *versus* laparoscopic surgery in men. This suggests that, when it comes to narrower pelvis, robotic surgery could be superior to the laparoscopic approach, bringing great benefits to patients. The authors concluded that robotic surgery does not confer an advantage in rectal cancer and has equivalent outcomes with increased costs (due to the price of robotic instruments and components).

A meta-analysis of five RCTs in 2018[12], including ROLARR, by Prete *et al*[12] compared laparoscopic *versus* robotic resection for rectal cancer. The results demonstrated no significant differences in circumferential radial margin positive rate, TME grade, postoperative leakage, number of lymph nodes harvested, mortality or complication rate. This meta-analysis highlighted that robotic procedures are connected to a decreased rate of conversion to open surgery but, at the same time, a significant increase in operating time.

Conversion rate is an important outcome that can influence other outcomes. The passage from minimally invasive to open surgery can influence postoperative complication rates. It can also be the cause of increased costs (due to longer hospital stay) and delays in chemotherapy, which can affect 5-year disease-free survival, leading to higher recurrence rates[9,13,14].

All the advantages and disadvantages of robotic surgery are summarized in the Table 1.

From the analysis of the literature, the following conclusions can be drawn regarding the different aspects taken into consideration.

Postoperative days until the first flatus and first oral diet

Robot-assisted colorectal surgery is associated with a shorter time to first flatus and to first oral intake[15-17].

Time of operation

The literature shows longer operating time for robotic surgery[15-20]. In most cases, the reason is probably related to the early learning phase of the surgeons. We believe that after an adequate learning curve, surgical times should be significantly reduced to be compared to laparoscopic surgery. Nevertheless, it is easy to imagine that overall operating time will be always slightly longer for robotic surgery due to longer time needed for patients' preparation and instrument placement and changing.

Length of hospital stay

The robotic approach had a shorter hospital stay in several studies[19-25].

Mortality (perioperative or 30 d after the operation)

A few studies have demonstrated that mortality rate is significantly reduced in robotic surgery[20-26], but, on the contrary, other systematic reviews and meta-analysis have not confirmed this result[16,21-23].

Conversion to open surgery

It has been demonstrated that, compared to laparoscopy, robotic surgery is associated

Table 1 Advantages and disadvantages of robotic surgery

Advantages	Disadvantages
High-resolution 3D view	Longer operating times due to patient preparation and positioning and docking time
Tool and wrist flexibility (seven degrees of freedom)	Lack of tactile sensation and stenic feedback
Elimination of hand tremors	High acquisition and maintenance cost
Ergonomic position which benefits the surgeon	
Faster learning curve	
Dual console and simulation software for training	
Integrated table motion	
Four trocars visualization with fluorescent/optical systems	
Robot-designed tools, like robotic stapler with smart-fire technology	

with a significantly lower rate of conversion to open surgery. This is more relevant in high-risk patients, such as men with a narrow pelvis, obese patients with lower rectal tumors, or those undergoing neoadjuvant therapy[13,16-23].

Intraoperative blood loss

In terms of blood loss, some studies have reported significantly lower rates in robotic surgery[17,18,20,24].

Anastomotic leakage

As far as we know, no significant differences regarding anastomotic leakage have been found in the literature. In our opinion, in the near future the introduction of new automatized stapling systems and new robotic technologies will reduce the rate of anastomotic leakage.

Resected lymph nodes

No differences have been reported in the number of lymph nodes resected using robotic *versus* laparoscopic surgery, although some studies have shown a higher number of harvested lymph nodes in the robotic approach[15].

Sexual and urological outcomes

Considering rectal cancer surgery, recovery of sexual and urological function is faster in patients who have undergone a robot-assisted approach compared to laparoscopic surgery. In one retrospective cohort study, rates of erectile dysfunction 1 mo after surgery were similar in both laparoscopic and robotic groups. However, 1 year after complete recovery, physiological functions were completely restored in all sexually active patients who underwent robotic resection and only in 43% of patients in the laparoscopic group[25-27].

Surgical wound infection

Review articles and clinical trials have not shown any significant difference between the robotic and laparoscopic groups for surgical wound infection. There is only one systematic review published in 2019 by Ng *et al*[16] that showed a significant difference in favor of the robotic approach. We believe that future technological innovation will allow an increasing number of full robotic procedures, and consequently, the size of the skin incisions will progressively reduce, therefore decreasing surgical wound infections.

Resection margins

Simillis *et al*[28] in a systematic review and network meta-analysis published in *Annals of Surgery* in 2019[28] demonstrated no significant differences regarding the involved resection margins. A study by Nixon *et al*[29] focusing on high-risk patients (preoperative chemoradiotherapy, male sex, tumor < 8 cm from the anal verge, body mass index > 30, and previous abdominal surgery) demonstrated that robotic surgery is related to higher rates of sphincter preservation, lower conversion rates, lower blood

loss and operating time, and consequently it is associated with shorter length of hospital stay.

THE PRESENT AND THE FUTURE

With advances in engineering and technology, surgical robots are constantly being improved. Exploration of new surgical approaches like the suprapubic approach or single port technique is of interest in the surgical field. The suprapubic surgical approach refers to a particular robotic technique in which ports used to perform colonic resection are placed in a horizontal line in the suprapubic area, and it is usually applied in robotic right colectomy. Recently, some authors have demonstrated[30,31] that the suprapubic approach has more advantages than the traditional port placement, with less console time and shorter hospital stay. Surgeons are attempting to reduce the number of ports used for robotic surgery. By reducing the number of surgical wounds, they aim to reduce the risk of postoperative wound infections. In this light, single port robotic surgery has begun to be performed more often. A systemic review[32] revealed that single port robotic surgery for colonic cancer is safe and feasible, with acceptable postoperative outcomes. These new changes have demonstrated promising potential in robotic surgery, in particular in colonic resection.

Until now, the surgical robot market has been monopolized, but it is easy to predict that the market for robotic platforms will rapidly grow in the near future as several manufactures are investing in the development of new robotic systems. For instance, MicroHand S is a robotic system produced in China and has recently entered clinical trials. Some studies have reported good performances and encouraging application prospects[33,34]. Senhance robotic system (TransEnterix Surgical Inc. Morrisville, NC, USA) has been recently introduced in Europe and approved for limited clinical use in the USA. Darwich *et al*[35] and Samalavicius *et al*[36] reported that procedures performed with this robotic system were safe and feasible and the robot could be used in general surgery. Versius from Cambridge Medical Robotics Ltd (Cambridge, UK), Hugo RAS from Medtronic Inc. (Dublin, Ireland), Meere Company (South Korea), Titan Medical (Toronto ON, Canada) and Virtual Incision (Pleasanton, CA, USA) have demonstrated potential in clinical applications. Competition between these new surgical robots from different manufacturers will surely change the market, leading to a reduction in costs with increased benefits for patients.

CONCLUSION

Robotic surgery offers a new minimally invasive approach in complex procedures or in anatomical areas that are difficult to reach. Robot-assisted procedures are not easier to perform, but robotic technology can make hard tasks feasible for less-experienced surgeons. In our opinion, robotic surgery could be considered the best option for rectal cancer surgical treatment, especially when compared to more traditional approaches (laparoscopic, open or transanal), since it offers the best combination of oncological, functional and patient recovery outcomes. Furthermore, the development of new approaches, like suprapubic and single port techniques, and the use of new devices, like the robotic stapler or vessels and lymph nodes tracers, will allow us to reach better results in oncological and clinical terms. The introduction of new surgical robots from multiple different suppliers will reduce their cost, leading to the widespread of the robot-assisted approach for colonic resection.

ACKNOWLEDGMENTS

We thank Dr. Gerardo Cioffi, native speaker, for reviewing the English language.

REFERENCES

1 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]

2 **Addison P**, Agnew JL, Martz J. Robotic Colorectal Surgery. *Surg Clin North Am* 2020; **100**: 337-360

[PMID: 32169183 DOI: 10.1016/j.suc.2019.12.012]
3 **Crippa J**, Grass F, Dozois EJ, Mathis KL, Merchea A, Colibaseanu DT, Kelley SR, Larson DW. Robotic Surgery for Rectal Cancer Provides Advantageous Outcomes Over Laparoscopic Approach: Results From a Large Retrospective Cohort. *Ann Surg* 2020 [PMID: 32068552 DOI: 10.1097/SLA.0000000000003805]
4 **Justiniano CF**, Becerra AZ, Xu Z, Aquina CT, Boodry CI, Schymura MJ, Boscoe FP, Noyes K, Temple LK, Fleming FJ. A Population-Based Study of 90-Day Hospital Cost and Utilization Associated With Robotic Surgery in Colon and Rectal Cancer. *J Surg Res* 2020; **245**: 136-144 [PMID: 31419638 DOI: 10.1016/j.jss.2019.07.052]
5 **De Wilde RL**, Herrmann A. Robotic surgery - advance or gimmick? *Best Pract Res Clin Obstet Gynaecol* 2013; **27**: 457-469 [PMID: 23357200 DOI: 10.1016/j.bpobgyn.2012.12.005]
6 **Kim CW**, Kim CH, Baik SH. Outcomes of robotic-assisted colorectal surgery compared with laparoscopic and open surgery: a systematic review. *J Gastrointest Surg* 2014; **18**: 816-830 [PMID: 24496745 DOI: 10.1007/s11605-014-2469-5]
7 **Tyler JA**, Fox JP, Desai MM, Perry WB, Glasgow SC. Outcomes and costs associated with robotic colectomy in the minimally invasive era. *Dis Colon Rectum* 2013; **56**: 458-466 [PMID: 23478613 DOI: 10.1097/DCR.0b013e31827085ec]
8 **Halabi WJ**, Kang CY, Jafari MD, Nguyen VQ, Carmichael JC, Mills S, Stamos MJ, Pigazzi A. Robotic-assisted colorectal surgery in the United States: a nationwide analysis of trends and outcomes. *World J Surg* 2013; **37**: 2782-2790 [PMID: 23564216 DOI: 10.1007/s00268-013-2024-7]
9 **Bhama AR**, Obias V, Welch KB, Vandewarker JF, Cleary RK. A comparison of laparoscopic and robotic colorectal surgery outcomes using the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP) database. *Surg Endosc* 2016; **30**: 1576-1584 [PMID: 26169638 DOI: 10.1007/s00464-015-4381-9]
10 **Tam MS**, Kaoutzanis C, Mullard AJ, Regenbogen SE, Franz MG, Hendren S, Krapohl G, Vandewarker JF, Lampman RM, Cleary RK. A population-based study comparing laparoscopic and robotic outcomes in colorectal surgery. *Surg Endosc* 2016; **30**: 455-463 [PMID: 25894448 DOI: 10.1007/s00464-015-4218-6]
11 **Jayne D**, Pigazzi A, Marshall H, Croft J, Corrigan N, Copeland J, Quirke P, West N, Rautio T, Thomassen N, Tilney H, Gudgeon M, Bianchi PP, Edlin R, Hulme C, Brown J. Effect of Robotic-Assisted vs Conventional Laparoscopic Surgery on Risk of Conversion to Open Laparotomy Among Patients Undergoing Resection for Rectal Cancer: The ROLARR Randomized Clinical Trial. *JAMA* 2017; **318**: 1569-1580 [PMID: 29067426 DOI: 10.1001/jama.2017.7219]
12 **Prete FP**, Pezzolla A, Prete F, Testini M, Marzaioli R, Patriti A, Jimenez-Rodriguez RM, Gurrado A, Strippoli GFM. Robotic Versus Laparoscopic Minimally Invasive Surgery for Rectal Cancer: A Systematic Review and Meta-analysis of Randomized Controlled Trials. *Ann Surg* 2018; **267**: 1034-1046 [PMID: 28984644 DOI: 10.1097/SLA.0000000000002523]
13 **Cleary RK**, Mullard AJ, Ferraro J, Regenbogen SE. The cost of conversion in robotic and laparoscopic colorectal surgery. *Surg Endosc* 2018; **32**: 1515-1524 [PMID: 28916895 DOI: 10.1007/s00464-017-5839-8]
14 **Rottoli M**, Bona S, Rosati R, Elmore U, Bianchi PP, Spinelli A, Bartolucci C, Montorsi M. Laparoscopic rectal resection for cancer: effects of conversion on short-term outcome and survival. *Ann Surg Oncol* 2009; **16**: 1279-1286 [PMID: 19252948 DOI: 10.1245/s10434-009-0398-4]
15 **Tagliabue F**, Burati M, Chiarelli M, Fumagalli L, Guttadauro A, Arborio E, De Simone M, Cioffi U. Robotic *vs* laparoscopic right colectomy - the burden of age and comorbidity in perioperative outcomes: An observational study. *World J Gastrointest Surg* 2020; **12**: 287-297 [PMID: 32774767 DOI: 10.4240/wjgs.v12.i6.287]
16 **Ng KT**, Tsia AKV, Chong VYL. Robotic Versus Conventional Laparoscopic Surgery for Colorectal Cancer: A Systematic Review and Meta-Analysis with Trial Sequential Analysis. *World J Surg* 2019; **43**: 1146-1161 [PMID: 30610272 DOI: 10.1007/s00268-018-04896-7]
17 **Lim S**, Kim JH, Baek SJ, Kim SH, Lee SH. Comparison of perioperative and short-term outcomes between robotic and conventional laparoscopic surgery for colonic cancer: a systematic review and meta-analysis. *Ann Surg Treat Res* 2016; **90**: 328-339 [PMID: 27274509 DOI: 10.4174/astr.2016.90.6.328]
18 **Barashi NS**, Pearce SM, Cohen AJ, Pariser JJ, Packiam VT, Eggener SE. Incidence, Risk Factors, and Outcomes for Rectal Injury During Radical Prostatectomy: A Population-based Study. *Eur Urol Oncol* 2018; **1**: 501-506 [PMID: 31158094 DOI: 10.1016/j.euo.2018.06.001]
19 **Ohtani H**, Maeda K, Nomura S, Shinto O, Mizuyama Y, Nakagawa H, Nagahara H, Shibutani M, Fukuoka T, Amano R, Hirakawa K, Ohira M. Meta-analysis of Robot-assisted *Versus* Laparoscopic Surgery for Rectal Cancer. *In Vivo* 2018; **32**: 611-623 [PMID: 29695568 DOI: 10.21873/invivo.11283]
20 **Lee SH**, Kim DH, Lim SW. Robotic vs laparoscopic intersphincteric resection for low rectal cancer: a systematic review and meta-analysis. *Int J Colorectal Dis* 2018; **33**: 1741-1753 [DOI: 10.1007/s00384-018-3145-0]
21 **Grass JK**, Perez DR, Izbicki JR, Reeh M. Systematic review analysis of robotic and transanal approaches in TME surgery- A systematic review of the current literature in regard to challenges in rectal cancer surgery. *Eur J Surg Oncol* 2019; **45**: 498-509 [PMID: 30470529 DOI: 10.1016/j.ejso.2018.11.010]
22 **Zheng B**, Zhang X, Wang X, Ge L, Wei M, Bi L, Deng X, Wang Q, Li J, Wang Z. A comparison of

open, laparoscopic and robotic total mesorectal excision: trial sequential analysis and network meta-analysis. *Colorectal Dis* 2020; **22**: 382-391 [PMID: 31600858 DOI: 10.1111/codi.14872]

23 **Phan K**, Kahlaee HR, Kim SH, Toh JWT. Laparoscopic vs. robotic rectal cancer surgery and the effect on conversion rates: a meta-analysis of randomized controlled trials and propensity-score-matched studies. *Tech Coloproctol* 2019; **23**: 221-230 [PMID: 30623315 DOI: 10.1007/s10151-018-1920-0]

24 **Tang B**, Gao GM, Zou Z, Liu DN, Tang C, Jiang QG, Lei X, Li TY. [Efficacy comparison between robot-assisted and laparoscopic surgery for mid-low rectal cancer: a prospective randomized controlled trial]. *Zhonghua Wei Chang Wai Ke Za Zhi* 2020; **23**: 377-383 [PMID: 32306606 DOI: 10.3760/cma.j.cn.441530-20190401-00135]

25 **Luca F**, Craigg DK, Senthil M, Selleck MJ, Babcock BD, Reeves ME, Garberoglio CA. Sexual and urinary outcomes in robotic rectal surgery: review of the literature and technical considerations. *Updates Surg* 2018; **70**: 415-421 [PMID: 30120743 DOI: 10.1007/s13304-018-0581-x]

26 **Wang G**, Wang Z, Jiang Z, Liu J, Zhao J, Li J. Male urinary and sexual function after robotic pelvic autonomic nerve-preserving surgery for rectal cancer. *Int J Med Robot* 2017; **13** [PMID: 26748601 DOI: 10.1002/rcs.1725]

27 **D'Annibale A**, Pernazza G, Monsellato I, Pende V, Lucandri G, Mazzocchi P, Alfano G. Total mesorectal excision: a comparison of oncological and functional outcomes between robotic and laparoscopic surgery for rectal cancer. *Surg Endosc* 2013; **27**: 1887-1895 [PMID: 23292566 DOI: 10.1007/s00464-012-2731-4]

28 **Simillis C**, Lal N, Thoukididou SN, Kontovounisios C, Smith JJ, Hompes R, Adamina M, Tekkis PP. Open Versus Laparoscopic Versus Robotic Versus Transanal Mesorectal Excision for Rectal Cancer: A Systematic Review and Network Meta-analysis. *Ann Surg* 2019; **270**: 59-68 [PMID: 30720507 DOI: 10.1097/SLA.0000000000003227]

29 **Nixon J**, Brown S, Smith IL, McGinnis E, Vargas-Palacios A, Nelson EA, Brown J, Coleman S, Collier H, Fernandez C, Gilberts R, Henderson V, McCabe C, Muir D, Rutherford C, Stubbs N, Thorpe B, Wallner K, Walker K, Wilson L, Hulme C. Comparing alternating pressure mattresses and high-specification foam mattresses to prevent pressure ulcers in high-risk patients: the PRESSURE 2 RCT. *Health Technol Assess* 2019; **23**: 1-176 [PMID: 31559948 DOI: 10.3310/hta23520]

30 **Hamilton AER**, Chatfield MD, Johnson CS, Stevenson ARL. Totally robotic right hemicolectomy: a multicentre case-matched technical and peri-operative comparison of port placements and da Vinci models. *J Robot Surg* 2020; **14**: 479-491 [PMID: 31468314 DOI: 10.1007/s11701-019-01014-0]

31 **Lee HJ**, Choi GS, Park JS, Park SY, Kim HJ, Woo IT, Park IK. A novel robotic right colectomy for colon cancer *via* the suprapubic approach using the da Vinci Xi system: initial clinical experience. *Ann Surg Treat Res* 2018; **94**: 83-87 [PMID: 29441337 DOI: 10.4174/astr.2018.94.2.83]

32 **Bae SU**, Jeong WK, Baek SK. Current status of robotic single-port colonic surgery. *Int J Med Robot* 2017; **13** [PMID: 26913985 DOI: 10.1002/rcs.1735]

33 **Yi B**, Wang G, Li J, Jiang J, Son Z, Su H, Zhu S, Wang S. Domestically produced Chinese minimally invasive surgical robot system "Micro Hand S" is applied to clinical surgery preliminarily in China. *Surg Endosc* 2017; **31**: 487-493 [PMID: 27194259 DOI: 10.1007/s00464-016-4945-3]

34 **Luo D**, Liu Y, Zhu H, Li X, Gao W, Zhu S, Yu X. The MicroHand S robotic-assisted versus Da Vinci robotic-assisted radical resection for patients with sigmoid colon cancer: a single-center retrospective study. *Surg Endosc* 2020; **34**: 3368-3374 [PMID: 31482355 DOI: 10.1007/s00464-019-07107-z]

35 **Darwich I**, Stephan D, Klöckner-Lang M, Scheidt M, Friedberg R, Willeke F. A roadmap for robotic-assisted sigmoid resection in diverticular disease using a Senhance™ Surgical Robotic System: results and technical aspects. *J Robot Surg* 2020; **14**: 297-304 [PMID: 31161448 DOI: 10.1007/s11701-019-00980-9]

36 **Samalavicius NE**, Janusonis V, Siaulys R, Jasėnas M, Deduchovas O, Venckus R, Ezerskiene V, Paskeviciute R, Klimaviciute G. Robotic surgery using Senhance® robotic platform: single center experience with first 100 cases. *J Robot Surg* 2020; **14**: 371-376 [PMID: 31301021 DOI: 10.1007/s11701-019-01000-6]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 117-126

DOI: 10.37126/aige.v2.i4.117

ISSN 2689-7164 (online)

MINIREVIEWS

Artificial intelligence in endoscopy: The challenges and future directions

Xiaohong Gao, Barbara Braden

ORCID number: Xiaohong Gao 0000-0002-8103-6624; Barbara Braden 0000-0002-8534-6873.

Author contributions: Gao XH and Braden B contributed to the literature research and writing of the manuscript; Both authors have read and approved the final manuscript.

Conflict-of-interest statement: The authors have no interests to declare.

Manuscript source: Unsolicited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: United

Xiaohong Gao, Department of Computer Science, Middlesex University, London NW4 4BT, United Kingdom

Barbara Braden, Translational Gastroenterology Unit, Oxford University Hospitals NHS Foundation Trust, Oxford OX3 9DU, United Kingdom

Corresponding author: Xiaohong Gao, PhD, Full Professor, Department of Computer Science, Middlesex University, The Burroughs, Hendon, London NW4 4BT, United Kingdom.
x.gao@mdx.ac.uk

Abstract

Artificial intelligence based approaches, in particular deep learning, have achieved state-of-the-art performance in medical fields with increasing number of software systems being approved by both Europe and United States. This paper reviews their applications to early detection of oesophageal cancers with a focus on their advantages and pitfalls. The paper concludes with future recommendations towards the development of a real-time, clinical implementable, interpretable and robust diagnosis support systems.

Key Words: Deep learning; Oesophageal cancer; Early detection; Squamous cell cancer; Barrett's oesophagus

Core Tip: Precancerous changes in the lining of the oesophagus are easily missed during endoscopy as these lesions usually grow flat with only subtle change in colour, surface pattern and microvessel structure. Many factors impair the quality of endoscopy and subsequently the early detection of oesophageal cancer. Artificial intelligence (AI) solutions provide independence from the skills and experience of the operator in lesion recognition. Recent developments have introduced promising AI systems that will support the clinician in recognising, delineating and classifying precancerous and early cancerous changes during the endoscopy of the oesophagus in real-time.

Citation: Gao X, Braden B. Artificial intelligence in endoscopy: The challenges and future

Kingdom

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): B
Grade C (Good): 0
Grade D (Fair): 0
Grade E (Poor): 0

Received: May 22, 2021
Peer-review started: May 22, 2021
First decision: June 18, 2021
Revised: June 20, 2021
Accepted: July 15, 2021
Article in press: July 15, 2021
Published online: August 28, 2021

P-Reviewer: Muneer A
S-Editor: Liu M
L-Editor: A
P-Editor: Zhang YL

directions. *Artif Intell Gastrointest Endosc* 2021; 2(4): 117-126
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/117.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.117

INTRODUCTION

AI is the artificial intelligence exhibited by computer machines, which is in opposition to the natural intelligence that is displayed by human being, including consciousness and emotionality. With the advances on both computer hardware and software technology, at present, we are able to model about 600 K neurons and their interlaced connections, leading to processing over 100 million parameters. Since the human brain contains about 100 billion neurons[1], there is still a long way to go to before AI models are close enough to a human brain. Hence machine learning (ML) techniques are developed to perform task specific modelling that is in part supervised by human. While this supervised ML process is transparent and understandable, the human's ability to comprehend large amounts of parameters, *e.g.*, in millions, is limited, from a calculation point of view. Hence the application areas are restricted by employing semi- or fully supervised ML approaches. More recently, propelled by the advances of computer hardware, including large memory and graphics processing unit (GPU), task specific learning by computer itself, *i.e.*, deep learning (DL), is realised, forming one of the most promising AI branches under the ML umbrella.

DL first made the headline when DL based computer program, AlphaGo, won the competition when playing board game Go with human players[2]. Since then, it has shown that nearly all winners in major competitions apply DL led methodologies, achieving state-of-the-art (SOTA) performance in nearly every domain, including natural language translation and image segmentation and classification. For example, the competition organised by Kaggle on detection of diabetics based on retinopathy has been won by DL based approach by a large margin in comparison with the other methods. While DL oriented methods have become a mainstream choice of methodology, there are advantaged and disadvantages, especially in the medical field. For example, a DL-based approach requires large amount of training datasets, better in millions, which is hardly met in medical domains. In addition, the training in deep layers demands higher computational power, leading to real-time processing a great challenge.

Hence this paper aims to review the latest development of application of AI to endoscopy realm and is organised below. Section 2 details the SOTA DL techniques and their application to medical domains. Section 3 explores the challenges facing early detection of oesophageal diseases from endoscopy and current solutions of computer aided systems. Section 4 points out future directions in achieving accurate diagnosis of oesophageal diseases with summaries provided in conclusion.

STATE OF THE ART DL TECHNIQUE AND ITS APPLICATION TO MEDICAL FIELD

DL neural networks refer to a class of computing algorithms that can learn a hierarchy of features by establishing high-level attributes from low-level ones. One of the most popular models remains the convolutional neural network (CNN)[3], which comprises several (deep) layers of processing involving learnable operators (both linear and non-linear), by automating the process of constructing discriminative information from learning hierarchies. In addition, recent advances in computer hardware technology (*e.g.*, the GPU) have propagated the implementation of CNNs in studying images. Usually, training a DL system to perform a task, *e.g.*, classification, employs an architecture in an end-to-end training fashion. As a result, by input of a raw datum, the trained system will output a classification label. The training activity takes place by processing the input data with known annotations (labels, or segmented regions) with a goal to establish a model to differentiate these annotated labels/region automatically by fine-tuning the relationship between parameters without the intervention of humans.

Conventionally, training a DL model requires large datasets and substantial training time. For example, the pre-trained CNN classifier, AlexNet[4], is built upon 7 Layers, simulating 500000 (K) neurons with 60 million (M) parameters and 630 M connections,

and trained on a subset (1.2 M with 1 K categories) of ImageNet with 15 M 2D images of 22 K categories, taking up 16 d on a CPU and 1.6 d on a GPU. Usually, more data will lead to more accurate systems. In the development of electric cars of Tesla's Autopilot, the training takes place with more than 780 million miles[5] whereas for playing AlphaGo[6] game using a computer, the training employed more than 100 million games.

DL-oriented approaches have recently been applied to medical images in a range of domains and achieved SOTA results. Although some doubt on DL has been casted on the 'black box' status while training without the embedding of human's knowledge in the middle stages (*e.g.,* hidden layers) apart from the initial input of labelled datasets, the performance of AI-led approached has been widely recognised, which is evidenced by the approval of medical devices by authorities. Between the year 2015 and 2020, 124 (about 15%) medical devices (mainly software) that are AL/ML/DL-based have been approved in Europe with *Conformité Européene* -marked and United States Food and Drug administration agency[7], highlighting the importance of AI/ML to the medical field, including an imaging system that uses algorithms to give diagnostic information for skin cancer and a smart electrocardiogram device that estimates the probability of a heart attack[8]. Table 1 summaries the recent achievements of DL-oriented approaches in medical domains.

Recently, AI or more specific DL-based approaches have won a number of competitions including the Kaggle competition on detection of diabetic retinopathy, segmentation of brain tumors from MRI images[9], analysis of severity of tuberculosis (TB) from high resolution 3D CT images in Image CLEFmed Competition[10] and detection of endoscopic artefacts from endoscopy video images in EAD2019[11] and EAD2020[12].

While applying AI/ML/DL approaches in medical domain, there are several challenges in need of responding. Firstly, in the medical domain, the number of datasets is limited, usually in hundreds whereas in other application, *e.g.,* self-driving cars, datasets are in millions. Secondly, images are in multiple dimensions ranging from 2D to 5D (*e.g.,* a moving heart at a specific location). And thirdly, perhaps the most outstanding obstacle is that medical data present subtle changes between normal and abnormal demanding the developed systems to be more precise.

Hence progress has been made to allow additional measures to be taken into account in order to apply DL techniques in medical fields. For example, for classification of 3D echocardiographic video images[13], a fused CNN architecture is established to incorporate both unsupervised CNN and hand-crafted features. For classification of 3D CT brain images[14], integration of both 2D and 3D CNN networks is in place. In addition, patch-based DL technique is designed to analyse 3D CT images for classification of TB types and analysis of multiple drug resistance[15,16] to overcome the sparse presence of diseased regions (< 10%). Another way to address small dataset issue is to employ transfer ML technique that is frequently implemented whereby a model developed built upon one dataset (*e.g.,* ImageNet) for a specific task is reused as a starting point for a model on a different task with completely different datasets [*e.g.,* coronavirus disease 2019 (COVID-19) computed tomography (CT) images]. Subsequently, most currently developed learning systems commence with a pre-trained model, such as VGG16[17] that is pre-trained on ImageNet datasets to extract initial feature maps that are then retrained to fit the new datasets and new tasks[18], capitalising on the accuracy a pre-trained model sustaining whilst saving considerable training times.

More recently, these AI techniques have been applied to predict COVID-19 virus and have demonstrated significant performance. With regard to medical images for diagnosis of COVID-19, CT and chest X-ray (CXR) represent the most common imaging tools. For 3D CT images, attention-based DL networks have shown effectiveness in classifying COVID-19 from normal subjects[19,20]. In relation to CXR, patch-based CNN is applied to study chest x-ray images[21] and to differentiate discriminatory features of COVID-19. In addition, COVID-Net[22], one of the pioneer studies, classifies COVID-19 from normal and pneumonia diseases through the application of a tailored DL network. To overcome the shortage of datasets, a number of researchers[23] apply generative adversarial neural network (GAN) to augment data first and subsequently to classify COVID-19.

In this paper, the application of AI/ML/DL techniques is exploited to endoscopy video images.

Table 1 Examples of deep learning-based approaches in application of medical tasks

Ref.	Medical domain	Tasks
Muehlematter *et al*[7]	Skin	Diagnosis of skin cancer
United States Food and Drug Administration[8]	Electrocardiogram	Detection of heart attack
Pereira *et al*[9]	Retinopathy	Detection of diabetics
Gao *et al*[10]	Pulmonary CT images	Detection of tuberculosis types and severity
Sharib *et al*[11], Ali *et al*[12]	Endoscopy	Detection of artefact.
Gao *et al*[14]	CT Brain images	Classification of Alzheimer's disease
Gao *et al*[15,16]	Pulmonary CT images	Analysis of multi-drug resistance
Gao *et al*[13]	Ultrasound	Classification of 3D echocardiographic video images
Wang *et al*[19], Ouyang *et al*[20]	Chest CT	Diagnosis of COVID-19
Oh *et al*[21], Wang *et al*[22], Waheed *et al*[23]	Chest X-Ray	Diagnosis of COVID-19
Everson *et al*[33], Horie *et al*[34], Ghatwary *et al*[35], Ohmori *et al*[38]	Endoscopy	Still image based cancer detection for 2 classes (normal vs abnormal)
de Groof *et al*[32], Everson *et al*[35], He *et al*[41], Guo *et al*[42]	Endoscopy	Video detection of SCC in real time
Gao *et al*[44], Tomita *et al*[45]	Endoscopy	Explainable AI for early detection of SCC

CT: Computed tomography; AI: Artificial intelligence; SCC: Squamous cell cancer; COVID-19: Coronavirus disease 2019.

ENDOSCOPY FOR DIAGNOSIS OF OESOPHAGEAL DISEASES

The oesophagus is the muscular tube that carries food and liquids from mouth to the stomach. The symptoms of oesophageal disorders include chest or back pain or having trouble swallowing. The most common problem with the oesophagus is gastroesophageal reflux disease which occurs when stomach contents frequently leak back, or reflux, into the oesophagus. The acidity of the fluids can irritate the lining of the oesophagus. Treatment of these disorders depends on the problem. Some problems get better with over-the-counter medicines or changes in diet. Others may need prescribed medicines or surgery.

As the 8th most common cancer worldwide[24], one of the most serious problems with regard to oesophagus is oesophageal cancer that constitutes the 6th leading cause of cancer-related death[25]. The main cancer types include adenocarcinoma and squamous cell carcinoma cancer (SCC). Globally, about 87% of all oesophageal cancers are in the form of SCC. The highest incidence rates often take place in Asia, the Middle East and Africa[26,27]. Early oesophageal cancer usually does not cause symptoms. At later stage, the symptoms might include swallowing difficulty, weight loss or continuous cough. Diagnosis of oesophageal cancer relies on imaging test, an upper endoscopy, and a biopsy.

Optical endoscopy or endoscopy is the primary diagnostic and therapeutic tool for management of gastrointestinal malignancies, in particular oesophagus cancers. As illustrated in Figure 1A, to perform an endoscopy procedure of monitoring oesophagus, an endoscopic camera along with a lighting inspection is inserted into the food pipe of the patient in concern, whereby the appearance inside the oesophageal tube in the form of video images can be visualised on a computer monitor that is linked to the camera image processing system, which is depicted in Figure 1B.

While Figure 1 presents the surface of oesophageal walls, it also shows the artefact in a number of frames. This is because the movements of the inserted camera is confined within the limited space of the food pipe. The most common artefacts include colour misalignment (C), burry (B), saturation (S), and device (D) as demonstrated in Figure 1B.

Challenges for detecting oesophageal squamous cancer

Commonly the five-year survival rate of oesophagus cancer is less than 20% as reported in[28]. However, this rate can be improved significantly to more than 90% if the cancer is detected in its early stages due to the fact that at this early stage,

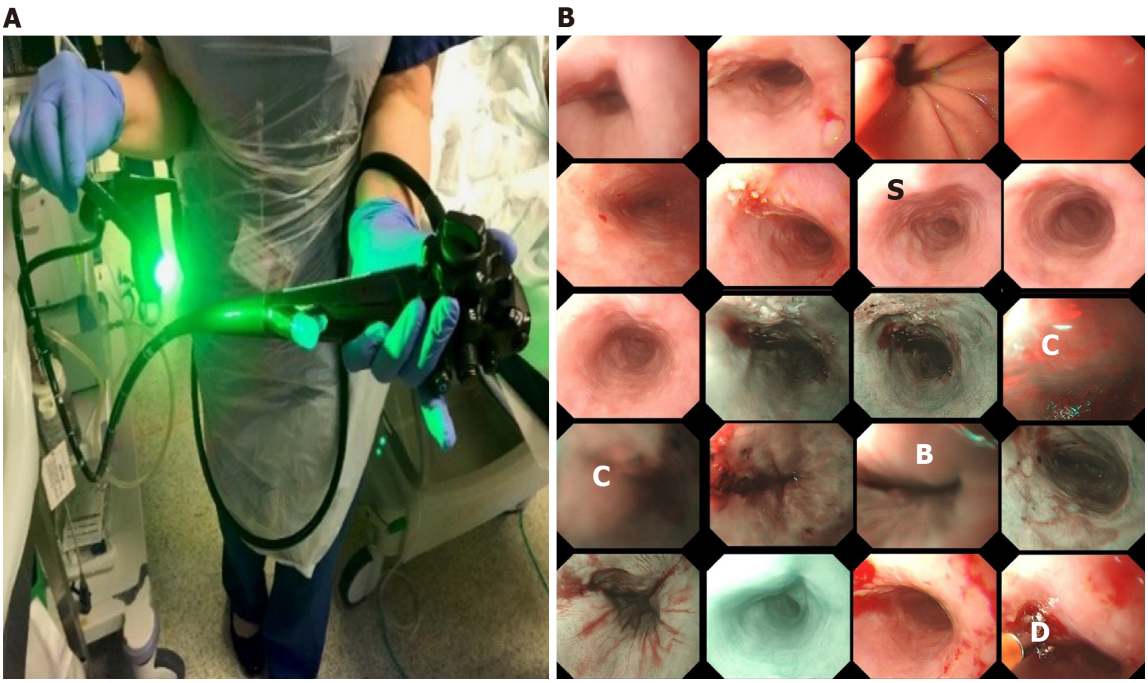

Figure 1 The endoscopy procedure. A: The oesophagus camera; B: A montage display of a clip of an endoscopic video including narrow-band imaging and conventional white light endoscopy (*e.g.*, top 2 rows). C: Colour misalignment; S: Saturation; B: Blurry; D: Device.

oesophageal cancer can be treated endoscopically[29], *e.g.*, by removing diseased tissues or administrating (spraying) treatment drugs. The challenge lies here is that precancerous stages (dysplasia in the oesophageal squamous epithelium) and early stages of SCC display subtle changes in appearance (*e.g.*, colour, surface structure) and in microvasculature, which therefore are easily missed at the time of conventional white light endoscopy (WLE) as illustrated in Figure 2A-D. To overcome this shortcoming while viewing WLE images, narrow-band imaging (NBI) can be turned on to display only two wavelengths [415 nm (blue) and 540 nm (green)] (Figure 2E-G) to improve the visibility of those suspected lesions by filtering out the rest of colour bands. Another approach is dye-based chromoendoscopy, *i.e.* Lugol's staining technique, which highlights dysplastic abnormalities by spraying iodine[30] (Figure 2H).

While NBI technique improves the visibility of the vascular network and surface structure, it mainly facilitates the detection of unique vascular and pit pattern morphology that are present in neoplastic lesions[31], whereas precancerous stages can take a variety of forms. With the Lugol's staining approach, many patients react uncomfortably to the spray.

It is therefore of clinical priority to have a computer assisted system to help clinicians to detect and highlight those potential suspected regions for further examinations. Currently, a number of promising results for computer-aided recognition of early neoplastic oesophageal lesions from endoscopic have been achieved based still images[32,33]. However, fewer less algorithms are applicable to real-time endoscopy to allow computer-aided decision-making during endoscopy at the point of examination. In addition, most of the existing studies focus mainly on the classification of endoscopic images between normal and abnormal stages with little work providing bounding boxes of the suspicious regions (detection) and delineating (segmentation).

Following challenges have been identified for the development of computerised algorithms for early detection of oesophageal cancers, which are inconspicuous changes on oesophageal surfaces artefacts of video images due to movement of endoscopic camera entering the food pipe limited time for patients undergoing each session of endoscopic procedure (about 20min) to minimise discomfort and invasiveness real time processing of video images to be in time to prompt endoscopist collecting biopsy samples while undertaking endoscopy limited datasets to train DL systems multiple modalities, including WLE, NBI and Lugol's multiple classes, including LD, GD, SCC, normal, and artefact.

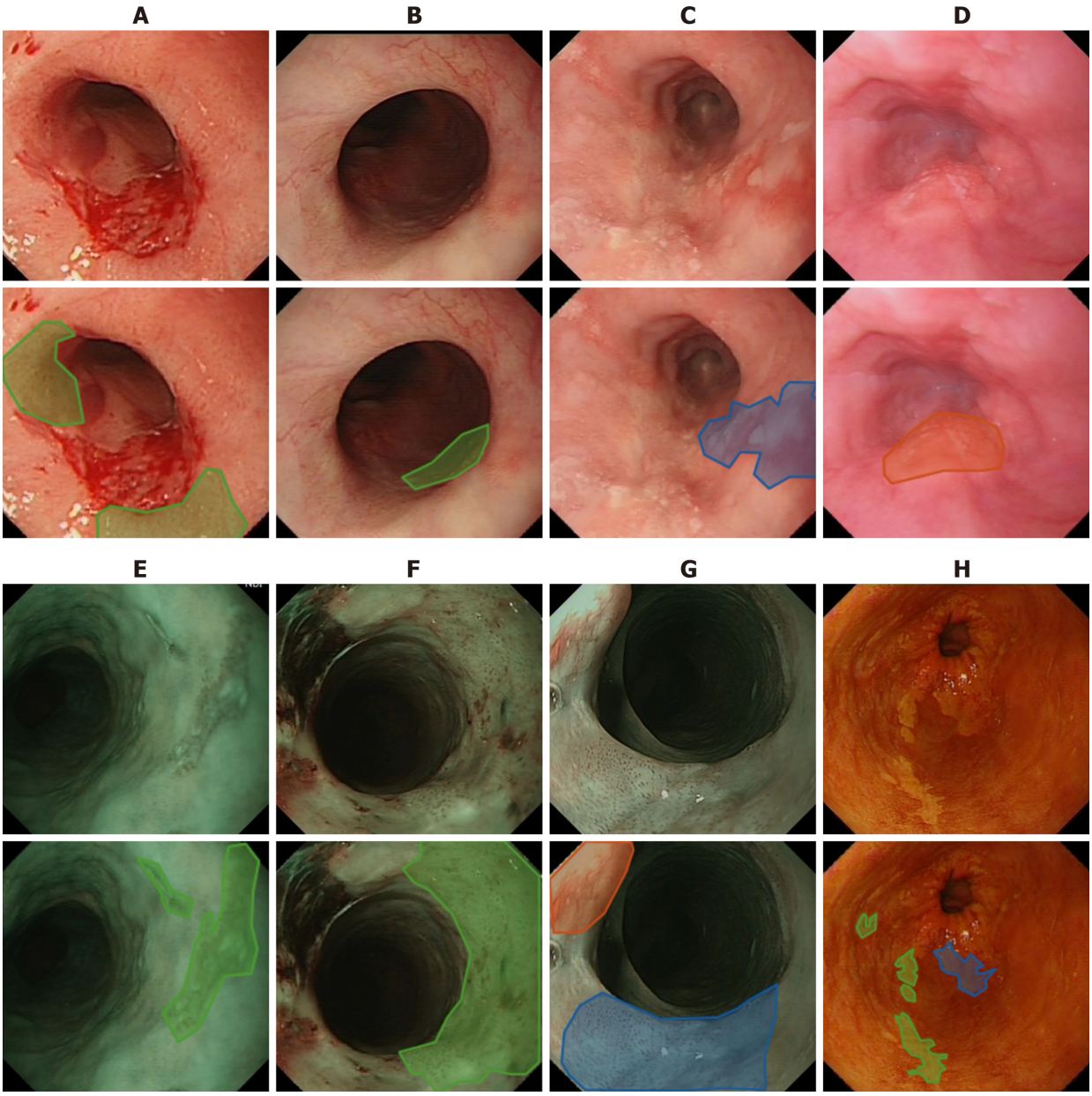

Figure 2 Examples of endoscopic images where green and blue masks refer to low and high grade dysplasia respectively and red for squamous cell cancer. A-D: White light endoscopy; E-G: Narrow band imaging; H: Lugol's. Mask colours: Green = low grade dysplasia; Blue = high grade dysplasia; Red = Squamous cell cancer.

Progress on the development of AI-based computer assisted supporting system for early detection of SCC

Progress on diagnosis of oesophageal cancer through the application of AI has been made by several research teams, mainly focusing on three directions, classification of abnormal from normal images, classification taking into consideration of processing speed, and detection of artefacts.

AI-based classification

Horie *et al*[34] conducted research to distinguish oesophageal cancers from non-cancer patients with an aim to reproduce diagnostic accuracy. While applying conventional CNN architecture to classify two classes, the researchers have achieved 98% sensitivity for cancer detection. In the study conducted by Ghatwary *et al*[35], researchers have evaluated several SOTA CNN approaches aiming to achieve early detection of SCC from high-definition WLE (HD-WLE) images and come to the conclusion that the approaches of single shot detection[36] and Faster R-CNN[37] perform better. They use one image modality of WLE. Again, two classes are investigated in their study, *i.e.*, cancerous and normal regions. While these studies demonstrate high accuracy of

classification, the main focus of those research remains on the binary classification distinguishing abnormal from normal. Similarly, in the study by Ohmori *et al*[38], while the authors studied oesophageal lesions on several imaging modes including blue-laser images, only two classes of either cancer or non-cancer are classified by employing a deep neural network. For detection of any potential suspected regions regardless how small they are, segmentation of abnormal regions also plays a key role in supporting clinical decisions.

Classification with near real-time processing

In addition, in order to assist clinicians in early diagnosis during endoscopic procedures, real-time processing of videos, *i.e.*, with processing speed of 24+ frames per second (fps) or at most 41 milliseconds (ms) per frame, should be realised. Everson *et al*[33] have achieved inference time between 26 to 37ms for an image of 696 × 308 pixels. The work conducted by de Groof *et al*[32] requires 240ms to process each frame (*i.e.*, 4.16 fps). For processing a video clip, frame processing and video playing back times all need to be considered to allow processed frames being played back seamlessly.

In order to ensure lesion detection takes place in time while patients undertaking endoscopy procedure, processing speed constitutes one of the key elements. Hence, comparisons are made to devalue the processing speed when detecting, classifying, and delineating multi-class (LD, HD, SCC) on multi-modality images (WLE, NBI, Lugol's)[39] employing DL architectures of YOLOv3[40] and mask-CNN[41]. In this study by applying YOLOv3, the average processing time is in the range of 0.064-0.101 s per frame, which leads to 10-15 frames per second while processing frames of endoscopic videos with a resolution of 1920 × 1080 pixels. This work was conducted under Windows 10 operating system with 1 GPU (GeForce GTX 1060). The averaged accuracies for classification and detection can be realised to 85% and 74% respectively. Since YOLOv3 only provides bounding boxes without masks, the approach of mask-RCNN is utilised to delineate lesioned regions, producing classification, segmentation (masks) and bounding boxes. As a result, mask-RCNN achieves better detection result (*i.e.*, bounding box) with 77% accuracy whereas the classification accuracy is similar to that obtained using YOLOYv3 with 84%. However, the processing speed applying mask-RCNN appears to be more than 10 times slower with an average of 1.2 s per frame, which is mainly stemmed from the time spent on the creation of masks. For the segmentation while employing mask-RCNN, the accuracy retains 63% measured on the overlapping regions between predicted and ground truth regions.

More recently, a research group by Guo *et al*[42] has developed a CAD system to aid decision making for early diagnosis of precancerous lesions. Their system can realise video processing time at 25 frames per second while applying narrow band images (NBI) that present clearer lesion structures than WLE. It appears that only one detection is identified for each frame, hence the study does not support localisation by bounding boxes.

Artefact detection

Due to the confined space to film the oesophageal tube, a number of artefacts are present, which not only hamper clinician's visual interpretation but also mislead training AI-based systems. Therefore, endoscopic artefact detection challenges were organised in 2019 (EAD2019)[11] and 2020 (EAD2020)[12] aiming to find solutions to these challenges. As expected, all top performant teams apply DL-based approaches to detect (bounding box), classify and segment artefacts including bubbles, saturation, blurry and artefacts[43].

FUTURE WORK

While significant progress has been made towards development of AI-enhanced systems to support clinicians' diagnosis, especially for early detection of oesophageal cancer, there is a still a considerable distance to go to benefit clinical diagnosis and to equip these assistant systems in an operative room. The following recommendations might shed light on future research directions.

Firstly, detection should be based on multi-classes, especially early onset lesions should be included. This is because most of the currently developed systems work on binary classifications between cancer and normal whereas cancers present most distinguishable visual features. At present, in 1 in 4 patients, the diagnosis of early stage oesophageal cancer is missed in their first visit[30]. Hence more work should emphasis

on the detection of early onset of SCC. Only in this way can patients' 5-year survival rates be increased to 90% from current 20%.

In addition, to circumvent data shortage, conventional data augmentation techniques appear to increase system accuracy by cropping, colour shifting, resizing and rotating. Due to the subtle change of early stages of SCC, data augmentation by inclusion of fake datasets generated by employing generative adversarial DL networks (GAN) appear to decrease the performance in this regard. Furthermore, when training with data that include samples with artefact, data augmentation with colour shifting also tend to hamper the system performance. Computational spectral imaging appears to benefit in this regard.

Secondly, to increase the wide acceptance by clinicians, the developed systems should be explainable and interpretable to a certain degree. For example, case-based reasoning[44] or attention-based modelling[45] are a way forward.

Lastly, real-time process should be achieved before the developed systems can make any real impact. This is because a collection of biopsy takes place only during the time of endoscopy. If those suspicious regions are overlooked, the patients in concern will miss the chances of correct diagnosis and appropriate treatment.

CONCLUSION

In conclusion, this paper overviews the current development of AI-based computer assisted systems for supporting early diagnosis of oesophageal cancers and proposes several future directions, expediting the clinical implementation and hence benefiting both patient and clinician communities.

REFERENCES

1 **von Bartheld CS**, Bahney J, Herculano-Houzel S. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *J Comp Neurol* 2016; **524**: 3865-3895 [PMID: 27187682 DOI: 10.1002/cne.24040]

2 **Silver D**, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T, Hassabis D. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016; **529**: 484-489 [PMID: 26819042 DOI: 10.1038/nature16961]

3 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

4 **Krizhevsky A**, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017; **60**: 84-90 [DOI: 10.1145/3065386]

5 **Bradley R**, Tesla Auopilot. 2016 Feb 23 [cited 20 April 2021]. In: MIT Technology Review. [Internet]. Available from: https://www.technologyreview.com/technology/tesla-autopilot/

6 **Gibney E**. Self-taught AI is best yet at strategy game Go, Nature News 2017 [DOI: 10.1038/nature.2017.22858]

7 **Muehlematter UJ**, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015-20): a comparative analysis. *Lancet Digit Health* 2021; **3**: e195-e203 [PMID: 33478929 DOI: 10.1016/S2589-7500(20)30292-2]

8 **US Food and Drug Administration**. Proposed regulatory framework for modifications to artificial intel-ligence/machine learning (AI/ML)-based software as a medical device (SaMD). 2019. [cited 5 February 2021]. [Internet]. https://www.fda.gov/media/122535/download

9 **Pereira S**, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans Med Imaging* 2016; **35**: 1240-1251 [PMID: 26960222 DOI: 10.1109/TMI.2016.2538465]

10 **Gao X**, James-Reynolds C, Currie E. Analysis TB severity with an enhanced deep residual learning depth-resnet. In: CLEF2018 Working Notes. CEUR Workshop Proceedings, Avignon, France. 2018. Available from: http://ceur-ws.org/Vol-2125/paper_175.pdf

11 **Ali S**, Zhou F, Braden B, Bailey A, Yang S, Cheng G, Zhang P, Li X, Kayser M, Soberanis-Mukul RD, Albarqouni S, Wang X, Wang C, Watanabe S, Oksuz I, Ning Q, Khan MA, Gao XW, Realdon S, Loshchenov M, Schnabel JA, East JE, Wagnieres G, Loschenov VB, Grisan E, Daul C, Blondel W, Rittscher J. An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy. *Sci Rep* 2020; **10**: 2748 [PMID: 32066744 DOI: 10.1038/s41598-020-59413-5]

12 **Ali S**, Dmitrieva M, Ghatwary N, Bano S, Polat G, Temizel A, Krenzer A, Hekalo A, Guo YB, Matuszewski B, Gridach M, Voiculescu I, Yoganand V, Chavan A, Raj A, Nguyen NT, Tran DQ, Huynh LD, Boutry N, Rezvy S, Chen H, Choi YH, Subramanian A, Balasubramanian V, Gao XW, Hu H, Liao Y, Stoyanov D, Daul C, Realdon S, Cannizzaro R, Lamarque D, Tran-Nguyen T, Bailey

A, Braden B, East JE, Rittscher J. Deep learning for detection and segmentation of artefact and disease instances in gastrointestinal endoscopy. *Med Image Anal* 2021; **70**: 102002 [PMID: 33657508 DOI: 10.1016/j.media.2021.102002]

13 **Gao X**, Li W, Loomes M, Wang L. A fused deep learning architecture for viewpoint classification of echocardiography. *Informat Fusion* 2017; **36**: 103-113 [DOI: 10.1016/j.inffus.2016.11.007]

14 **Gao XW**, Hui R, Tian Z. Classification of CT brain images based on deep learning networks. *Comput Methods Programs Biomed* 2017; **138**: 49-56 [PMID: 27886714 DOI: 10.1016/j.cmpb.2016.10.007]

15 **Gao**, X, Carl James-Reynolds, Edward Currie. Analysis of Tuberculosis Severity Levels From CT Pulmonary Images Based on Enhanced Residual Deep Learning Architecture. *Neuro Computing* 2021; 392: 233-244 [DOI: 10.1016/j.neucom.2018.12.086]

16 **Gao X**, Quan Y. Prediction of multi-drug resistant TB from CT pulmonary Images based on deep learning techniques. *Mol Pharmaceutics* 2018; **15**: 4326-4335 [DOI: 10.1021/acs.molpharmaceut.7b00875]

17 **Simonyan K**, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2014 Sep 4 [cited 20 April 2021]. In: arXiv [Internet]. Available from: https://arxiv.org/abs/1409.1556

18 **Rezvy S**, Zebin T, Braden B, Pang W, Taylor S, Gao X. Transfer Learning For Endoscopy Disease Detection & Segmentation With Mask-RCNN Benchmark Architecture. Proceedings of the 2nd International Workshop and Challenge on Computer Vision in Endoscopy (EndoCV2020). 3 Apr 2020. Available from: http://ceur-ws.org/Vol-2595/endoCV2020_paper_id_17.pdf

19 **Wang J**, Bao Y, Wen Y, Lu H, Luo H, Xiang Y, Li X, Liu C, Qian D. Prior-Attention Residual Learning for More Discriminative COVID-19 Screening in CT Images. *IEEE Trans Med Imaging* 2020; **39**: 2572-2583 [PMID: 32730210 DOI: 10.1109/TMI.2020.2994908]

20 **Ouyang X**, Huo J, Xia L, Shan F, Liu J, Mo Z, Yan F, Ding Z, Yang Q, Song B, Shi F, Yuan H, Wei Y, Cao X, Gao Y, Wu D, Wang Q, Shen D. Dual-Sampling Attention Network for Diagnosis of COVID-19 From Community Acquired Pneumonia. *IEEE Trans Med Imaging* 2020; **39**: 2595-2605 [PMID: 32730212 DOI: 10.1109/TMI.2020.2995508]

21 **Oh Y**, Park S, Ye JC. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Trans Med Imaging* 2020; **39**: 2688-2700 [PMID: 32396075 DOI: 10.1109/TMI.2020.2993291]

22 **Wang L**, Lin Z, Wong A, COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images, Sci Rep 2020; 10: 19549 [DOI: 10. Wang L, Lin Z, Wong A, COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images, Sci Rep 2020; 10: 19549 [DOI: 10.1038/s41598-020-76550-z]

23 **Waheed A**, Goyal M. , Gupta D, Khanna A, Al-Turjman F, Pinheiro PR. *IEEE Access* 2020; **8**: 91916-91923 [DOI: 10.1109/access.2020.2994762]

24 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]

25 **Pennathur A**, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. *Lancet* 2013; **381**: 400-412 [PMID: 23374478 DOI: 10.1016/S0140-6736(12)60643-6]

26 **Arnold M**, Laversanne M, Brown LM, Devesa SS, Bray F. Predicting the Future Burden of Esophageal Cancer by Histological Subtype: International Trends in Incidence up to 2030. *Am J Gastroenterol* 2017; **112**: 1247-1255 [PMID: 28585555 DOI: 10.1038/ajg.2017.155]

27 **Arnold M**, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* 2015; **64**: 381-387 [PMID: 25320104 DOI: 10.1136/gutjnl-2014-308124]

28 **Siegel R**, Desantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin* 2014; **64**: 104-117 [PMID: 24639052 DOI: 10.3322/caac.21220]

29 **Shimizu Y**, Tsukagoshi H, Fujita M, Hosokawa M, Kato M, Asaka M. Long-term outcome after endoscopic mucosal resection in patients with esophageal squamous cell carcinoma invading the muscularis mucosae or deeper. *Gastrointest Endosc* **56**: 387-390 [DOI: 10.1067/mge.2002.127100]

30 **Trivedi PJ**, Braden B. Indications, stains and techniques in chromoendoscopy. *QJM* 2013; **106**: 117-131 [PMID: 23097386 DOI: 10.1093/qjmed/hcs186]

31 **van der Sommen F**, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, Bergman JJ, de With PH, Schoon EJ. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. *Endoscopy* 2016; **48**: 617-624 [PMID: 27100718 DOI: 10.1055/s-0042-105284]

32 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929.e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

33 **Everson M**, Herrera L, Li W, Luengo IM, Ahmad O, Banks M, Magee C, Alzoubaidi D, Hsu HM, Graham D, Vercauteren T, Lovat L, Ourselin S, Kashin S, Wang HP, Wang WL, Haidry RJ. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. *United European Gastroenterol J* 2019; **7**: 297-306 [PMID: 31080614 DOI: 10.1177/2050640618821800]

34 **Horie Y**, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019; **89**: 25-32 [PMID: 30120958 DOI: 10.1016/j.gie.2018.07.037]

35 **Ghatwary N**, Zolgharni M, Ye X. Early esophageal adenocarcinoma detection using deep learning methods. *Int J Comput Assist Radiol Surg* 2019; **14**: 611-621 [PMID: 30666547 DOI: 10.1007/s11548-019-01914-4]

36 W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC. SSD: Single Shot MultiBox Detector. *Eur Conf Comput Vision* 2016; 21-37 [DOI: 10.1007/978-3-319-46448-0_2]

37 **Girshick R**. Fast R-CNN. *Int Conf Comput Vision* 2015; 1440-1448 [DOI: 10.1109/iccv.2015.169]

38 **Ohmori M**, Ishihara R, Aoyama K, Nakagawa K, Iwagami H, Matsuura N, Shichijo S, Yamamoto K, Nagaike K, Nakahara M, Inoue T, Aoi K, Okada H, Tada T. Endoscopic detection and differentiation of esophageal lesions using a deep neural network. *Gastrointest Endosc* 2020; **91**: 301-309.e1 [PMID: 31585124 DOI: 10.1016/j.gie.2019.09.034]

39 **Gao XW**, Braden B, Taylor S, Pang W. Towards Real-Time Detection of Squamous Pre-Cancers from Oesophageal Endoscopic Videos. *Int Conf Comput Vision* 2019; 1606-1612 [DOI: 10.1109/icmla.2019.00264]

40 **Redmon J**, Farhadi A. YOLOv3: An incremental improvement. 2018 Apr 8 [cited 20 April 2021]. In: arXiv [Internet]. Available from https://arxiv.org/abs/1804.02767

41 **He K**, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. *Int Conf Comput Vision* 2017; 2980-2988 [DOI: 10.1109/iccv.2017.322]

42 **Guo L**, Xiao X, Wu C, Zeng X, Zhang Y, Du J, Bai S, Xie J, Zhang Z, Li Y, Wang X, Cheung O, Sharma M, Liu J, Hu B. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest Endosc* 2020; **91**: 41-51 [PMID: 31445040 DOI: 10.1016/j.gie.2019.08.018]

43 **Bolya D**, Zhou C, Xiao F, and Lee YJ. YOLACT: Real-time Instance Segmentation. *Int Conf Comput Vision* 2019; 9156-916 [DOI: 10.1109/iccv.2019.00925]

44 **Gao XW**, Braden B, Zhang L, Taylor S, Pang W, Petridis M. Case-based Reasoning of a Deep Learning Network for Prediction of Early Stage of Oesophageal Cancer, UKCBR 2019, AI-2019, Cambridge. 2019. Available from: http://www.expertupdate.org/papers/20-1/2019_paper_1.pdf

45 **Tomita N**, Abdollahi B, Wei J, Ren B, Suriawinata A, Hassanpour S. Attention-Based Deep Neural Networks for Detection of Cancerous and Precancerous Esophagus Tissue on Histopathological Slides. *JAMA Netw Open* 2019; **2**: e1914645 [PMID: 31693124 DOI: 10.1001/jamanetworkopen.2019.14645]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 127-135

DOI: 10.37126/aige.v2.i4.127

ISSN 2689-7164 (online)

MINIREVIEWS

Deep learning applied to the imaging diagnosis of hepatocellular carcinoma

Vinícius Remus Ballotin, Lucas Goldmann Bigarella, John Soldera, Jonathan Soldera

ORCID number: Vinícius Remus Ballotin 0000-0002-2659-2249; Lucas Goldmann Bigarella 0000-0001-8087-0070; John Soldera 0000-0002-4000-903X; Jonathan Soldera 0000-0001-6055-4783.

Author contributions: All authors contributed to study concept and design, to drafting of the manuscript and to critical revision of the manuscript for important intellectual content.

Conflict-of-interest statement: The authors have no conflict of interest to disclose.

Manuscript source: Invited manuscript

Specialty type: Gastroenterology

Vinícius Remus Ballotin, Lucas Goldmann Bigarella, School of Medicine, Universidade de Caxias do Sul, Caxias do Sul 95070-560, RS, Brazil

John Soldera, Computer Science, Federal Institute of Education, Science and Technology Farroupilha, Santo Ângelo 98806-700, RS, Brazil

Jonathan Soldera, Clinical Gastroenterology, Universidade de Caxias do Sul, Caxias do Sul 95070-560, RS, Brazil

Corresponding author: Jonathan Soldera, MD, MSc, Associate Professor, Staff Physician, Clinical Gastroenterology, Universidade de Caxias do Sul, Rua Francisco Getúlio Vargas 1130, Caxias do Sul 95070-560, RS, Brazil. jonathansoldera@gmail.com

Abstract

Each year, hepatocellular carcinoma is diagnosed in more than half a million people worldwide. It is the fifth most common cancer in men and the seventh most common cancer in women. Its diagnosis is currently made using imaging techniques, such as computed tomography and magnetic resonance imaging. For most cirrhotic patients, these methods are enough for diagnosis, foregoing the necessity of a liver biopsy. In order to improve outcomes and bypass obstacles, many companies and clinical centers have been trying to develop deep learning systems that could be able to diagnose and classify liver nodules in the cirrhotic liver, in which the neural networks are one of the most efficient approaches to accurately diagnose liver nodules. Despite the advances in deep learning systems for the diagnosis of imaging techniques, there are many issues that need better development in order to make such technologies more useful in daily practice.

Key Words: Hepatocellular carcinoma; Cirrhosis; Machine learning; Artificial intelligence

Core Tip: Hepatocellular carcinoma is diagnosed using imaging techniques, such as computed tomography and magnetic resonance imaging. In order to improve outcomes and bypass obstacles, many companies and clinical centers have been trying to develop deep learning systems that could be able to diagnose and classify liver nodules in the

and hepatology

Country/Territory of origin: Brazil

Peer-review report's scientific quality classification

Grade A (Excellent): A
Grade B (Very good): 0
Grade C (Good): C, C, C
Grade D (Fair): 0
Grade E (Poor): E

Received: April 21, 2021
Peer-review started: April 21, 2021
First decision: May 19, 2021
Revised: June 5, 2021
Accepted: July 19, 2021
Article in press: July 19, 2021
Published online: August 28, 2021

P-Reviewer: Hameed MM, Raut V, Zhang L, Zhu YY
S-Editor: Fan JR
L-Editor: Filipodia
P-Editor: Ma YJ

cirrhotic liver. Neural networks have become one of the most efficient approaches to accurately diagnose liver nodules using deep learning systems. Therefore, with the improvement of these techniques in the long term, they could be applicable in daily practice, modifying outcomes.

Citation: Ballotin VR, Bigarella LG, Soldera J, Soldera J. Deep learning applied to the imaging diagnosis of hepatocellular carcinoma. *Artif Intell Gastrointest Endosc* 2021; 2(4): 127-135
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/127.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.127

INTRODUCTION

Each year, hepatocellular carcinoma (HCC) is diagnosed in more than half a million people worldwide, and it is the fifth most common cancer in men and the seventh most common cancer in women[1]. The greatest burden of this disease is in developing countries, such as Southeast Asia and Sub-Saharan Africa, where hepatitis B is endemic[2,3].

The incidence of HCC has been rising, unlike many other types of neoplasms[4]. This is expected to change, as the worldwide incidence of viral hepatitis B and C is expected to subdue in the next generation *via* vaccination and treatment, respectively. Nevertheless, the acute rise in the prevalence of nonalcoholic steatohepatitis in the last couple of decades might become a key risk factor for HCC and could become solely responsible for sustaining its incidence, both in the Western and Eastern population[5,6].

Therefore, understanding the diagnostic and therapeutic approaches to this disease is essential, especially if we keep in mind the quintessential basics of prevention and early detection to improve results[7,8].

DIAGNOSIS OF HCC

HCC diagnosis is currently made using imaging techniques, such as computed tomography and magnetic resonance imaging (MRI). For most cirrhotic patients, these methods are enough for diagnosis, foregoing the necessity of a liver biopsy[9-11]. Nevertheless, the precise diagnosis of a liver nodule *via* imaging techniques is a rather challenging task, requiring a highly trained and specialized multidisciplinary team of radiologists, hepatologists and oncologists.

In order to facilitate communication between professionals of such a team, a system for reporting imaging of liver nodules has been developed and adopted world-wide–the Liver Imaging Reporting And Data System (LI-RADS)[12]. The LI-RADS classification[13] can be found in Table 1. Although this was an attempt into standard-ization, a high discordance rate among radiologists has been described[14]. Inter-rater reliability has varied greatly in studies, with Cohen's kappa coefficients ranging from 0.35 to 0.73[15-19]. This is expected, since this classification requires high-quality imaging and radiologists with vast experience[19,20]. Another very important argument is that where HCC incidence is higher (developing countries), highly specialized radiologists are scarcest despite a high volume of patients[21]. In order to improve outcomes and bypass these obstacles, many companies and clinical centers have been trying to develop deep learning systems (DLS) intended to accurately diagnose liver nodules in the cirrhotic liver[22].

DLS AND HCC

There are many DLS approaches available in the literature, where neural networks are gaining much attention currently as one of the best approaches to accurately diagnose liver nodules. Particularly, a DLS based on convolutional neural networks (CNN) could achieve such capacities after machine learning (ML) by using examples of images with and without the disease in question[8]. Unlike other DLS, CNN does not

Table 1 Liver imaging reporting and data system classification[13]

Category	Description
LR-1	Definitely benign
LR-2	Probably benign
LR-3	Intermediate probability of HCC
LR-4	High probability of HCC, not 100%
LR-5	Definitely HCC
LR-5V	Definite venous invasion regardless of other imaging features
LR treated	LR-5 lesion status post-locoregional treatment
LR-M	Non-HCC malignancies that may occur in cirrhosis: metastases, lymphoma, cholangiocarcinoma, PTLD

HCC: Hepatocellular carcinoma; PTLD: Post-transplant lymphoproliferative disorder.

demand a clear definition of the lesion in order to interpret the images[23], which might lead to discovery of additional differential characteristics that are not currently known by radiologists[24]. Table 2 summarizes the main characteristics about the studies in diagnosis of liver tumors with images and clinical data using DLS.

There are several DLS applied in the recognition of image patterns[25,26], from which CNN-based approaches have achieved the highest performance[25]. While conventional deep learning algorithms require specific features to be extracted from images before the learning process, the application of CNNs requires rather a simpler feature representation based on the original image pixel intensities, also allowing to use all available image information in the learning process[27]. Moreover, CNNs can process extracted image features by several convolution filters, which allow analysis of the image at different granularities. Therefore, CNN is one of the most advanced techniques for artificial intelligence[25], which has been implemented with success for imaging and clinical interpretation in many medical fields. For example, CNN has been validated to identify liver tumors[28], the prognosis of esophageal variceal bleeding in cirrhotic patients[29], to predict the mortality of liver transplantation[30,31], to predict the prognosis of HCC[32-37] *Helicobacter pylori* infection[38], colonic polyps[39], to help classify mammary cancer, head and neck cancer and gliomas[36] and to focal liver disease detection[40].

In the topic of liver tumors, many studies have shown that CNNs performed the same or better when compared to experienced radiologists. Hamm *et al*[8] developed and validated a CNN that classified six types of common hepatic lesions on multiphasic MRI, achieving better sensitivity and specificity when compared to board-certified radiologists[8]. Nevertheless, this study was developed in only one center, using local and typical images, with no external validation. In a follow-up to this study, Wang *et al*[41] used a pre-trained CNN in a model-agonistic approach capable of distinguishing among several types of lesions and developed a post-hoc algorithm with the purpose of standardizing the lesion features used in the diagnosis. Such a tool could interact with other standardized scales, such as LI-RADS, validating auxiliary resources and improving clinical practicality[41]. This study found a sensitivity of 82.9% for adequate identification of imaging characteristics when analyzing lesions from a databank. It is expected that this type of DLS that can be transparent regarding its steps towards the diagnosis will have better clinical acceptance.

Yamashita *et al*[14] developed a DLS applied to diagnose liver carcinoma by using two CNNs: a pre-trained network with an input of triple-phase images (trained with transfer learning from other CNNs) and a custom-made network with an input of quadruple-phase images (trained from scratch from internal data)[14]. However, by using external data from other pre-trained CNNs, Zech *et al*[42] showed that the performance of the DLS worsened when compared to CNNs trained with internal data, showing that it is not still proved that CNNs trained on X-rays from one hospital or one group of hospitals will work equally well at different hospitals. This has also been demonstrated for the detection of pneumonia in chest X-rays, where CNN performed worse when exposed to external data with a wide range of diseases and radiological findings[42]. Besides, such CNNs could be used for the determination of LI-RADS category, which has been shown to be possible[14], even from a small data set. Nevertheless, external validation seems to be a major obstacle for the dissem-

Table 2 Main characteristics of the studies that evaluate deep learning for liver tumor diagnosis throughout images or clinical data

Ref.	Country	Deep learning method	Accuracy	Sensitivity	Specificity	AUROC	DLS performance compared	Multicenter validation	Conclusion
Hamm *et al*[8], 2019	United States	Proof-of-concept validation CNN	92%	92%	98%	0.992	Better than radiologists	Not done	DLS was feasibility for classifying lesions with typical imaging features from six common hepatic lesion types
Yamashita *et al*[14], 2020	United States	CNN architectures: custom-made network and transfer learning-based network	60.4%	NA	NA	LR-1/2: 0.85. LR-3: 0.90. LR-4: 0.63. LR-5: 0.82	Transfer learning model was better	Performed	There is a feasibility of CNN for assigning LI-RADS categories from a relatively small dataset but highlights the challenges of model development and validation
Shi *et al* [23], 2020	China	Three CDNs	Model-A: 83.3%, B: 81.1%, C: 85.6%	NA	NA	Model-A: 0.925; B: 0.862; C: 0.920	Three model compared, A and C with better results	Not done	Three-phase CT protocol without precontrast showed similar diagnosis accuracy as four-phase protocol in differentiating HCC. It can reduce the radiation dose
Yasaka *et al*[25], 2018	Japan	CNN	84%	Category[1]: A: 71%; B: 33%; C: 94%; D: 90%; E: 100%	NA	0.92	Not applicable	Not done	Deep learning with CNN showed high diagnostic performance in differentiation of liver masses at dynamic CT
Trivizakis *et al*[28], 2019	Greece	3D and 2D CNN	83%	93%	67%	0.80	Superior compared with 2D CNN model	Not done	3D CNN architecture can bring significant benefit in DW-MRI liver discrimination and potentially in numerous other tissue classification problems based on tomographic data, especially in size-limited, disease specific clinical datasets
Wang *et al* [41], 2019	United States	Proof-of-concept "interpretable" CNN	88%	82.9%	NA	NA	Not applicable	Not done	This interpretable deep learning system demonstrates proof of principle for illuminating portions of a pre-trained deep neural network's decision-making, by analyzing inner layers and automatically describing features contributing to predictions
Frid-Adar *et al*[45], 2018	Israel	GANs	Classic data: 78.6%. Synthetic data: 85.7%	Classic data: 78.6%. Synthetic data: 85.7%	Classic data: 88.4%. Synthetic data: 92.4%	NA	Synthetic data augmentation is better than classic data augmentation	Not done	This approach to synthetic data augmentation can generalize to other medical classification applications and thus support radiologists' efforts to improve diagnosis
Wang *et al*	Japan	CNN with	NA	NA	NA	Clinical	Combined	Not done	The AUC of the

[47], 2019		clinical data				model: 0.723. Model: A: 0.788; B: 0.805; C: 0825.	model C present with better results		combined model is about 0.825, which is much better than the models using clinical data only or CT image only
Sato *et al* [48], 2019	Japan	Fully connected neural network with 4 layers of neurons using only biomarkers, gradient boosting (non-linear model) and others	DLS: 83.54%. Gradient boosting: 87.34%	Gradient boosting: 93.27%	Gradient boosting: 75.93%	DLS: 0.884. Gradient boosting: 0.940	Deep learning was not the optimal classifier in the current study	Not done	The gradient boosting model reduced the misclassification rate by about half compared with a single tumor marker. The model can be applied to various kinds of data and thus could potentially become a translational mechanism between academic research and clinical practice
Naeem *et al*[49], 2020	Pakistan	MLP, SVM, RF, and J48 using ten-fold cross-validation	MLP: 99%	NA	NA	MLP: 0.983. SVM: 0.966. RF: 0.964. J48: 0.959	MLP model present with better results	Radiopaedia dataset	Our proposed system has the capability to verify the results on different MRI and CT scan databases, which could help radiologists to diagnose liver tumors

[1]Five categories: A: Classic hepatocellular carcinomas; B: Malignant liver tumors other than classic and early hepatocellular carcinomas; C: Indeterminate masses or mass like lesions (including early hepatocellular carcinomas and dysplastic nodules) and rare benign liver masses other than hemangiomas and cysts; D: Hemangiomas; E: Cysts. AUC: Area under the curve; AUROC: Area under the receiver operating characteristic curve; CDNs: Convolutional dense networks CNN: Convolutional neural network; CT: Computed tomography; DLS: Deep learning system; DW-MRI: Diffusion weighted magnetic resonance imaging; GANs: Generative adversarial networks; HCC: Hepatocellular carcinoma; LI-RADS: Liver Imaging Reporting and Data System; LR: LI-RADS; MLP: Multiplayer perceptron; MRI: Magnetic resonance imaging; NA: Not available; RF: Random forest; SVM: Support vector machine.

ination of ML tools. There are many devices that produce images, and there are many ways to store data from these exams.

When compared to other DLS, another advantage of the use of CNNs is that it can improve the diagnosis by using less images for ML, reducing the time of exam and the amount of exposure to radiation[23,43,44]. Moreover, by generating additional training samples through data augmentation, the liver lesion classification sensitivity and accuracy are enhanced whilst less images are required in the ML process[45]. Moreover, the sensitivity, specificity, and accuracy can be manually calculated with the confusion matrix. In Table 3, we compare the best ML algorithms for classification [46].

A DLS has been proposed for the prediction of HCC recurrence, using data from computed tomography combined with clinical information[47]. The triple layer model including imaging studies, clinical data and a filtering of this data has had the better performance, with an area under the receiver operating characteristic curve (AUROC) of 0.825. This is way more precise than the current tools are. Furthermore, Sato *et al*[48] proposed a ML model for predicting HCC using data obtained during clinical practice [48]. The AUROC of the optimal hyperparameter, gradient boosting model, involving multiple laboratories and tumor markets was 0.940. However, when compared with single tumor markers the AUROC to the prediction of HCC for alpha-fetoprotein, des-gamma-carboxy prothrombin and alpha-fetoprotein-L3 were 0.766, 0.644 and 0.683, respectively. Accordingly, a combination of multiple data can provide a reliable diagnostic tool.

A preliminary study has attempted to diagnose liver masses using a CNN without the aid of a radiologist, achieving a high accuracy to differentiate HCC from benign liver masses, achieving an AUROC of 0.92[25]. In another study, a CNN was designed to differentiate HCC from metastatic liver masses on MRI, but this time the DLS used a 3-D representation, with higher accuracy (83.0% of the 3-D model *vs* 65.2% of the 2-D model)[28]. Nevertheless, the authors stressed that more studies with larger databanks are needed to verify the accuracy of this method. Besides that, Naeem *et al*[49] performed a hybrid-feature analysis between computed tomography scans and MRI for differentiation of liver tumors using DLS. The accuracy of multilayer perceptron

Table 3 Best machine learning algorithms for classification[36]

Algorithm	Pros	Cons
Naïve Bayes Classifier	Simple, easy and fast. Not sensitive to irrelevant features. Works great in practice. Needs less training data. For both multi-class and binary classification. Works with continuous and discrete data	Accepts every feature as independent. This is not always the truth
Decision Trees	Easy to understand. Easy to generate rules. There are almost no hyperparameters to be tuned. Complex decision tree models can be significantly simplified by its visualizations	Might suffer from overfitting. Does not easily work with nonnumerical data. Low prediction accuracy for a dataset in comparison with other algorithms. When there are many class labels, calculations can be complex
Support Vector Machines	Fast algorithm. Effective in high dimensional spaces. Great accuracy. Power and flexibility from kernels. Works very well with a clear margin of separation. Many applications	Does not perform well with large data sets. Not so simple to program. Does not perform so well when the data comes with more noise *i.e.* target classes are overlapping
Random Forest Classifier	The overfitting problem does not exist. Can be used for feature engineering *i.e.* for identifying the most important features among all available features in the training dataset. Runs very well on large databases. Extremely flexible and have very high accuracy. No need for preparation of the input data	Complexity. Requires a lot of computational resources. Time-consuming. Need to choose the number of trees
KNN Algorithm	Simple to understand and easy to implement. Zero to little training time. Works easily with multi-class data sets. Has good predictive power. Does well in practice	Computationally expensive testing phase. Can have skewed class distributions. The accuracy can be decreased when it comes to high-dimension data. Needs to define a value for the parameter k

KNN: K-nearest neighbors.

model for hepatoblastoma, cyst, hemangioma, hepatocellular adenoma, HCC and metastasis were 99.67%, 99.33%, 98.33%, 99.67%, 97.33% and 99.67% respectively[49]. This method can be helpful to reduce human error.

Therefore, despite the advances in DLS for the diagnosis of imaging techniques, there are many points that need better development in order to become useful and common tools in daily practice. These techniques currently require comparison with trained radiologists and the application for many databanks with atypical images to achieve better results and the use of less radiation for HCC diagnosis.

We previously presented several DLS applied to liver nodule diagnosis; however, they are not able to segment the nodule from the liver in the analyzed images. Moreover, automatic nodule segmentation in an image is a challenging task since this kind of lesion may show a high variability in shape, appearance and localization and is dependent on the equipment, contrast, lesion type, lesion stage and so on[50].

There are some liver nodule segmentation methods available in the literature, and in one of them[50] a fully convolutional network architecture was adopted to determine an approximation for where the nodule was located on the image. This CNN works on four resolution levels, learning local and global image features. The final nodule segmentation is obtained by using post-processing techniques and a random forest classifier, achieving a quality comparable to a human expert.

However, this method uses hand-crafted features that need the supervision of an expert. There are also automatic approaches that can segment the nodule[51], where a CNN is used for ML. To refine the segmentation results, this method applies conditional random fields to eliminate the false segmentation points in the segmentation results, improving accuracy. However, liver nodule segmentation in general still needs improvements to achieve a better accuracy and practical applicability. Furthermore, it is necessary for more research effort in DLS to at the same time detect the tumor in the liver and segment it on the image.

CONCLUSION

In conclusion, the goal of statistical methods is to achieve conclusions about a population from data that are collected from a representative sample of that population, such as linear and logistic regression. Therefore, the objective is to comprehend the associations among variables. However, as reported by Sidey-Gibbons and Sidey-Gibbons[36], the primary concern about DLS is an accurate prediction. Moreover, explaining the relationship between predictors and outcomes when the relationship is non-linear is difficult. However, in several DLS as improving navigation, translating documents or recognizing objects in videos, understanding the relationship between features and outcomes is less important[46]. In summary,

enhancement of DLS features will allow more accurate diagnosis in the medical field. For future research, we recommend to test deep learning methods in other datasets (*e.g.*, other hospitals), develop an easy usable interface and introduce the tool in daily medical practice.

REFERENCES

1 **Soldera J**, Balbinot SS, Balbinot RA, Cavalcanti AG. Diagnostic and Therapeutic Approaches to Hepatocellular Carcinoma: Understanding the Barcelona Clínic Liver Cancer Protocol. *Clin Med Insights Gastroenterol* 2016; **9**: 67-71 [PMID: 27812296 DOI: 10.4137/CGast.S30190]

2 **El-Serag HB**, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 2007; **132**: 2557-2576 [PMID: 17570226 DOI: 10.1053/j.gastro.2007.04.061]

3 **Bosetti C**, Turati F, La Vecchia C. Hepatocellular carcinoma epidemiology. *Best Pract Res Clin Gastroenterol* 2014; **28**: 753-770 [PMID: 25260306 DOI: 10.1016/j.bpg.2014.08.007]

4 **Siegel RL**, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020; **70**: 7-30 [PMID: 31912902 DOI: 10.3322/caac.21590]

5 **Onzi G**, Moretti F, Balbinot SS, Balbinot RA, Soldera J. Hepatocellular carcinoma in non-alcoholic fatty liver disease with and without cirrhosis. *Hepatoma Res* 2019; **5** [DOI: 10.20517/2394-5079.2018.114]

6 **Margini C**, Dufour JF. The story of HCC in NAFLD: from epidemiology, across pathogenesis, to prevention and treatment. *Liver Int* 2016; **36**: 317-324 [PMID: 26601627 DOI: 10.1111/liv.13031]

7 **Soldera J**, Balbinot SS, Balbinot RA, Furlan RG, Terres AZ. Advanced hepatocellular carcinoma. *Austin J Gastroenterol* 2017; **4**: 1088 Available from: URL: https://austinpublishinggroup.com/gastroenterology/fulltext/ajg-v4-id1088.php

8 **Hamm CA**, Wang CJ, Savic LJ, Ferrante M, Schobert I, Schlachter T, Lin M, Duncan JS, Weinreb JC, Chapiro J, Letzen B. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. *Eur Radiol* 2019; **29**: 3338-3347 [PMID: 31016442 DOI: 10.1007/s00330-019-06205-9]

9 **Marrero JA**, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, Roberts LR, Heimbach JK. Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases. *Hepatology* 2018; **68**: 723-750 [PMID: 29624699 DOI: 10.1002/hep.29913]

10 **European Association for the Study of the Liver**. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol* 2018; **69**: 182-236 [PMID: 29628281 DOI: 10.1016/j.jhep.2018.03.019]

11 **Chagas AL**, Mattos AA, Carrilho FJ, Bittencourt PL; Members of the Panel of the 2nd Consensus of the Brazilian Society of Hepatology on the Diagnosis and Management of Hepatocellular Carcinoma, Vezozzo DCP, Horvat N, Rocha MS, Alves VAF, Coral GP, Alvares-DA-Silva MR, Barros FMDR, Menezes MR, Monsignore LM, Coelho FF, Silva RFD, Silva RCMA, Boin IFSF, D Albuquerque LAC, Garcia JHP, Felga GEG, Moreira AM, Braghiroli MIFM, Hoff PMG, Mello VB, Dottori MF, Branco TP, Schiavon LL, Costa TFA. Brazilian society of hepatology updated recommendations for diagnosis and treatment of hepatocellular carcinoma. *Arq Gastroenterol* 2020; **57**: 1-20 [PMID: 32294682 DOI: 10.1590/S0004-2803.202000000-20]

12 **Elsayes KM**, Kielar AZ, Chernyak V, Morshid A, Furlan A, Masch WR, Marks RM, Kamaya A, Do RKG, Kono Y, Fowler KJ, Tang A, Bashir MR, Hecht EM, Jambhekar K, Lyshchik A, Rodgers SK, Heiken JP, Kohli M, Fetzer DT, Wilson SR, Kassam Z, Mendiratta-Lala M, Singal AG, Lim CS, Cruite I, Lee J, Ash R, Mitchell DG, McInnes MDF, Sirlin CB. LI-RADS: a conceptual and historical review from its beginning to its recent integration into AASLD clinical practice guidance. *J Hepatocell Carcinoma* 2019; **6**: 49-69 [PMID: 30788336 DOI: 10.2147/JHC.S186239]

13 **Elsayes KM**, Kielar AZ, Agrons MM, Szklaruk J, Tang A, Bashir MR, Mitchell DG, Do RK, Fowler KJ, Chernyak V, Sirlin CB. Liver Imaging Reporting and Data System: an expert consensus statement. *J Hepatocell Carcinoma* 2017; **4**: 29-39 [PMID: 28255543 DOI: 10.2147/JHC.S125396]

14 **Yamashita R**, Mittendorf A, Zhu Z, Fowler KJ, Santillan CS, Sirlin CB, Bashir MR, Do RKG. Deep convolutional neural network applied to the liver imaging reporting and data system (LI-RADS) version 2014 category classification: a pilot study. *Abdom Radiol (NY)* 2020; **45**: 24-35 [PMID: 31696269 DOI: 10.1007/s00261-019-02306-7]

15 **Fowler KJ**, Tang A, Santillan C, Bhargavan-Chatfield M, Heiken J, Jha RC, Weinreb J, Hussain H, Mitchell DG, Bashir MR, Costa EAC, Cunha GM, Coombs L, Wolfson T, Gamst AC, Brancatelli G, Yeh B, Sirlin CB. Interreader Reliability of LI-RADS Version 2014 Algorithm and Imaging Features for Diagnosis of Hepatocellular Carcinoma: A Large International Multireader Study. *Radiology* 2018; **286**: 173-185 [PMID: 29091751 DOI: 10.1148/radiol.2017170376]

16 **Schellhaas B**, Hammon M, Strobel D, Pfeifer L, Kielisch C, Goertz RS, Cavallaro A, Janka R, Neurath MF, Uder M, Seuss H. Interobserver and intermodality agreement of standardized algorithms for non-invasive diagnosis of hepatocellular carcinoma in high-risk patients: CEUS-LI-RADS versus MRI-LI-RADS. *Eur Radiol* 2018; **28**: 4254-4264 [PMID: 29675659 DOI: 10.1007/s00330-018-5379-1]

17 **Barth BK**, Donati OF, Fischer MA, Ulbrich EJ, Karlo CA, Becker A, Seifert B, Reiner CS.

Reliability, Validity, and Reader Acceptance of LI-RADS-An In-depth Analysis. *Acad Radiol* 2016; **23**: 1145-1153 [PMID: 27174029 DOI: 10.1016/j.acra.2016.03.014]

18 **Davenport MS**, Khalatbari S, Liu PS, Maturen KE, Kaza RK, Wasnik AP, Al-Hawary MM, Glazer DI, Stein EB, Patel J, Somashekar DK, Viglianti BL, Hussain HK. Repeatability of diagnostic features and scoring systems for hepatocellular carcinoma by using MR imaging. *Radiology* 2014; **272**: 132-142 [PMID: 24555636 DOI: 10.1148/radiol.14131963]

19 **Blachar A**, Federle MP, Ferris JV, Lacomis JM, Waltz JS, Armfield DR, Chu G, Almusa O, Grazioli L, Balzano E, Li W. Radiologists' performance in the diagnosis of liver tumors with central scars by using specific CT criteria. *Radiology* 2002; **223**: 532-539 [PMID: 11997564 DOI: 10.1148/radiol.2232010801]

20 **Dubus L**, Gayet M, Zappa M, Abaleo L, De Cooman A, Orieux G, Vilgrain V. Comparison of semi-automated and manual methods to measure the volume of liver tumours on MDCT images. *Eur Radiol* 2011; **21**: 996-1003 [PMID: 21132500 DOI: 10.1007/s00330-010-2013-2]

21 **Nayak A**, Baidya Kayal E, Arya M, Culli J, Krishan S, Agarwal S, Mehndiratta A. Computer-aided diagnosis of cirrhosis and hepatocellular carcinoma using multi-phase abdomen CT. *Int J Comput Assist Radiol Surg* 2019; **14**: 1341-1352 [PMID: 31062266 DOI: 10.1007/s11548-019-01991-5]

22 **Azer SA**. Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: A systematic review. *World J Gastrointest Oncol* 2019; **11**: 1218-1230 [PMID: 31908726 DOI: 10.4251/wjgo.v11.i12.1218]

23 **Shi W**, Kuang S, Cao S, Hu B, Xie S, Chen S, Chen Y, Gao D, Zhu Y, Zhang H, Liu H, Ye M, Sirlin CB, Wang J. Deep learning assisted differentiation of hepatocellular carcinoma from focal liver lesions: choice of four-phase and three-phase CT imaging protocol. *Abdom Radiol (NY)* 2020; **45**: 2688-2697 [PMID: 32232524 DOI: 10.1007/s00261-020-02485-8]

24 **Greenspan H**, Van Ginneken B, Summers RM. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans Med Imag* 2016; **35**: 1153-1159 [DOI: 10.1109/TMI.2016.2553401]

25 **Yasaka K**, Akai H, Abe O, Kiryu S. Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology* 2018; **286**: 887-896 [PMID: 29059036 DOI: 10.1148/radiol.2017170706]

26 **Fukushima K**, Miyake S. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pat Recog* 1982; **15**: 455-469 [DOI: 10.1016/0031-3203(82)90024-3]

27 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

28 **Trivizakis E**, Manikis GC, Nikiforaki K, Drevelegas K, Constantinides M, Drevelegas A, Marias K. Extending 2-D Convolutional Neural Networks to 3-D for Advancing Deep Learning Cancer Classification With Application to MRI Liver Tumor Differentiation. *IEEE J Biomed Health Inform* 2019; **23**: 923-930 [PMID: 30561355 DOI: 10.1109/JBHI.2018.2886276]

29 **Soldera J**, Tomé F, Corso LL, Rech MM, Ferrazza AD, Terres AZ, Cini BT, Eberhardt LZ, Balensiefer JIL, Balbinot RS, Muscope ALF, Longen ML, Schena B, Rost GL Jr, Furlan RG, Balbinot RA, Balbinot SS. Use of a machine learning algorithm to predict rebleeding and mortality for oesophageal variceal bleeding in cirrhotic patients. *EMJ Gastroenterol* 2020; **9**: 46-48. [cited 20 March 2021] Available from: URL: https://www.emjreviews.com/gastroenterology/abstract/use-of-a-machine-learning-algorithm-to-predict-rebleeding-and-mortality-for-oesophageal-variceal-bleeding-in-cirrhotic-patients/

30 **Soldera J**, Tomé F, Corso LL, Ballotin VR, Bigarella LG, Balbinot RS, Rodriguez S, Brandão AB, Hochhegger B. 590 Predicting 30 and 365-day mortality after liver transplantation using a machine learning algorithm. *Gastroenterology* 2021; **160**: S-789 [DOI: 10.1016/S0016-5085(21)02602-0]

31 **Wingfield LR**, Ceresa C, Thorogood S, Fleuriot J, Knight S. Using Artificial Intelligence for Predicting Survival of Individual Grafts in Liver Transplantation: A Systematic Review. *Liver Transpl* 2020; **26**: 922-934 [PMID: 32274856 DOI: 10.1002/lt.25772]

32 **Lai Q**, Spoletini G, Mennini G, Laureiro ZL, Tsilimigras DI, Pawlik TM, Rossi M. Prognostic role of artificial intelligence among patients with hepatocellular cancer: A systematic review. *World J Gastroenterol* 2020; **26**: 6679-6688 [PMID: 33268955 DOI: 10.3748/wjg.v26.i42.6679]

33 **Sato M**, Tateishi R, Yatomi Y, Koike K. Artificial intelligence in the diagnosis and management of hepatocellular carcinoma. *J Gastroenterol Hepatol* 2021; **36**: 551-560 [PMID: 33709610 DOI: 10.1111/jgh.15413]

34 **Yi PS**, Hu CJ, Li CH, Yu F. Clinical value of artificial intelligence in hepatocellular carcinoma: Current status and prospect. *Artif Intell Gastroenterol* 2021; **2**: 42-55 [DOI: 10.35712/aig.v2.i2.42]

35 **Chang KP**, Lin SH, Chu YW. Artificial intelligence in gastrointestinal radiology: A review with special focus on recent development of magnetic resonance and computed tomography. *Artif Intell Gastroenterol* 2021; **2**: 27-41 [DOI: 10.35712/aig.v2.i2.27]

36 **Verde F**, Romeo V, Stanzione A, Maurea S. Current trends of artificial intelligence in cancer imaging. *Artif Intell Med Imaging* 2020; **1**: 87-93 [DOI: 10.35711/aimi.v1.i3.87]

37 **Kudou M**, Kosuga T, Otsuji E. Artificial intelligence in gastrointestinal cancer: Recent advances and future perspectives. *Artif Intell Gastroenterol* 2020; **1**: 71-85 [DOI: 10.35712/aig.v1.i4.71]

38 **Morreale GC**, Sinagra E, Vitello A, Shahini E, Maida M. Emerging artificial intelligence applications in gastroenterology: A review of the literature. *Artif Intell Gastrointest Endosc* 2020; **1**: 6-18 [DOI: 10.37126/aige.v1.i1.6]

39 **Li JW**, Ang TL. Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged? *Artif Intell Gastrointest Endosc* 2021; **2**: 36-49 [DOI: 10.37126/aige.v2.i2.36]

40 **Masuzaki R**, Kanda T, Sasaki R, Matsumoto N, Nirei K, Ogawa M, Moriyama M. Application of artificial intelligence in hepatology: Minireview. *Artif Intell Gastroenterol* 2020; **1**: 5-11 [DOI: 10.35712/aig.v1.i1.5]

41 **Wang CJ**, Hamm CA, Savic LJ, Ferrante M, Schobert I, Schlachter T, Lin M, Weinreb JC, Duncan JS, Chapiro J, Letzen B. Deep learning for liver tumor diagnosis part II: convolutional neural network interpretation using radiologic imaging features. *Eur Radiol* 2019; **29**: 3348-3357 [PMID: 31093705 DOI: 10.1007/s00330-019-06214-8]

42 **Zech JR**, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLoS Med* 2018; **15**: e1002683 [PMID: 30399157 DOI: 10.1371/journal.pmed.1002683]

43 **Li M**, Hsu W, Xie X, Cong J, Gao W. SACNN: Self-Attention Convolutional Neural Network for Low-Dose CT Denoising With Self-Supervised Perceptual Loss Network. *IEEE Trans Med Imaging* 2020; **39**: 2289-2301 [PMID: 31985412 DOI: 10.1109/TMI.2020.2968472]

44 **Chen H**, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, Zhou J, Wang G. Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network. *IEEE Trans Med Imaging* 2017; **36**: 2524-2535 [PMID: 28622671 DOI: 10.1109/TMI.2017.2715284]

45 **Frid-Adar M**, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018; **321**: 321-331 [DOI: 10.1016/j.neucom.2018.09.013]

46 **Sidey-Gibbons JAM**, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol* 2019; **19**: 64 [PMID: 30890124 DOI: 10.1186/s12874-019-0681-4]

47 **Wang W**, Chen Q, Iwamoto Y, Han X, Zhang Q, Hu H, Lin L, Chen YW. Deep Learning-Based Radiomics Models for Early Recurrence Prediction of Hepatocellular Carcinoma with Multi-phase CT Images and Clinical Data. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; **2019**: 4881-4884 [PMID: 31946954 DOI: 10.1109/EMBC.2019.8856356]

48 **Sato M**, Morimoto K, Kajihara S, Tateishi R, Shiina S, Koike K, Yatomi Y. Machine-learning Approach for the Development of a Novel Predictive Model for the Diagnosis of Hepatocellular Carcinoma. *Sci Rep* 2019; **9**: 7704 [PMID: 31147560 DOI: 10.1038/s41598-019-44022-8]

49 **Naeem S**, Ali A, Qadri S, Mashwani WK, Tairan N, Shah H, Fayaz M, Jamal F, Chesneau C, Anam S. Machine-Learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images. *Appl Sci* 2020; **10**: 3134 [DOI: 10.3390/app10093134]

50 **Bousabarah K**, Letzen B, Tefera J, Savic L, Schobert I, Schlachter T, Staib LH, Kocher M, Chapiro J, Lin M. Automated detection and delineation of hepatocellular carcinoma on multiphasic contrast-enhanced MRI using deep learning. *Abdom Radiol* 2020; 1-10 [DOI: 10.1007/s00261-020-02604-5]

51 **Meng L**, Tian Y, Bu S. Liver tumor segmentation based on 3D convolutional neural network with dual scale. *J Appl Clin Med Phys* 2020; **21**: 144-157 [PMID: 31793212 DOI: 10.1002/acm2.12784]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 136-148

DOI: 10.37126/aige.v2.i4.136

ISSN 2689-7164 (online)

MINIREVIEWS

Role of capsule endoscopy in inflammatory bowel disease: Anything new?

Edith Pérez de Arce, Rodrigo Quera, Paulina Núñez F, Raúl Araya

ORCID number: Edith Pérez de Arce 0000-0002-8990-7594; Rodrigo Quera 0000-0001-5854-0526; Paulina Núñez F 0000-0003-3727-1851; Raúl Araya 000-0002-1155-3984.

Author contributions: Pérez de Arce E, Quera R, Núñez F P, and Araya R equally contributed to this review with the conception and design of the study, literature review and analysis, drafting and critical revision and editing, and approval of the final version.

Conflict-of-interest statement: The authors declare no conflict of interest for this article.

Manuscript source: Invited manuscript

Edith Pérez de Arce, Department of Gastroenterology, Hospital Clínico Universidad de Chile, Santiago 8380456, Chile

Rodrigo Quera, Paulina Núñez F, Raúl Araya, Digestive Disease Center, Inflammatory Bowel Disease Program, Clínica Universidad de los Andes, Santiago 7620157, Chile

Paulina Núñez F, Department of Gastroenterology, Hospital San Juan De Dios, Santiago 8350488, Chile

Corresponding author: Rodrigo Quera, MD, Assistant Professor, Digestive Disease Center, Clínica Universidad de los Andes, 2501 Plaza Avenue, Las Condes, Santiago 7620157, Chile. rquera@clinicauandes.cl

Abstract

Capsule endoscopy (CE) is a recently developed diagnostic method for diseases of the small bowel that is non-invasive, safe, and highly tolerable. Its role in patients with inflammatory bowel disease has been widely validated in suspected and established Crohn's disease (CD) due to its ability to assess superficial lesions not detected by cross-sectional imaging and proximal lesions of the small bowel not evaluable by ileocolonoscopy. Because CE is a highly sensitive but less specific technique, differential diagnoses that can simulate CD must be considered, and its interpretation should be supported by other clinical and laboratory indicators. The use of validated scoring systems to characterize and estimate lesion severity (Lewis score, Capsule Endoscopy Crohn's Disease Activity Index), as well as the standardization of the language used to define the lesions (Delphi Consensus), have reduced the interobserver variability in CE reading observed in clinical practice, allowing for the optimization of diagnoses and clinical management strategies. The appearance of the panenteric CE, the incorporation of artificial intelligence, magnetically-guided capsules, and tissue biopsies are elements that contribute to CE being a promising, unique diagnostic tool in digestive tract diseases.

Key Words: Capsule endoscopy; Inflammatory bowel disease; Crohn's disease; Artificial intelligence; Capsule Endoscopy Crohn's Disease Activity Index; Lewis score

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: Chile

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: May 1, 2021
Peer-review started: May 1, 2021
First decision: June 18, 2021
Revised: June 21, 2021
Accepted: August 16, 2021
Article in press: August 16, 2021
Published online: August 28, 2021

P-Reviewer: Cotter J
S-Editor: Liu M
L-Editor: A
P-Editor: Wang LYT

Core Tip: Capsule endoscopy (CE) is the non-invasive diagnostic method of choice for visualizing the small bowel. Its utility is widely validated in both suspected and established Crohn's disease (CD) due to its high sensitivity for detecting early lesions and a high negative predictive value. CE enables estimating the activity and extent of disease, establishing prognosis, and evaluating the therapeutic response in patients with CD. New technologies, such as the panenteric CE and the recent incorporation of artificial intelligence to CE image analysis, render CE an attractive, unique diagnostic tool for diseases of the digestive tract in the future.

Citation: Pérez de Arce E, Quera R, Núñez F P, Araya R. Role of capsule endoscopy in inflammatory bowel disease: Anything new? *Artif Intell Gastrointest Endosc* 2021; 2(4): 136-148
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/136.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.136

INTRODUCTION

Capsule endoscopy (CE) is a non-invasive diagnostic method of increasing development in the study of small bowel diseases. Since its appearance in 2000[1], it has shown its greatest utility in studying obscure gastrointestinal bleeding, celiac disorder, polyposis syndromes, and Crohn's disease (CD). The role of CE in inflammatory bowel disease (IBD), especially in CD, has been extensively investigated in the diagnosis of suspected CD and the management of established CD for the evaluation of disease severity, extent, and response to treatment[2]. The main advantage of CE over ileocolonoscopy is its ability to visualize the mucosa of the proximal small bowel, and compared to imaging studies, its ability to detect superficial mucosal ulcerations missed on magnetic resonance enterography (MRE)[3] and computed tomography enterography[4]. This point is fundamental, considering that studies report that the involvement of the small bowel affects up to 66% of patients diagnosed with CD[5], corresponding to up to 90% of lesions located in the terminal ileum accessible by ileocolonoscopy[6]. Before CE development, the proximal small bowel was examined using indirect imaging methods such as radiography, cross-sectional imaging, and enteroscopy. Detection of lesions in the proximal small bowel is critical due to the implications for managing patients with CD. Jejunal lesions visualized by CE have been found in up to 56% of patients with CD, and these are associated with more severe disease and more rapid progression[7]. Moreover, the role of CE in ulcerative colitis (UC) is not well established because the evidence remains limited. In recent years, remarkable advances in CE technology and design, and the recent use of artificial intelligence, have improved its diagnostic yield in CD.

This review aims to assess the role of CE in IBD and discusses advances in the field and their implications for clinical practice going forward.

DIAGNOSTIC YIELD OF CE IN PATIENTS WITH CD

Studies comparing the diagnostic yield of CE with other diagnostic techniques in patients with CD conclude that CE has a high sensitivity[4,8] and a high negative predictive value (NPV)[9]. However, the diagnostic accuracy of CE has not been determined due to the lack of a gold standard for the diagnosis of CD. A meta-analysis [8] found that CE had a better diagnostic yield than small bowel radiography [52% *vs* 16%; incremental yield (IY) 32%, $P < 0.0001$, 95% confidence interval (CI) = 16%-48%], computed tomography enterography (68 *vs* 21%; IY 47%, w = 47%, $P < 0.00001$, 95%CI = 31%-63%), and ileocolonoscopy (47 *vs* 25%; IY 22%, $P = 0.009$, 95%CI = 5%-39%) in unsuspected CD patients. Similarly, in patients with established CD, CE also outperformed these diagnostic tests[8]. Furthermore, CE was superior to MRE in detecting small bowel lesions in patients with CD, mainly superficial and proximal lesions[10]. A subsequent meta-analysis of 13 studies[3] compared the diagnostic performance of CE with MRE and small bowel contrast ultrasound imaging for the evaluation of small bowel CD. These authors found that the diagnostic yield of CE was similar to MRE

[odds ratio (OR) 1.17; 95%CI: 0.83–1.67] and small bowel contrast ultrasound (OR 0.88; 95%CI: 0.51–1.53) when detecting lesions in the small bowel for both established and suspected CD. However, CE was superior to MRE in detecting proximal small bowel lesions (OR 2.79; 95%CI: 1.2–6.48).

In a recent study among the pediatric population, CE was as sensitive as MRE in identifying inflammatory activity in the terminal ileum and the proximal small bowel; however, the distribution of small bowel inflammation was more extensive when characterized by CE[11].

In summary, the diagnostic yield of CE is at least similar to MRE for established CD in the evaluation of the small bowel. However, the main advantage of CE is the detection of the most proximal and superficial lesions missed on MRE. In suspected CD, CE is more useful when the ileocolonoscopy results are negative.

DIAGNOSTIC SCORES IN CE

At the moment, there are no established diagnostic criteria for the diagnosis of CD by CE. Currently, the Lewis score (LS)[12,13] and Capsule Endoscopy Crohn's Disease Activity Index (CECDAI)[14] are the two validated diagnostic indexes for the evaluation of CE images. Their results must be interpreted in the patient's clinical setting because lesions are not pathognomonic for CD and can be found in other inflammatory conditions. The LS was the first and most widely used index for evaluating inflammatory changes in the mucosa of the small intestine, which is divided into three tertiles according to the transit time estimated using CE. Each characteristic CD finding (villous edema, ulceration, stenosis) is assigned a score for each tertile. The final result of the LS corresponds to the tertile with the highest score, in addition to the stenoses score. A score < 135 is considered normal or clinically insignificant inflammation; from 135 to 790 indicates mild inflammation; and > 790 moderate to severe inflammation[12]. The CECDAI evaluates the proximal and distal segments of the small bowel using an inflammation score (A; 0–5), an extent score (B; 0–3), and a stricture score (C; 0–3), which are combined using the formula A × B + C. The total score (from 0–26) results from adding both the proximal and distal segments. A higher CECDAI score reflects more severe mucosal inflammation[14]. Although there is a good correlation between the LS and CECDAI (Pearson's = 0.81, P = 0.0001) [15], a recent study of 102 patients with CD found that CECDAI was superior to LS in reflecting active intestinal inflammation[16]. Recently, Eliakim *et al*[17] published the Eliakim score, a quantitative measure for PillCam™ Crohn's with excellent reliability that significantly correlates with LS and fecal calprotectin (FC).

CE READING METHOD

So far, manual video review is the method of choice for the detection of lesions in CE. However, a fast-reading method is offered by TOP100, a new software tool in RAPID Reader version 9.0[18]. TOP100 automatically selects the 100 best images from the video with relevant findings, allowing the LS to be calculated quickly. An initial study that compared both reading techniques found agreement in 89.6% of cases calculated by TOP100 as having LS > 135 and those calculated by manual review of the video. Despite these encouraging results, TOP100 should not replace the traditional reading method but rather constitutes a complementary tool for quick LS calculation[18].

DESCRIPTION OF LESIONS WITH CE

Although studies have shown the usefulness of CE in identifying small bowel lesions, one of the difficulties in IBD studies was the lack of nomenclature and descriptions of small bowel lesions. The high interobserver variability in the interpretation and evaluation of the severity of the lesions has both clinical and research implications. Published in 2005, the Capsule Endoscopy Structured Terminology (CEST)[19] is an international consensus on standardized terminology for the findings or lesions detected by CE and also contains guidelines for reporting these findings (structure and content). However, the description of ulcerative and inflammatory lesions in the CEST is ambiguous and limited and, as such, fails to inform clinicians as to which type of lesion is most suspicious for the diagnosis of CD. Therefore, the international Delphi

consensus statement established seven definitions describing the ulcerative and inflammatory lesions seen in CD by CE: aphthoid erosion, deep ulceration, superficial ulceration, stenosis, edema, hyperemia, and denudation[20]. The use of a common language enables standardizing the results of clinical studies and improves patients' health care (Figure 1).

The use of virtual chromoendoscopy, such as flexible spectral color enhancement (FICE), can also be applied in the revision of CE images to improve the visualization of any lesions. FICE enhances mucosal surface patterns using software to convert white light images to certain ranges of wavelengths (red, blue, green). A systematic review and meta-analysis of 13 studies found that the use of FICE failed to significantly improve the injury detection rate in CE[21].

RETENTION OF CE

The CE retention rate (not passed in more than two weeks post-ingestion or less if endoscopic or surgical intervention is required)[22] in the general population ranges from 1.0% to 2.5%[23]. Due to the potential occurrence of stenosis in patients with CD, the retention rate in patients with suspected CD is 2.35%, and with established CD is up to 4.63%[24]. The risk of EC retention can be estimated with a patency capsule (Pillcam™), a capsule with a lactose body, and a barium section for follow-up by fluoroscopy. The disintegration that induces deformation of the capsule or non-expulsion after 30 h suggests small bowel stenosis[23]. The NPV of the patency capsule to predict CE retention ranges from 98% to 100%[25,26]. Given the high risk of CE retention in patients with established CD, and due to the impossibility of distinguishing high from low-risk retention in the clinic, the use of a patency capsule is recommended before CE[24].

CE IN IBD: CLINICAL SCENARIOS

The main clinical scenarios for the application of EC for IBD are both suspected and established CD. CE studies in UC are limited.

CE and suspected CD

The European Society of Gastrointestinal Endoscopy Clinical Guideline[23] and the Clinical practice guidelines for the use of CE[27] recommend the use of CE in patients with suspected CD and negative ileocolonoscopy[23,27] and imaging results[27] as a diagnostic method for the evaluation of the small bowel, in the absence of obstructive symptoms or known stricture.

In a study with 95 patients, CE excluded the diagnosis of CD if the result was negative (NPV of 96%). Only 3% of the cases with negative CEs were diagnosed with CD after 15 mo of follow-up[9]. Moreover, minor lesions detected by CE may be present in more than 10% of healthy subjects[28]. Non-steroidal anti-inflammatory drug (NSAID)-induced enteropathy is one of the main differential diagnoses of small bowel lesions. In this setting, lesions can appear as early as 2 wk from the onset of NSAID therapy[29,30]. Other differential diagnoses include radiation enteritis, ischemia, Bechet's disease, lymphoma, and gastrointestinal infections[30]. Then, the interpretation of the findings from CE against suspected CD must be supported for other clinical elements due to the impossibility of obtaining tissue samples by CE.

The use of biomarkers as a screening method for intestinal inflammation, such as FC, could be useful in patients with suspected CD. FC is a cytosolic protein present in neutrophils that is released during inflammation; as such, its elevation in stool samples is a good indicator of intestinal inflammation[31]. Although it is highly sensitive, it is not specific since its levels can increase in IBD, colon cancer, ischemic colitis, and NSAID-induced enteropathy, among others[31]. Although FC has shown higher sensitivity and a stronger correlation with inflammatory activity in UC[32], in CD, the usefulness of FC is less established[33,34], particularly in the small bowel. However, recent studies have shown that FC could be a useful tool for selecting which patients should undergo CE for suspected CD when the ileocolonoscopy results are negative due to its ability to predict inflammatory activity in CE in patients with suspected CD [35-37]. Monteiro *et al*[35] found a moderate positive correlation ($r = 0.56$, $P < 0.0019$) between FC and the LS. FC > 100 μg/g were correlated with LS > 135 in 89% of patients, showing a sensitivity of 78.6%, specificity of 87.9%, positive predictive value

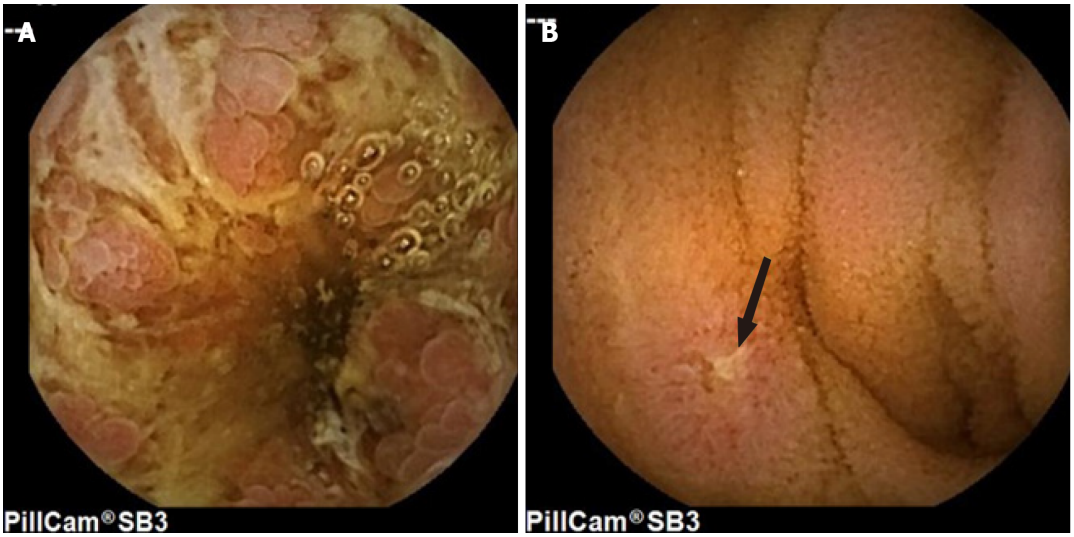

Figure 1 Capsule endoscopy findings in Crohn's disease. A: Deep ulceration; B: Aphthoid erosion (superficial lesion).

of 89.2%, and NPV of 76.3%[35]. Similar findings for FC[27,36,38] and, to a lesser degree, for CRP[36,38] were described by other authors. In a subsequent meta-analysis of 463 patients from seven studies, FC had a significant diagnostic accuracy in detecting small bowel CD, and with FC values < 50 μg/g, the probability of a positive diagnosis was very low[39].

Considering the available evidence and due to CE's ability to diagnose early disease, in patients with suspected CD (typical symptoms, elevated fecal and plasma biomarkers, anemia, or extraintestinal manifestations of CD), CE should be performed even if ileocolonoscopy results are positive due to the need to evaluate proximal lesions that could determine prognosis and treatment strategies (Figure 2).

CE in established CD

The American consensus guidelines for the use of CE recommends its use in patients with established CD when: (1) Clinical features unexplained by ileocolonoscopy or imaging studies are present; (2) The assessment of small bowel mucosal healing (not evaluable by ileocolonoscopy) is needed; and (3) Small bowel recurrence of CD after colectomy is suspected, undiagnosed by ileocolonoscopy or imaging studies[27]. Recently, the European Crohn's and Colitis Organisation and the European Society of Gastrointestinal and Abdominal Radiology guidelines recommend CE along with intestinal ultrasound and MRE for initial evaluation and follow-up of established CD [40] (Figure 3).

CE in patients with CD with unexplained clinical features

The persistence of irritable bowel disease-like symptoms in patients with IBD in remission can occur in almost one-third of patients[41,42], being more frequent in patients with CD[42]. In a scenario where traditional diagnostics tests (ileocolonoscopy and cross-sectional imaging) are normal, CE could play a role in evaluating the small bowel to rule out disease activity as the symptom origin. Another clinical scenario is the study of persistent anemia in patients with CD in remission.

CE in follow-up and prediction of relapse

Studies have shown that the clinical response to treatment does not correlate with mucosal healing in patients with CD of the small bowel evaluated by CE[43]. Therefore, objective monitoring of disease activity in the small bowel is necessary. Hall *et al*[43] conducted the first prospective study in 43 patients with CD evaluated with CE at baseline and after 52 wk of treatment. The authors found that 90% of the patients had an active CD in their small bowel at baseline, yet only 65% at week 52 of treatment, with 42% of the patients achieving complete mucosal healing at week 52 (P < 0.0001, 95%CI: 0.62-0.22). Stenosis detected by CE was a poor prognostic factor for the response to treatment in this study[43]. In a subsequent prospective study in 43 patients with CD in clinical remission, fecal biomarkers (FC, lactoferrin, and S100A12) were good predictors of mucosal healing assessed by CE, proving useful in monitoring the CD progression[44]. Finally, a recent prospective observational cohort study

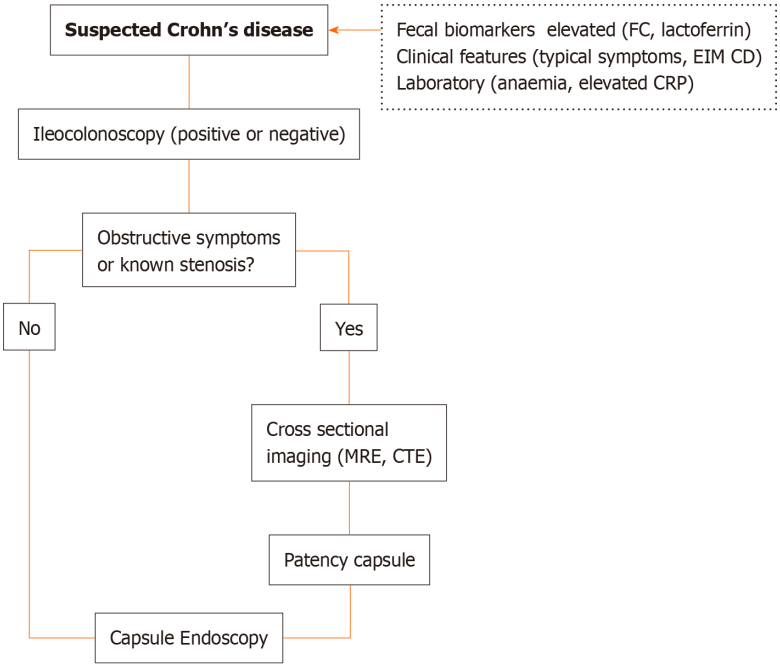

Figure 2 Diagnostic algorithm recommended for suspected Crohn's disease. MRE: Magnetic resonance enterography; CTE: Computed tomography enterography; FC: Fecal calprotectin; CRP: C-reactive protein; EIM: Extraintestinal manifestations.

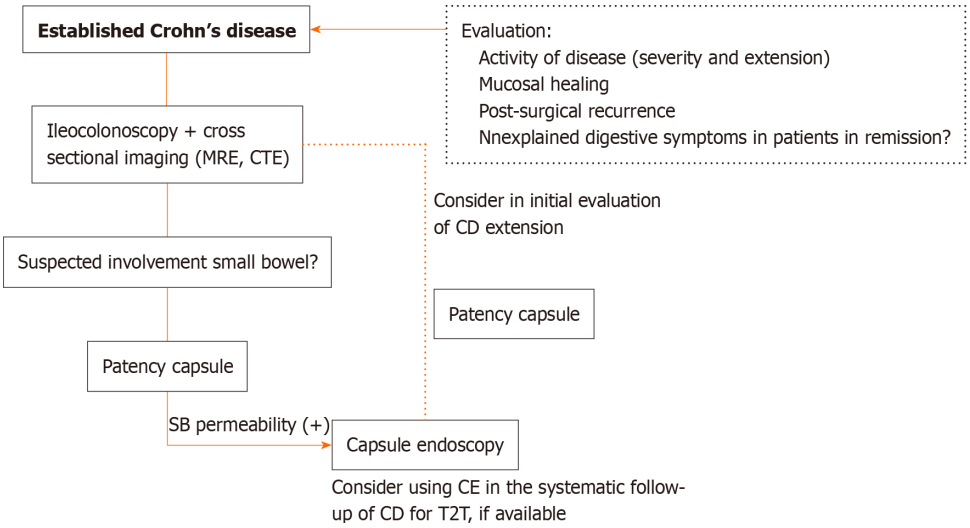

Figure 3 Diagnostic algorithm recommended for established Crohn's disease. T2T: Treat to target; MRE: Magnetic resonance enterography; CTE: Computed tomography enterography; CD: Crohn's disease; SB: Small bowel.

assessed the ability of MRE, FC, and CE to predict flare-ups in patients with quiescent CD. CE predicted both short-term (3 mo) and long-term [24 mo, area under curve (AUC) 0.79, 95%CI: 0.66–0.88; P = 0.0001] flares, while FC only predicted short-term flares within 3 mo (AUC 0.81, 95%CI: 0.76-0.85), and MRE correlated with 2-year flare risk (AUC 0.71, 95%CI: 0.58-0.82; P = 0.024)[45].

CE and post-surgical recurrence of CD

In a recent study, Shiga *et al*[46] compared the postoperative follow-up for CE in patients with CD who underwent intestinal resection with the appearance of clinical symptoms for treatment adjustment. In the CE group, 87% residual or recurrent lesions were found at the 3rd postoperative month. Adjusted treatment based on EC findings revealed a strong protective effect (0.30, 0.10–0.75)[46]. This study did not compare the use of CE with ileocolonoscopy in the postoperative follow-up. However, it included 37% of small bowel resections not evaluable by ileocolonoscopy. Previous studies have shown post-surgical recurrence by CE that was not detected by

ileocolonoscopy, which has allowed active treatment in this group of patients[47]. Although ileocolonoscopy continues to be the gold standard for the search for postoperative recurrence, CE is an excellent complementary tool, if available, that improves diagnostic performance in this clinical setting.

CE and "treat to target" in CD

The "treat to target" strategy in CD[48] is based on the regular assessment of disease activity by using validated outcome measures and the subsequent adjustment of treatment of disease activity, following targets, where the main target is mucosal healing. A recent systematic review that included 47 studies highlighted CE as an objective method of evaluating CD activity that enables reclassifying patients with CD, monitoring the effect of medical treatment through the evaluation of mucosal healing, and detecting postoperative recurrence[49]. Owing to its diagnostic accuracy, CE could be incorporated into the "treat to target" management of patients with CD[49]. However, larger, randomized, controlled trials are necessary to confirm these findings.

CE and IBD – undefined

CE allows the classification of patients with a diagnosis of IBD-undefined (IBD-U)[50-53], where the inflammatory involvement of the colon cannot differentiate between UC and CD. IBD-U occurs in up to 10% to 15% of patients[54], and at least 15% to 30% of patients will be reclassified as having CD during their disease[50,55]. Establishing this difference is important from a surgical point of view regarding the selection of the type of surgery and which complications to expect in patients with CD and, from the medical point of view, in the selection of the type of biological therapies.

CE and UC

The role of CE in the evaluation of the colonic mucosa in UC is unclear. Colon CE (CCE and later CCE-2 or second-generation) was developed in 2006 and was designed for non-invasive visualization of the colon[56]. A systematic review showed that the diagnostic accuracy of CCE in the colon is comparable with ileocolonoscopy in assessing the severity and extent of the disease[57]. However, some studies with a small number of patients have found a weak correlation between the findings from CCE and colonoscopy, which supports the latter for the evaluation of the mucosa in UC[58,59].

Regarding the evaluation of the small intestine in UC, a prospective observational study (capcolitis) on CE in 127 patients with known UC found that only 4% of the diagnoses changed to CD upon evaluating the small bowel with CE[60].

PANENTERIC CE

Panenteric CE (PCE) is a new type of CE similar to PillCam™ COLON 2 (CCE-2) and is currently known as PillCam™ Crohn's System (Medtronic, Dublin, Ireland)[61]. PillCam™ Crohn's System is designed for the evaluation of the mucosa of patients with CD. This capsule has a field of view that allows for a 344° view between both capsule heads to provide a pan-intestinal panoramic visualization. The rate frame of PillCam™ Crohn's System ranges from 4–35 frames per second depending on the speed of the capsule into the gut and has an operating time of more than 12 h[61]. PCE was first described in a multicenter prospective study where it demonstrated a better diagnostic yield of PCE than ileocolonoscopy in 66 patients with active CD who underwent both modalities[62]. The authors found that the per-subject diagnostic yield rate for active CD lesions was 83.3% for PCE and 69.7% for ileocolonoscopy (yield difference 13.6%; 95%CI: 2.6%–24.7%), and the per-segment diagnostic yield rate was 40.6% for PCE and 32.7% for ileocolonoscopy (yield difference 7.9%; 95%CI: 3.3%–12.4%)[62].

In an observational cohort study performed on 93 patients (established CD: 71 and suspected CD: 22), the use of PCE allowed to change the treatment in 38.7% of patients [63]. Moreover, Montreal classification was up-staged in 33.8% of patients with established CD, and identifying proximal small bowel disease in 12.7% predicted treatment intensification[63]. A recent prospective, multicenter study in patients with established CD found that sensitivity of PCE was superior to MRE for proximal small bowel inflammation (97% *vs* 71%, $P = 0.021$) and similar to MRE and/or ileocolonoscopy in the terminal ileum and colon[64]. However, the overall sensitivity for active enteric inflammation for CE *vs* MRE and/or ileocolonoscopy was similar (94% *vs* 100%, $P = 0.125$), but the specificity was 74% *vs* 22%, respectively ($P = 0.001$)

[64]. In the pediatric population, a prospective study in 48 children with CD found that PCE led to a change in therapy for 71% of patients at baseline and 23% at 24 wk. A "treat to target" strategy in these children led to increased mucosal healing and deep remission from 21% at baseline to 54% at week 24 and 58% at week 52[65]. A recent multicenter study[66] compared the 344° panoramic-view recorded by PillCam™ Crohn's System (lesions detected by cameras A and B) with the standard 172°-view (lesions detected by one camera only) in 41 patients who underwent CE for suspected or established CD. The study found that the panoramic 344°-view increased small bowel CE accuracy *vs* the standard 172°-view, detecting a greater number of relevant lesions (56.1% *vs* 39.0%; $P = 0.023$), resulting in higher LS (222.8 *vs* 185.7; $P = 0.031$), and improved clinical management (48.8% *vs* 31.7%, $P = 0.023$)[66].

PCE, as the only study modality, could reduce costs associated with the evaluation of patients with CD, considering the need for MRE and ileocolonoscopy for the complete evaluation of the intestine in these patients. Furthermore, PCE is a safe method preferred by patients[64] that does not require sedation, representing advantages for the pediatric population[65].

Table 1 presents the main characteristics of the capsules used in IBD.

In summary, based on the available literature, CE is essential in evaluating patients with CD. The finding of lesions in the small bowel detected by CE and not observed in conventional studies (cross-sectional imaging, ileo-colonoscopy) determines changes in the Montreal classification in patients with CD[10]. This leads to a modification of the therapeutic strategies, with the earlier introduction of immunomodulators and/or biological therapy, improving the prognosis of these patients[67].

ARTIFICIAL INTELLIGENCE IN CE AND ITS APPLICATION IN IBD

In recent years, the development of artificial intelligence (AI) in medicine has made it possible to apply this technology to the automated identification of images on CE. AI, through deep learning artificial neural network (ANN) algorithms[68], facilitates image recognition according to which characteristics the algorithm chooses for itself based on what it considers best for that task, which requires much less time than conventional readings by endoscopists (5.9 min *vs* 96.6 min)[69]. Convolutional neural network (CNN), a type of ANN[68] applied to CE, has shown excellent performance for the detection of ulcers, polyps, celiac disease, and bleeding[69].

A recent study by Klang *et al*[70] evaluated the accuracy of CNN for the detection of ulcers in CD on CE for image sets from 49 patients. They reported an AUC of 0.99 for split images and accuracies ranging from 95.4% to 96.7%. The AUC for individual patients was 0.94 to 0.99[69]. Also, the use of CNN enabled characterizing the severity of ulcers on CE images in patients with CD with high accuracy in the detection of severe CD ulcerations and better differentiation between mild and severe ulceration (accuracy 0.91, 95%CI: 0.867-0.954) but a less accurate separation of moderate from severe: (Accuracy 0.78, 95%CI: 0.716–0.844) and mild *vs* moderate (accuracy 0.624, 95%CI: 0.547–0.701)[71]. Undoubtedly, this technology provides accurate and rapid detection of ulcers from CE images, thereby decreasing reading times. Moreover, deep neural networks are highly accurate in detecting stenosis in CE images (accuracy 93.5%) and differentiating between stenosis and healthy mucosa (AUC 0.989), stenosis, and all ulcers (AUC 0.942), and stenosis and different degrees of ulcer severity[72]. In another area, recent studies suggest that CNN would allow for the automatic evaluation of the degree of intestinal cleansing in CE studies, which could serve as a means of comparing different intestinal preparation methods and thus design recommendations[73].

Despite the encouraging results on the use of AI on CE in IBD, prospective studies are necessary to evaluate its usefulness in the diagnosis and follow-up in CD.

OTHER NEWS IN CE

Because CE passage is passive and dependent on the peristalsis of the intestine, only 80 to 90% of patients have their entire intestine visualized. Thus, up to 30% of minor injuries may not be seen during the study[23]**.** One of the new challenges is the possibility of directing the navigation of the CE in the intestine. Magnetically-assisted CE (MACE) has been tested as a screening tool in gastric cancer[74], Barrett's esophagus, and esophageal varix[75]. MACE has generated results comparable with esophagogastroduodenoscopy in detecting focal lesions[76] and the study of iron

Table 1 Characteristics of available capsule endoscopy systems for the study of inflammatory bowel disease

	SB CE (Pillcam SB3Ò)	Colon CE	PillCam CrohnÒ
Dimensions	26 mm × 11 mm	32 mm × 11 mm	32.3 mm ± 0.5 mm × 11.6 mm
Weight	3.0 g	2.9 g	2.9 g
Camera	One	2-one at each end	2-one at each end
Field of view	156º ISO-8600-3	344º: 172° ISO-8600-3 per camera	344º: 172° ISO-8600-3 per camera
Frame rate	2-6 fps (2-6)	4-35 fps (AFR)	4-35 fps (AFR)
operating time	≥ 8 or longer (max.15)	10 h	Minimum of 10 hr
Operating temperature	20-40 ºC	20-40 ºC	20-40 ºC

SB: Small bowel; CE: Capsule endoscopy; AFR: Adaptative frame rate; fps: Frames per second.

deficiency anemia[77]; however, it has not been evaluated in patients with IBD.

Other CE prototypes in development include biopsy[78] and drug delivery[79] capabilities, which could be clinically relevant for patients with IBD in the future.

CONCLUSION

The use of CE has played a fundamental role in evaluating the small bowel of patients with IBD, mainly in those with suspected CD and established CD. The development of new types of capsules, such as the panenteric capsule, and the integration of AI into CE image analysis, have improved the visualization and automated the identification of lesions in the digestive tract using a non-invasive, safe, highly tolerated method. Treatment optimization for patients with CD, thanks to CE findings, has improved the course of the disease. More studies are needed to support the use of CE in the evaluation of all patients with CD.

REFERENCES

1 **Iddan G**, Meron G, Glukhovsky A, Swain P. Wireless capsule endoscopy. *Nature* 2000; **405**: 417 [PMID: 10839527 DOI: 10.1038/35013140]

2 **McCain JD**, Pasha SF, Leighton JA. Role of Capsule Endoscopy in Inflammatory Bowel Disease. *Gastrointest Endosc Clin N Am* 2021; **31**: 345-361 [PMID: 33743930 DOI: 10.1016/j.giec.2020.12.004]

3 **Kopylov U**, Yung DE, Engel T, Vijayan S, Har-Noy O, Katz L, Oliva S, Avni T, Battat R, Eliakim R, Ben-Horin S, Koulaouzidis A. Diagnostic yield of capsule endoscopy versus magnetic resonance enterography and small bowel contrast ultrasound in the evaluation of small bowel Crohn's disease: Systematic review and meta-analysis. *Dig Liver Dis* 2017; **49**: 854-863 [PMID: 28512034 DOI: 10.1016/j.dld.2017.04.013]

4 **Dionisio PM**, Gurudu SR, Leighton JA, Leontiadis GI, Fleischer DE, Hara AK, Heigh RI, Shiff AD, Sharma VK. Capsule endoscopy has a significantly higher diagnostic yield in patients with suspected and established small-bowel Crohn's disease: a meta-analysis. *Am J Gastroenterol* 2010; **105**: 1240-8; quiz 1249 [PMID: 20029412 DOI: 10.1038/ajg.2009.713]

5 **Van Assche G**, Dignass A, Panes J, Beaugerie L, Karagiannis J, Allez M, Ochsenkühn T, Orchard T, Rogler G, Louis E, Kupcinskas L, Mantzaris G, Travis S, Stange E; European Crohn's and Colitis Organisation (ECCO). The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. *J Crohns Colitis* 2010; **4**: 7-27 [PMID: 21122488 DOI: 10.1016/j.crohns.2009.12.003]

6 **Jensen MD**, Nathan T, Rafaelsen SR, Kjeldsen J. Ileoscopy reduces the need for small bowel imaging in suspected Crohn's disease. *Dan Med J* 2012; **59**: A4491 [PMID: 22951194]

7 **Flamant M**, Trang C, Maillard O, Sacher-Huvelin S, Le Rhun M, Galmiche JP, Bourreille A. The prevalence and outcome of jejunal lesions visualized by small bowel capsule endoscopy in Crohn's disease. *Inflamm Bowel Dis* 2013; **19**: 1390-1396 [PMID: 23552764 DOI: 10.1097/MIB.0b013e31828133c1]

8 **Triester SL**, Leighton JA, Leontiadis GI, Gurudu SR, Fleischer DE, Hara AK, Heigh RI, Shiff AD, Sharma VK. A meta-analysis of the yield of capsule endoscopy compared to other diagnostic

modalities in patients with non-stricturing small bowel Crohn's disease. *Am J Gastroenterol* 2006; **101**: 954-964 [PMID: 16696781 DOI: 10.1111/j.1572-0241.2006.00506.x]

9 **Hall B**, Holleran G, Costigan D, McNamara D. Capsule endoscopy: High negative predictive value in the long term despite a low diagnostic yield in patients with suspected Crohn's disease. *United European Gastroenterol J* 2013; **1**: 461-466 [PMID: 24917998 DOI: 10.1177/2050640613508551]

10 **González-Suárez B**, Rodriguez S, Ricart E, Ordás I, Rimola J, Díaz-González Á, Romero C, de Miguel CR, Jáuregui A, Araujo IK, Ramirez A, Gallego M, Fernández-Esparrach G, Ginés Á, Sendino O, Llach J, Panés J. Comparison of Capsule Endoscopy and Magnetic Resonance Enterography for the Assessment of Small Bowel Lesions in Crohn's Disease. *Inflamm Bowel Dis* 2018; **24**: 775-780 [PMID: 29506048 DOI: 10.1093/ibd/izx107]

11 **Prichard DO**, Hamilton Z, Savage T, Smyth M, Penner C, Lakhani A, Carroll MW, Al Sarkhy A, Lemberg DA, Enns R, Jamieson D, Jacobson K. Capsule Endoscopy Complements Magnetic Resonance Enterography and Endoscopy in Evaluating Small Bowel Crohn's Disease. *J Can Assoc Gastroenterol* 2020; **3**: 279-287 [PMID: 33241181 DOI: 10.1093/jcag/gwz028]

12 **Gralnek IM**, Defranchis R, Seidman E, Leighton JA, Legnani P, Lewis BS. Development of a capsule endoscopy scoring index for small bowel mucosal inflammatory change. *Aliment Pharmacol Ther* 2008; **27**: 146-154 [PMID: 17956598 DOI: 10.1111/j.1365-2036.2007.03556.x]

13 **Cotter J**, Dias de Castro F, Magalhães J, Moreira MJ, Rosa B. Validation of the Lewis score for the evaluation of small-bowel Crohn's disease activity. *Endoscopy* 2015; **47**: 330-335 [PMID: 25412092 DOI: 10.1055/s-0034-1390894]

14 **Gal E**, Geller A, Fraser G, Levi Z, Niv Y. Assessment and validation of the new capsule endoscopy Crohn's disease activity index (CECDAI). *Dig Dis Sci* 2008; **53**: 1933-1937 [PMID: 18034304 DOI: 10.1007/s10620-007-0084-y]

15 **Yablecovitch D**, Lahat A, Neuman S, Levhar N, Avidan B, Ben-Horin S, Eliakim R, Kopylov U. The Lewis score or the capsule endoscopy Crohn's disease activity index: which one is better for the assessment of small bowel inflammation in established Crohn's disease? *Therap Adv Gastroenterol* 2018; **11**: 1756283X17747780 [PMID: 29399042 DOI: 10.1177/1756283X17747780]

16 **Omori T**, Kambayashi H, Murasugi S, Ito A, Yonezawa M, Nakamura S, Tokushige K. Comparison of Lewis Score and Capsule Endoscopy Crohn's Disease Activity Index in Patients with Crohn's Disease. *Dig Dis Sci* 2020; **65**: 1180-1188 [PMID: 31541367 DOI: 10.1007/s10620-019-05837-7]

17 **Eliakim R**, Yablecovitch D, Lahat A, Ungar B, Shachar E, Carter D, Selinger L, Neuman S, Ben-Horin S, Kopylov U. A novel PillCam Crohn's capsule score (Eliakim score) for quantification of mucosal inflammation in Crohn's disease. *United European Gastroenterol J* 2020; **8**: 544-551 [PMID: 32213037 DOI: 10.1177/2050640620913368]

18 **Freitas M**, Arieira C, Carvalho PB, Rosa B, Moreira MJ, Cotter J. Simplify to improve in capsule endoscopy - TOP 100 is a swift and reliable evaluation tool for the small bowel inflammatory activity in Crohn's disease. *Scand J Gastroenterol* 2020; **55**: 408-413 [PMID: 32228199 DOI: 10.1080/00365521.2020.1745880]

19 **Korman LY**, Delvaux M, Gay G, Hagenmuller F, Keuchel M, Friedman S, Weinstein M, Shetzline M, Cave D, de Franchis R. Capsule endoscopy structured terminology (CEST): proposal of a standardized and structured terminology for reporting capsule endoscopy procedures. *Endoscopy* 2005; **37**: 951-959 [PMID: 16189767 DOI: 10.1055/s-2005-870329]

20 **Leenhardt R**, Buisson A, Bourreille A, Marteau P, Koulaouzidis A, Li C, Keuchel M, Rondonotti E, Toth E, Plevris JN, Eliakim R, Rosa B, Triantafyllou K, Elli L, Wurm Johansson G, Panter S, Ellul P, Pérez-Cuadrado Robles E, McNamara D, Beaumont H, Spada C, Cavallaro F, Cholet F, Fernandez-Urien Sainz I, Kopylov U, McAlindon ME, Németh A, Tontini GE, Yung DE, Niv Y, Rahmi G, Saurin JC, Dray X. Nomenclature and semantic descriptions of ulcerative and inflammatory lesions seen in Crohn's disease in small bowel capsule endoscopy: An international Delphi consensus statement. *United European Gastroenterol J* 2020; **8**: 99-107 [PMID: 32213061 DOI: 10.1177/2050640619895864]

21 **Yung DE**, Boal Carvalho P, Giannakou A, Kopylov U, Rosa B, Rondonotti E, Toth E, Plevris JN, Koulaouzidis A. Clinical validity of flexible spectral imaging color enhancement (FICE) in small-bowel capsule endoscopy: a systematic review and meta-analysis. *Endoscopy* 2017; **49**: 258-269 [PMID: 28122387 DOI: 10.1055/s-0042-122015]

22 **Cave D**, Legnani P, de Franchis R, Lewis BS; ICCE. ICCE consensus for capsule retention. *Endoscopy* 2005; **37**: 1065-1067 [PMID: 16189792 DOI: 10.1055/s-2005-870264]

23 **Pennazio M**, Spada C, Eliakim R, Keuchel M, May A, Mulder CJ, Rondonotti E, Adler SN, Albert J, Baltes P, Barbaro F, Cellier C, Charton JP, Delvaux M, Despott EJ, Domagk D, Klein A, McAlindon M, Rosa B, Rowse G, Sanders DS, Saurin JC, Sidhu R, Dumonceau JM, Hassan C, Gralnek IM. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy* 2015; **47**: 352-376 [PMID: 25826168 DOI: 10.1055/s-0034-1391855]

24 **Pasha SF**, Pennazio M, Rondonotti E, Wolf D, Buras MR, Albert JG, Cohen SA, Cotter J, D'Haens G, Eliakim R, Rubin DT, Leighton JA. Capsule Retention in Crohn's Disease: A Meta-analysis. *Inflamm Bowel Dis* 2020; **26**: 33-42 [PMID: 31050736 DOI: 10.1093/ibd/izz083]

25 **Cebrián García A**, Elosua González A, Fernández-Urién Sainz I. Use of patency capsule in daily practice. *Rev Esp Enferm Dig* 2019; **111**: 491-492 [PMID: 31021169 DOI: 10.17235/reed.2019.5952/2018]

26 **Albuquerque A**, Cardoso H, Marques M, Rodrigues S, Vilas-Boas F, Lopes S, Dias CC, Macedo G.

Predictive factors of small bowel patency in Crohn's disease patients. *Rev Esp Enferm Dig* 2016; **108**: 65-70 [PMID: 26838487 DOI: 10.17235/reed.2015.3957/2015]

27 **Enns RA**, Hookey L, Armstrong D, Bernstein CN, Heitman SJ, Teshima C, Leontiadis GI, Tse F, Sadowski D. Clinical Practice Guidelines for the Use of Video Capsule Endoscopy. *Gastroenterology* 2017; **152**: 497-514 [PMID: 28063287 DOI: 10.1053/j.gastro.2016.12.032]

28 **Goldstein JL**, Eisen GM, Lewis B, Gralnek IM, Zlotnick S, Fort JG; Investigators. Video capsule endoscopy to prospectively assess small bowel injury with celecoxib, naproxen plus omeprazole, and placebo. *Clin Gastroenterol Hepatol* 2005; **3**: 133-141 [PMID: 15704047 DOI: 10.1016/s1542-3565(04)00619-6]

29 **Maiden L**, Thjodleifsson B, Seigal A, Bjarnason II, Scott D, Birgisson S, Bjarnason I. Long-term effects of nonsteroidal anti-inflammatory drugs and cyclooxygenase-2 selective agents on the small bowel: a cross-sectional capsule enteroscopy study. *Clin Gastroenterol Hepatol* 2007; **5**: 1040-1045 [PMID: 17625980 DOI: 10.1016/j.cgh.2007.04.031]

30 **Bar-Meir S**. Review article: capsule endoscopy - are all small intestinal lesions Crohn's disease? *Aliment Pharmacol Ther* 2006; **24** Suppl 3: 19-21 [PMID: 16961739 DOI: 10.1111/j.1365-2036.2006.03054.x]

31 **Pérez de Arce E**, Sedano R, Quera R. [Biomarkers in inflammatory bowel disease]. *Rev Med Chil* 2020; **148**: 362-370 [PMID: 32730381 DOI: 10.4067/S0034-98872020000300362]

32 **Mosli MH**, Zou G, Garg SK, Feagan SG, MacDonald JK, Chande N, Sandborn WJ, Feagan BG. C-Reactive Protein, Fecal Calprotectin, and Stool Lactoferrin for Detection of Endoscopic Activity in Symptomatic Inflammatory Bowel Disease Patients: A Systematic Review and Meta-Analysis. *Am J Gastroenterol* 2015; **110**: 802-19; quiz 820 [PMID: 25964225 DOI: 10.1038/ajg.2015.120]

33 **Stawczyk-Eder K**, Eder P, Lykowska-Szuber L, Krela-Kazmierczak I, Klimczak K, Szymczak A, Szachta P, Katulska K, Linke K. Is faecal calprotectin equally useful in all Crohn's disease locations? *Arch Med Sci* 2015; **11**: 353-361 [PMID: 25995752 DOI: 10.5114/aoms.2014.43672]

34 **Sipponen T**, Haapamäki J, Savilahti E, Alfthan H, Hämäläinen E, Rautiainen H, Koskenpato J, Nuutinen H, Färkkilä M. Fecal calprotectin and S100A12 have low utility in prediction of small bowel Crohn's disease detected by wireless capsule endoscopy. *Scand J Gastroenterol* 2012; **47**: 778-784 [PMID: 22519419 DOI: 10.3109/00365521.2012.677953]

35 **Monteiro S**, Barbosa M, Cúrdia Gonçalves T, Boal Carvalho P, Moreira MJ, Rosa B, Cotter J. Fecal Calprotectin as a Selection Tool for Small Bowel Capsule Endoscopy in Suspected Crohn's Disease. *Inflamm Bowel Dis* 2018; **24**: 2033-2038 [PMID: 29722829 DOI: 10.1093/ibd/izy098]

36 **Egea Valenzuela J**, Pereñíguez López A, Pérez Fernández V, Alberca de Las Parras F, Carballo Álvarez F. Fecal calprotectin and C-reactive protein are associated with positive findings in capsule endoscopy in suspected small bowel Crohn's disease. *Rev Esp Enferm Dig* 2016; **108**: 394-400 [PMID: 27312194 DOI: 10.17235/reed.2016.4318/2016]

37 **Koulaouzidis A**, Nemeth A, Johansson GW, Toth E. Dissecting Lewis score under the light of fecal calprotectin; an analysis of correlation of score components with calprotectin levels in capsule endoscopy. *Ann Gastroenterol* 2015; **28**: 259-264 [PMID: 25830236]

38 **Höög CM**, Bark LÅ, Broström O, Sjöqvist U. Capsule endoscopic findings correlate with fecal calprotectin and C-reactive protein in patients with suspected small-bowel Crohn's disease. *Scand J Gastroenterol* 2014; **49**: 1084-1090 [PMID: 24853318 DOI: 10.3109/00365521.2014.920915]

39 **Kopylov U**, Yung DE, Engel T, Avni T, Battat R, Ben-Horin S, Plevris JN, Eliakim R, Koulaouzidis A. Fecal calprotectin for the prediction of small-bowel Crohn's disease by capsule endoscopy: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol* 2016; **28**: 1137-1144 [PMID: 27415156 DOI: 10.1097/MEG.0000000000000692]

40 **Maaser C**, Sturm A, Vavricka SR, Kucharzik T, Fiorino G, Annese V, Calabrese E, Baumgart DC, Bettenworth D, Borralho Nunes P, Burisch J, Castiglione F, Eliakim R, Ellul P, González-Lama Y, Gordon H, Halligan S, Katsanos K, Kopylov U, Kotze PG, Krustinš E, Laghi A, Limdi JK, Rieder F, Rimola J, Taylor SA, Tolan D, van Rheenen P, Verstockt B, Stoker J; European Crohn's and Colitis Organisation [ECCO] and the European Society of Gastrointestinal and Abdominal Radiology [ESGAR]. ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: Initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis* 2019; **13**: 144-164 [PMID: 30137275 DOI: 10.1093/ecco-jcc/jjy113]

41 **Isgar B**, Harman M, Kaye MD, Whorwell PJ. Symptoms of irritable bowel syndrome in ulcerative colitis in remission. *Gut* 1983; **24**: 190-192 [PMID: 6826101 DOI: 10.1136/gut.24.3.190]

42 **Halpin SJ**, Ford AC. Prevalence of symptoms meeting criteria for irritable bowel syndrome in inflammatory bowel disease: systematic review and meta-analysis. *Am J Gastroenterol* 2012; **107**: 1474-1482 [PMID: 22929759 DOI: 10.1038/ajg.2012.260]

43 **Hall B**, Holleran G, Chin JL, Smith S, Ryan B, Mahmud N, McNamara D. A prospective 52 week mucosal healing assessment of small bowel Crohn's disease as detected by capsule endoscopy. *J Crohns Colitis* 2014; **8**: 1601-1609 [PMID: 25257546 DOI: 10.1016/j.crohns.2014.09.005]

44 **Aggarwal V**, Day AS, Connor S, Leach ST, Brown G, Singh R, Friedman A, Zekry A, Craig PI. Role of capsule endoscopy and fecal biomarkers in small-bowel Crohn's disease to assess remission and predict relapse. *Gastrointest Endosc* 2017; **86**: 1070-1078 [PMID: 28947363 DOI: 10.1016/j.gie.2017.09.011]

45 **Ben-Horin S**, Lahat A, Amitai MM, Klang E, Yablecovitch D, Neuman S, Levhar N, Selinger L, Rozendorn N, Turner D, Chowers Y, Odes S, Schwartz D, Yanai H, Dotan I, Braun T, Haberman Y, Kopylov U, Eliakim R; Israeli IBD Research Nucleus (IIRN). Assessment of small bowel mucosal

healing by video capsule endoscopy for the prediction of short-term and long-term risk of Crohn's disease flare: a prospective cohort study. *Lancet Gastroenterol Hepatol* 2019; **4**: 519-528 [PMID: 31080097 DOI: 10.1016/S2468-1253(19)30088-3]

46 **Shiga H**, Abe I, Kusaka J, Shimoyama Y, Moroi R, Kuroha M, Kakuta Y, Kinouchi Y, Masamune A. Capsule Endoscopy Is Useful for Postoperative Tight Control Management in Patients with Crohn's Disease. *Dig Dis Sci* 2021 [PMID: 33495918 DOI: 10.1007/s10620-021-06841-6]

47 **Han ZM**, Qiao WG, Ai XY, Li AM, Chen ZY, Feng XC, Zhang J, Wan TM, Xu ZM, Bai Y, Li MS, Liu SD, Zhi FC. Impact of capsule endoscopy on prevention of postoperative recurrence of Crohn's disease. *Gastrointest Endosc* 2018; **87**: 1489-1498 [PMID: 29355520 DOI: 10.1016/j.gie.2018.01.017]

48 **Bouguen G**, Levesque BG, Feagan BG, Kavanaugh A, Peyrin-Biroulet L, Colombel JF, Hanauer SB, Sandborn WJ. Treat to target: a proposed new paradigm for the management of Crohn's disease. *Clin Gastroenterol Hepatol* 2015; **13**: 1042-50.e2 [PMID: 24036054 DOI: 10.1016/j.cgh.2013.09.006]

49 **Le Berre C**, Trang-Poisson C, Bourreille A. Small bowel capsule endoscopy and treat-to-target in Crohn's disease: A systematic review. *World J Gastroenterol* 2019; **25**: 4534-4554 [PMID: 31496630 DOI: 10.3748/wjg.v25.i31.4534]

50 **Mehdizadeh S**, Chen G, Enayati PJ, Cheng DW, Han NJ, Shaye OA, Ippoliti A, Vasiliauskas EA, Lo SK, Papadakis KA. Diagnostic yield of capsule endoscopy in ulcerative colitis and inflammatory bowel disease of unclassified type (IBDU). *Endoscopy* 2008; **40**: 30-35 [PMID: 18058654 DOI: 10.1055/s-2007-995359]

51 **Maunoury V**, Savoye G, Bourreille A, Bouhnik Y, Jarry M, Sacher-Huvelin S, Ben Soussan E, Lerebours E, Galmiche JP, Colombel JF. Value of wireless capsule endoscopy in patients with indeterminate colitis (inflammatory bowel disease type unclassified). *Inflamm Bowel Dis* 2007; **13**: 152-155 [PMID: 17206697 DOI: 10.1002/ibd.20060]

52 **Cohen SA**, Gralnek IM, Ephrath H, Saripkin L, Meyers W, Sherrod O, Napier A, Gobin T. Capsule endoscopy may reclassify pediatric inflammatory bowel disease: a historical analysis. *J Pediatr Gastroenterol Nutr* 2008; **47**: 31-36 [PMID: 18607266 DOI: 10.1097/MPG.0b013e318160df85]

53 **Monteiro S**, Dias de Castro F, Boal Carvalho P, Rosa B, Moreira MJ, Pinho R, Saraiva MM, Cotter J. Essential role of small bowel capsule endoscopy in reclassification of colonic inflammatory bowel disease type unclassified. *World J Gastrointest Endosc* 2017; **9**: 34-40 [PMID: 28101306 DOI: 10.4253/wjge.v9.i1.34]

54 **Guindi M**, Riddell RH. Indeterminate colitis. *J Clin Pathol* 2004; **57**: 1233-1244 [PMID: 15563659 DOI: 10.1136/jcp.2003.015214]

55 **Eliakim R**. The impact of wireless capsule endoscopy on gastrointestinal diseases. *South Med J* 2007; **100**: 235-236 [PMID: 17396720 DOI: 10.1097/01.smj.0000257405.87268.48]

56 **Eliakim R**, Fireman Z, Gralnek IM, Yassin K, Waterman M, Kopelman Y, Lachter J, Koslowsky B, Adler SN. Evaluation of the PillCam Colon capsule in the detection of colonic pathology: results of the first multicenter, prospective, comparative study. *Endoscopy* 2006; **38**: 963-970 [PMID: 17058158 DOI: 10.1055/s-2006-944832]

57 **Shi HY**, Chan FKL, Higashimori A, Kyaw M, Ching JYL, Chan HCH, Chan JCH, Chan AWH, Lam KLY, Tang RSY, Wu JCY, Sung JJY, Ng SC. A prospective study on second-generation colon capsule endoscopy to detect mucosal lesions and disease activity in ulcerative colitis (with video). *Gastrointest Endosc* 2017; **86**: 1139-1146.e6 [PMID: 28713062 DOI: 10.1016/j.gie.2017.07.007]

58 **Meister T**, Heinzow HS, Domagk D, Dortgolz A, Lenze F, Ross M, Domschke W, Lügering A. Colon capsule endoscopy versus standard colonoscopy in assessing disease activity of ulcerative colitis: a prospective trial. *Tech Coloproctol* 2013; **17**: 641-646 [PMID: 23307507 DOI: 10.1007/s10151-012-0965-8]

59 **Manes G**, Ardizzone S, Cassinotti A. PillCam Colon and ulcerative colitis: what do physicians need to know? *Endoscopy* 2013; **45**: 325 [PMID: 23533079 DOI: 10.1055/s-0032-1326410]

60 **Bokemeyer B**, Luehr D, Helwig U, Maaser C, Jessen P, Schreiber S. Small bowel capsule endoscopy in ulcerative colitis: the capcolitis study: a prospective observational study. *Eur J Gastroenterol Hepatol* 2019; **31**: 766-772 [PMID: 31082999 DOI: 10.1097/MEG.0000000000001410]

61 **Eliakim R**, Spada C, Lapidus A, Eyal I, Pecere S, Fernández-Urién I, Lahat A, Costamagna G, Schwartz A, Ron Y, Yanai H, Adler S. Evaluation of a new pan-enteric video capsule endoscopy system in patients with suspected or established inflammatory bowel disease - feasibility study. *Endosc Int Open* 2018; **6**: E1235-E1246 [PMID: 30302381 DOI: 10.1055/a-0677-170]

62 **Leighton JA**, Helper DJ, Gralnek IM, Dotan I, Fernandez-Urien I, Lahat A, Malik P, Mullin GE, Rosa B. Comparing diagnostic yield of a novel pan-enteric video capsule endoscope with ileocolonoscopy in patients with active Crohn's disease: a feasibility study. *Gastrointest Endosc* 2017; **85**: 196-205.e1 [PMID: 27658907 DOI: 10.1016/j.gie.2016.09.009]

63 **Tai FWD**, Ellul P, Elosua A, Fernandez-Urien I, Tontini GE, Elli L, Eliakim R, Kopylov U, Koo S, Parker C, Panter S, Sidhu R, McAlindon M. Panenteric capsule endoscopy identifies proximal small bowel disease guiding upstaging and treatment intensification in Crohn's disease: A European multicentre observational cohort study. *United European Gastroenterol J* 2021; **9**: 248-255 [PMID: 32741315 DOI: 10.1177/2050640620948664]

64 **Bruining DH**, Oliva S, Fleisher MR, Fischer M, Fletcher JG; BLINK study group. Panenteric capsule endoscopy versus ileocolonoscopy plus magnetic resonance enterography in Crohn's disease: a multicentre, prospective study. *BMJ Open Gastroenterol* 2020; **7** [PMID: 32499275 DOI: 10.1136/bmjgast-2019-000365]

65 **Oliva S**, Aloi M, Viola F, Mallardo S, Civitelli F, Maccioni F, Hassan C, Papoff P, Cucchiara S, Cohen SA. A Treat to Target Strategy Using Panenteric Capsule Endoscopy in Pediatric Patients With Crohn's Disease. *Clin Gastroenterol Hepatol* 2019; **17**: 2060-2067.e1 [PMID: 30326301 DOI: 10.1016/j.cgh.2018.10.015]

66 **Tontini GE**, Rizzello F, Cavallaro F, Bonitta G, Gelli D, Pastorelli L, Salice M, Vecchi M, Gionchetti P, Calabrese C. Usefulness of panoramic 344°-viewing in Crohn's disease capsule endoscopy: a proof of concept pilot study with the novel PillCam™ Crohn's system. *BMC Gastroenterol* 2020; **20**: 97 [PMID: 32264831 DOI: 10.1186/s12876-020-01231-0]

67 **Cotter J**, Dias de Castro F, Moreira MJ, Rosa B. Tailoring Crohn's disease treatment: the impact of small bowel capsule endoscopy. *J Crohns Colitis* 2014; **8**: 1610-1615 [PMID: 24631311 DOI: 10.1016/j.crohns.2014.02.018]

68 **Klang E**. Deep learning and medical imaging. *J Thorac Dis* 2018; **10**: 1325-1328 [PMID: 29708147 DOI: 10.21037/jtd.2018.02.76]

69 **Ding Z**, Shi H, Zhang H, Meng L, Fan M, Han C, Zhang K, Ming F, Xie X, Liu H, Liu J, Lin R, Hou X. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* 2019; **157**: 1044-1054.e5 [PMID: 31251929 DOI: 10.1053/j.gastro.2019.06.025]

70 **Klang E**, Barash Y, Margalit RY, Soffer S, Shimon O, Albshesh A, Ben-Horin S, Amitai MM, Eliakim R, Kopylov U. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc* 2020; **91**: 606-613.e2 [PMID: 31743689 DOI: 10.1016/j.gie.2019.11.012]

71 **Barash Y**, Azaria L, Soffer S, Margalit Yehuda R, Shlomi O, Ben-Horin S, Eliakim R, Klang E, Kopylov U. Ulcer severity grading in video capsule images of patients with Crohn's disease: an ordinal neural network solution. *Gastrointest Endosc* 2021; **93**: 187-192 [PMID: 32535191 DOI: 10.1016/j.gie.2020.05.066]

72 **Klang E**, Grinman A, Soffer S, Margalit Yehuda R, Barzilay O, Amitai MM, Konen E, Ben-Horin S, Eliakim R, Barash Y, Kopylov U. Automated Detection of Crohn's Disease Intestinal Strictures on Capsule Endoscopy Images Using Deep Neural Networks. *J Crohns Colitis* 2021; **15**: 749-756 [PMID: 33216853 DOI: 10.1093/ecco-jcc/jjaa234]

73 **Noorda R**, Nevárez A, Colomer A, Pons Beltrán V, Naranjo V. Automatic evaluation of degree of cleanliness in capsule endoscopy based on a novel CNN architecture. *Sci Rep* 2020; **10**: 17706 [PMID: 33077755 DOI: 10.1038/s41598-020-74668-8]

74 **Zhao AJ**, Qian YY, Sun H, Hou X, Pan J, Liu X, Zhou W, Chen YZ, Jiang X, Li ZS, Liao Z. Screening for gastric cancer with magnetically controlled capsule gastroscopy in asymptomatic individuals. *Gastrointest Endosc* 2018; **88**: 466-474.e1 [PMID: 29753039 DOI: 10.1016/j.gie.2018.05.003]

75 **Beg S**, Card T, Warburton S, Rahman I, Wilkes E, White J, Ragunath K. Diagnosis of Barrett's esophagus and esophageal varices using a magnetically assisted capsule endoscopy system. *Gastrointest Endosc* 2020; **91**: 773-781.e1 [PMID: 31678203 DOI: 10.1016/j.gie.2019.10.031]

76 **Liao Z**, Hou X, Lin-Hu EQ, Sheng JQ, Ge ZZ, Jiang B, Hou XH, Liu JY, Li Z, Huang QY, Zhao XJ, Li N, Gao YJ, Zhang Y, Zhou JQ, Wang XY, Liu J, Xie XP, Yang CM, Liu HL, Sun XT, Zou WB, Li ZS. Accuracy of Magnetically Controlled Capsule Endoscopy, Compared With Conventional Gastroscopy, in Detection of Gastric Diseases. *Clin Gastroenterol Hepatol* 2016; **14**: 1266-1273.e1 [PMID: 27211503 DOI: 10.1016/j.cgh.2016.05.013]

77 **Ching HL**, Hale MF, Kurien M, Campbell JA, Chetcuti Zammit S, Healy A, Thurston V, Hebden JM, Sidhu R, McAlindon ME. Diagnostic yield of magnetically assisted capsule endoscopy versus gastroscopy in recurrent and refractory iron deficiency anemia. *Endoscopy* 2019; **51**: 409-418 [PMID: 30360012 DOI: 10.1055/a-0750-5682]

78 **Son D**, Gilbert H, Sitti M. Magnetically Actuated Soft Capsule Endoscope for Fine-Needle Biopsy. *Soft Robot* 2020; **7**: 10-21 [PMID: 31418640 DOI: 10.1089/soro.2018.0171]

79 **Stewart FR**, Newton IP, Näthke I, Huang Z, Cox BF, Cochran S. Development of a therapeutic capsule endoscope for treatment in the gastrointestinal tract: bench testing to translational trial. 2017 IEEE International Ultrasonics Symposium (IUS); 2017 Sep 6-9; Washington, D.C., United States. Piscataway (NJ): IEEE, 2017: 1-4 [DOI: 10.1109/ULTSYM.2017.8091791]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 149-156

DOI: 10.37126/aige.v2.i4.149

ISSN 2689-7164 (online)

MINIREVIEWS

Role of optical coherence tomography in Barrett's esophagus

Nikhil Gupta, Raghav Yelamanchi, Himanshu Agrawal, Nitin Agarwal

ORCID number: Nikhil Gupta 0000-0001-7265-8168; Raghav Yelamanchi 0000-0001-6786-8056; Himanshu Agrawal 0000-0001-7994-2356; Nitin Agarwal 0000-0002-4820-9896.

Author contributions: Gupta N performed the literature review and critically reviewed the manuscript; Yelamanchi R performed the literature review and drafted the manuscript; Agrawal H and Agarwal N reviewed the manuscript; all authors read and approved the final manuscript.

Conflict-of-interest statement: None of the authors has any conflict of interest or financial ties to disclose.

Manuscript source: Invited manuscript

Nikhil Gupta, Raghav Yelamanchi, Himanshu Agrawal, Department of Surgery, Atal Bihari Vajpayee Institute of Medical Sciences and Dr. Ram Manohar Lohia Hospital, Delhi 110001, India

Nitin Agarwal, Department of Surgical Disciploines, Postgraduate Institute of Medical Education and Research and Dr. Ram Manohar Lohia Hospital, Delhi 110001, India

Corresponding author: Nikhil Gupta, FACS, FRCS (Gen Surg), Professor, Surgeon, Department of Surgery, Atal Bihari Vajpayee Institute of Medical Sciences and Dr. Ram Manohar Lohia Hospital, BKS Marg, Delhi 110001, India. nikhil_ms26@yahoo.co.in

Abstract

Traditional endoscopic techniques for Barrett's esophagus (BE) surveillance relied on factor of probability as endoscopists performed cumbersome random biopsies of low yield. Optical coherence tomography (OCT) is a novel technique based on tissue light interference and is set to break conventional barriers. OCT was initially introduced in ophthalmology but was soon adopted by other areas of medicine. When applied to endoscopy, OCT can render images of the superficial layers of the gastrointestinal tract and is highly sensitive in detecting dysplasia in BE. Volumetric laser endomicroscopy is a second generation OCT endoscope device which is able to identify buried glands after ablation. Addition of artificial intelligence to OCT has rendered it more productive. The newer additions to OCT such as angiogram and laser marking will increase the accuracy of investigation. In spite of the few inevitable drawbacks associated with the technology, it presently outperforms all newer endoscopic techniques for the surveillance of BE.

Key Words: Optical coherence tomography; Volume laser endomicroscopy; Esophageal adenocarcinoma; Endoscopy; Gastroesophageal reflux disease

Core Tip: Surveillance of Barrett's esophagus for dysplasia is a long-debated and intensively researched topic. Optical coherence tomography (OCT) is a breakthrough technology in the medical field that enables the visualization of the layers of a structure in an office setting. The application of artificial intelligence (AI) to OCT endoscopy is the latest addition to the armamentarium of endoscopists. AI-based diagnostic algorithm scores are proven to be better than clinical scores. The accuracy of AI-based

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: India

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: May 9, 2021
Peer-review started: May 9, 2021
First decision: May 19, 2021
Revised: May 20, 2021
Accepted: July 19, 2021
Article in press: July 19, 2021
Published online: August 28, 2021

P-Reviewer: Morozov S
S-Editor: Fan JR
L-Editor: A
P-Editor: Zhang YL

system is enhanced further by using color coding software and convolutional neural networks. Multi-center randomized control trials validating these technologies is the need of the hour.

Citation: Gupta N, Yelamanchi R, Agrawal H, Agarwal N. Role of optical coherence tomography in Barrett's esophagus. *Artif Intell Gastrointest Endosc* 2021; 2(4): 149-156
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/149.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.149

INTRODUCTION

Barrett's esophagus (BE) is defined as columnar metaplasia (or intestinal metaplasia, as some authorities prefer to call it) of the stratified squamous epithelium, lining the lower end of the esophagus[1]. It occurs due to chronic exposure of the distal esophagus to acidic contents as a part of gastroesophageal reflux disease. The prevalence of BE is around 1.6% in general population[2]. It varies in different regions of the world with increased prevalence in the western population. Apart from the gastroesophageal reflux disease, other risk factors for BE include advancing age, male gender, obesity, tobacco consumption, and Caucasian race[3].

Once the diagnosis of BE is made based on endoscopy, the endoscopist evaluates its extent as *per* the Prague C and M classification. All cases of BE should be biopsied at multiple levels as *per* the Seattle biopsy protocol to identify the presence of dysplasia or adenocarcinoma, which is the main concern. Traditional endoscopic techniques relied on the chance factor as endoscopists performed random cumbersome biopsies of low yield. The early diagnosis of esophageal neoplasia is important because it helps to initiate curative therapies for cancer. This has directed the path of research to identify newer techniques and technologies to increase the accuracy of biopsies during endoscopy[4]. Optical coherence tomography (OCT) is one of such techniques which is set to break conventional barriers.

Humans are prone to do errors due to fatigue, increased workload and working environment. The use of artificial intelligence (AI) has grown rapidly in the past few decades from using technology to perform simple household tasks to piloting aircraft. AI is also adopted into the medical field in the form of surgical robots in the last decade. The application of AI to endoscopy is widely researched as newer technologies of endoscopy are being developed. The purpose of this narrative review is to enlighten the readers about the principles of OCT and its application to BE and the use of AI in the OCT endoscopy.

OCT

OCT is an imaging modality based on light interference. It is used to produce cross-sectional images of a structure based on the differential properties of various layers with respect to light refraction[5]. The basic setup of OCT consists of a light source which is a low-coherence semiconductor super-luminescent diode. The light is split into two beams by an optical splitter: A reference beam and a sample beam. The reference beam is reflected back by a mirror, while the sample beam is focused onto the tissue to be imaged. Based on the refractory properties of the layers of the tissue, the sample beam is variably reflected back. The reflected light from the reference and sample beams are coupled in a coupler, producing interference patterns which are analyzed, after which a cross-sectional image is created (Figure 1). The axial resolution of OCT will depend on the spectral band of the light source with large spectral bands having better resolution[5]. The transverse resolution is independent of axial resolution and will depend on the numerical aperture of the lens through which the light beam passes[5].

The conventional OCT technology is based on the time-domain (TD-OCT) concept in which variations in the time of the travelled beams of light are analyzed to form an image with the help of moving mirrors. The technology has now evolved into the Fourier-domain (FD-OCT) which uses static mirrors so an image is formed based on the modulations in the source spectrum. The FD-OCT has higher image acquisition

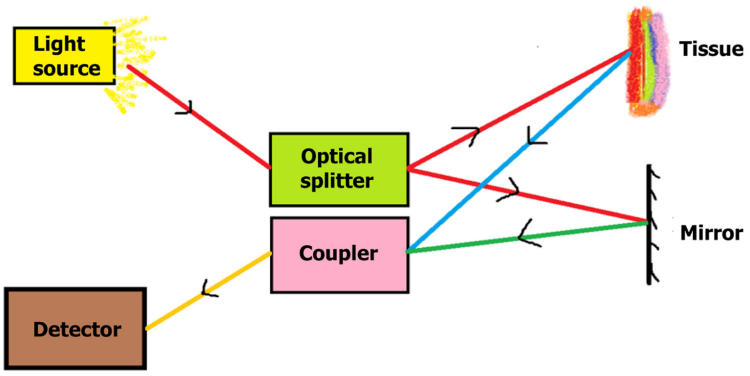

Figure 1 Basic schematic representation of the principle of optical coherence tomography.

speeds than TD-OCT. The resolution of FD-OCT is 1-3 μm, which is far better than the 10 μm resolution of TD-OCT. The FD-OCT is based on either charge-coupled device-based image acquisition (spectral-domain OCT) or photodetector-based image detection with longer wavelengths of the light source (swept-source OCT)[6]. The swept-source OCT has better resolution and twice the image acquisition speed compared to spectral-domain OCT[6].

OCT was initially introduced in ophthalmology as a method to visualize the layers of the retina but it was soon adopted into other areas of medicine. Nevertheless, the utility of OCT is still only the "tip of the iceberg" with its vast potential yet to be unleashed. When applied to endoscopy, OCT is able to render images of the superficial layers of the gastrointestinal tract. OCT can be combined with either a forward-viewing endoscope or a side-viewing endoscope, with the forward-viewing endoscope enabling the sampling of the desired tissue[7]. There are two main types of OCT endoscopes: The proximal scanning rotating endoscope, which is less expensive but has lower capture speed, and the distal scanning endoscope, which comes with a micromotor, acquires images at a much higher speed but comes at a cost higher than the proximal scanning endoscope[7].

Volumetric laser endomicroscopy (VLE) is a second generation OCT endoscope device presently used for imaging[8] (Figure 2). It uses balloon centered imaging probes for imaging with a high axial resolution of 7 μm and a depth of 3 mm, which is 10 times greater compared to the standard endoscopic ultrasound[9]. It images the esophagus in six-centimeter intervals and is quite fast in image acquisition compared to the conventional OCT. It images about 1200 cross-sectional areas in the six cm span which are reconstructed. The application of VLE in BE is mainly to diagnose suspicious areas of mucosal abnormalities and in the post-treatment surveillance of BE and early neoplastic lesions.

PREDICTIVE FEATURES OF DYSPLASIA IN BE USING OCT/VLE AND THE USE OF AI

The absence of layering, surface maturation, and gland maturation are the three independent predictive factors for dysplasia in OCT imaging. The surface maturation is assessed in terms of the surface OCT signal, which if equal or stronger than the sub-surface signal, is predictive of dysplasia. Gland maturation is assessed in terms of the number of abnormal glands identified in imaging with more than five glands predictive of dysplasia.

AI is based on computer algorithms which provide result based on the received input. The algorithms are created based on previous OCT images which are correlated with histological diagnosis. The AI system has been automated to evolve with time, based on its previous results just as a human brain which is known as machine learning. Machine learning may be supervised, semi-supervised or unsupervised. Hence, AI is said to as good as a human brain and sometimes even better. Swager *et al* [10] created an AI-based VLE prediction score using multivariable logistic regression analysis of 60 VLE images[10]. The components of the score were: the lack of layering of superficial layers, higher surface intensity than sub-surface intensity, and the number of abnormal glands (Table 1). A cut-off score of ≥ 8 was predicative of dysplasia with a sensitivity and specificity of 83% and 71% respectively[10]. This VLE

Table 1 Volumetric laser endomicroscopy prediction score and diagnostic algorithm[9]

VLE prediction score		
Parameter		Score
Layering	Layering present-more than 50%	0
	Layering present–less than 50%	8
Surface signal	Surface signal < subsurface signal	0
	Surface signal = subsurface signal	6
	Surface signal > subsurface signal	8
Abnormal glands	0-5	0
	> 5	5
VLE-diagnostic algorithm		
Mucosal layer partial effacement	Abnormal glands > 5	Dysplasia
	Abnormal glands ≤ 5	Non-dysplasia
Mucosal layer complete effacement	Surface intensity > subsurface intensity	Dysplasia
	Surface intensity ≤ subsurface intensity	Non-dysplasia

VLE: Volumetric laser endomicroscopy.

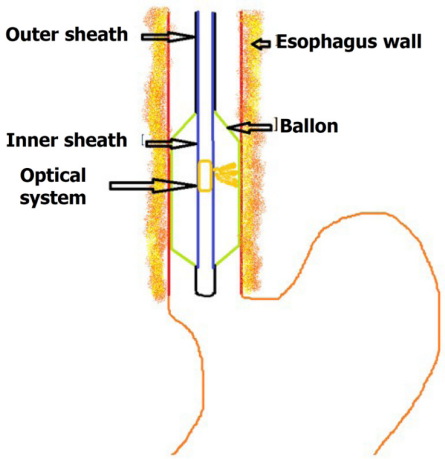

Figure 2 Parts of the volumetric laser endomicroscopy device.

prediction score based on computer-based VLE diagnostic algorithm (VLE-DA) was more sensitive (86%) and specific (88%) than the clinical VLE predication score[10-12]. The components of VLE-DA are listed in Table 1.

Outcomes of first generation OCT

The traditional OCT criteria were found to be 97% sensitive and 93% specific when applied to BE surveillance prospectively in a study by Poneros *et al*[13] in 2001. The accuracy of OCT in diagnosing dysplasia in BE was about 78% in a double-blinded study by Isenberg *et al*[14] in 2005. The utility of OCT in diagnosing dysplasia was also confirmed in a study by Evans *et al*[15] using the dysplasia index which was 83% sensitive and 75% specific[15]. Chen *et al*[16] used ultra-high-resolution OCT for diagnosing dysplasia and adenocarcinoma with an accuracy of 83.3% and 100% respectively[16]. The utility of ultra-high-resolution OCT was also confirmed in the study by Cobb *et al*[17].

OUTCOMES OF SECOND-GENERATION OCT

The imaging capability of three-dimensional OCT is faster than conventional OCT. Its utility was proved in the study by Adler *et al*[18]. VLE has been found to be more sensitive and specific than random blind biopsies as *per* Seattle protocol. The role of VLE was initially proved in a study by Vakoc *et al*[19], while in a study by Trindade *et al*[20] five out of six patients were upstaged due to the diagnosis of dysplasia which was missed by conventional endoscopy and narrow band imaging[19,20]. The sensitivity and specificity of VLE in diagnosing dysplasia was 86% and 88% in a study by Leggett *et al*[11]. In a study by Jain *et al*[21], VLE was compared with histology; the sensitivity in diagnosing BE-related dysplasia was 50% and specificity was 47.1%[21]. The false negative rate was 2.9%. Even though the specificity was low in the study, it is far better than the random biopsies. In a systematic review by Kohli *et al*[22], the sensitivity and specificity of OCT in diagnosing dysplasia and early malignancy was in the ranges of 68%-83% and 75%-82% respectively[22].

POST-ABLATION BE SURVEILLANCE USING OCT

A variety of ablation therapies such as radiofrequency ablation, cryoablation, laser ablation, photodynamic therapy, *etc.* are used for the treatment of high-grade BE dysplasia and insitu carcinoma. One of the main disadvantages of these procedures is the occurrence of buried glands or subsquamous glandular structures[23,24]. These glands, present beneath the epithelium, may undergo dysplastic changes and turn malignant, but are not visualized on routine endoscopy as the surface epithelium appears normal. OCT is one of the few techniques able to diagnose buried glands[25]. The sensitivity and specificity in identifying buried glands in post-treatment BE using VLE was shown to be 92.3% and 23.8% in a study by Jain *et al*[21]. However, in the study by Swager *et al*[26], most of the subsquamous glandular structures identified on OCT were histologically normal[26]. The role of OCT in post-ablative surveillance was also proved in a study by Benjamin *et al*[27].

Doppler-OCT is useful in detecting the changes in the sub-mucosal micro-vascular network, which further improves the accuracy of OCT. Doppler-OCT is also used to detect the change in the vascular pattern during post-photodynamic therapy for BE. Doppler-OCT helps to monitor the dose of photodynamic therapy[28,29].

NEWER ADDITIONS TO OCT

As neoplasia is associated with neovascularization, this is one of the features used to distinguish benign epithelium form malignancy. OCT angiography is used to image the subsurface vasculature without the need for any contrast and is useful in diagnosing neoplasia[30]. The changes in the OCT signal caused by the movement of erythrocytes are quantified by calculating the decorrelation. However, this makes the OCT signal susceptible to artifacts due to respiratory and cardiac movements.

As a balloon is used to augment the scanning speed in VLE, simultaneous sampling of mucosa is not possible. The biopsy taken from the mucosa may not be the original mucosa intended on imaging. This disadvantage is overcome by using laser marking along with VLE. The laser fiber is used for creating point coagulation spots which act as markers for biopsy after the scan[31,32]. Simultaneous laser coagulation along with OCT is also possible[32].

The addition of deep learning to AI-based OCT systems further improved the accuracy of prediction of BE related dysplasia. Deep learning is one kind of machine learning where multiple diagnostic algorithms are layered to form a convolutional neural network just as a human brain. The output from one layer is fed to the next layer which further processes it and feeds it to the next layer to produce a refined output[33]. Deep learning also increases the speed of processing the images.

Trindade *et al*[34] used an AI-based new software termed intelligent real-time image segmentation for BE surveillance. The software provided color codes based on the degree of dysplasia using the previously mentioned VLE prediction features[34]. A multi-center randomized control trial with trial number NCT03814824 is going on, validating the above software, the results of which are awaited.

OCT IN COMPARISON TO OTHER ADVANCED ENDOSCOPIC IMAGING

VLE has been proved to be better than confocal laser endomicroscopy (CLE), which is one of the emerging endoscopic imaging techniques for BE and associated dysplasia. The sensitivity and specificity of VLE using VLE-DA were higher than CLE in a study by Leggett *et al*[11]. CLE is also disadvantageous as it requires injection of contrast into the blood and a limited field of view and imaging depth[35]. Endoscopic ultrasound is an excellent imaging modality for assessing the depth of tumor involvement. However, its accuracy is lower in differentiating early invasive carcinoma (T1 and T2). In a study by Kahn *et al*[36], VLE showed good results in differentiating T1a lesions from T1b lesions[36].

DRAWBACKS OF OCT

All technologies have one or more drawbacks and OCT is no exception. The main drawback of OCT is the absence of real-time imaging, as it is the case with other imaging modalities. Even the fastest OCT technology and probes require seconds to process the reflected waves. VLE requires balloon apposition and although perfect apposition is theoretically possible, it is rare in reality. The mucous layer on the surface epithelium, the contractions of the esophagus, and the presence of blood interfere with the close approximation resulting in artifacts. Simultaneous biopsy is not possible during imaging in VLE probes, which may pose a difficulty in biopsying the originally identified area. Movement artifacts are common in Doppler-OCT and OCT angiography. Unlike endoscopic ultrasound, OCT cannot be used to image the deeper tissues. Finally, cost is one of the main limiting factors for the widespread usage in all institutes.

The application of AI to OCT requires inputs from a large number of experts with expertise in this new technology who are fewer at present. The accuracy of the AI systems is based on the data fed which requires advanced imaging techniques and higher quality images. As AI and machine learning require input from humans it may be the victim of human errors during data input. Much of the knowledge of AI in OCT is based on pilot studies and case series. The number of randomized control trials and multi-center trials are very less due to concerns raised by ethical committees.

CONCLUSION

Surveillance of BE for dysplasia is a long-debated and intensively researched topic. OCT is a breakthrough technology in the medical field that enables the visualization of the layers of a structure in an office setting. The application of OCT to endoscopy is the latest addition to the armamentarium of endoscopists. Even though earlier OCT instruments were slow to image tissues, the newer AI-based technologies are fast enough to add only a few minutes to the conventional endoscopy time and are highly accurate compared to clinical diagnosis. OCT is highly sensitive in detecting dysplasia in BE. Even though the specificity in diagnosing dysplasia is lower, it is far more efficient than the conventional blind biopsy protocol. An especially important feature is the ability of VLE to identify buried glands after ablation. The newer additions to OCT, such as angiogram and laser marking, will help to increase the accuracy of the investigation. The AI software systems and deep learning systems are evolving over time. However, the utility of AI to BE surveillance is still at its bud stage. In spite of the few unavoidable drawbacks associated with the technology, AI-based OCT system is presently the most promising of all newer endoscopic techniques for the surveillance of BE.

REFERENCES

1 **Naini BV**, Souza RF, Odze RD. Barrett's Esophagus: A Comprehensive and Contemporary Review for Pathologists. *Am J Surg Pathol* 2016; **40**: e45-e66 [PMID: 26813745 DOI: 10.1097/PAS.0000000000000598]

2 **Dam AN**, Klapman J. A narrative review of Barrett's esophagus in 2020, molecular and clinical update. *Ann Transl Med* 2020; **8**: 1107 [PMID: 33145326 DOI: 10.21037/atm-20-4406]

3 **Qumseya BJ**, Bukannan A, Gendy S, Ahemd Y, Sultan S, Bain P, Gross SA, Iyer P, Wani S.

Systematic review and meta-analysis of prevalence and risk factors for Barrett's esophagus. *Gastrointest Endosc* 2019; **90**: 707-717.e1 [PMID: 31152737 DOI: 10.1016/j.gie.2019.05.030]

4 **Sharma P**. New endoscopic techniques for detecting dysplasia in barrett esophagus. *Gastroenterol Hepatol (NY)* 2008; **4**: 460-461 [PMID: 21960917]

5 **Popescu DP**, Choo-Smith LP, Flueraru C, Mao Y, Chang S, Disano J, Sherif S, Sowa MG. Optical coherence tomography: fundamental principles, instrumental designs and biomedical applications. *Biophys Rev* 2011; **3**: 155 [PMID: 28510064 DOI: 10.1007/s12551-011-0054-7]

6 **Gabriele ML**, Wollstein G, Ishikawa H, Kagemann L, Xu J, Folio LS, Schuman JS. Optical coherence tomography: history, current status, and laboratory work. *Invest Ophthalmol Vis Sci* 2011; **52**: 2425-2436 [PMID: 21493951 DOI: 10.1167/iovs.10-6312]

7 **Gora MJ**, Suter MJ, Tearney GJ, Li X. Endoscopic optical coherence tomography: technologies and clinical applications [Invited]. *Biomed Opt Express* 2017; **8**: 2405-2444 [PMID: 28663882 DOI: 10.1364/BOE.8.002405]

8 **Wolfsen HC**. Volumetric Laser Endomicroscopy in Patients With Barrett Esophagus. *Gastroenterol Hepatol (NY)* 2016; **12**: 719-722 [PMID: 28035200]

9 **Rodriguez MAC**, de Moura DTH, Ribeiro IB, Bernardo WM, Morita FHA, Marques SB, Sakai P, de Moura EGH. Volumetric laser endomicroscopy and optical coherence tomography in Barrett's esophagus: a systematic review and meta-analysis. *Endosc Int Open* 2019; **7**: E1078-E1091 [PMID: 31475224 DOI: 10.1055/a-0965-6487]

10 **Swager AF**, Tearney GJ, Leggett CL, van Oijen MGH, Meijer SL, Weusten BL, Curvers WL, Bergman JJGHM. Identification of volumetric laser endomicroscopy features predictive for early neoplasia in Barrett's esophagus using high-quality histological correlation. *Gastrointest Endosc* 2017; **85**: 918-926.e7 [PMID: 27658906 DOI: 10.1016/j.gie.2016.09.012]

11 **Leggett CL**, Gorospe EC, Chan DK, Muppa P, Owens V, Smyrk TC, Anderson M, Lutzke LS, Tearney G, Wang KK. Comparative diagnostic performance of volumetric laser endomicroscopy and confocal laser endomicroscopy in the detection of dysplasia associated with Barrett's esophagus. *Gastrointest Endosc* 2016; **83**: 880-888.e2 [PMID: 26344884 DOI: 10.1016/j.gie.2015.08.050]

12 **Struyvenberg MR**, van der Sommen F, Swager AF, de Groof AJ, Rikos A, Schoon EJ, Bergman JJ, de With PHN, Curvers WL. Improved Barrett's neoplasia detection using computer-assisted multiframe analysis of volumetric laser endomicroscopy. *Dis Esophagus* 2020; **33** [PMID: 31364700 DOI: 10.1093/dote/doz065]

13 **Poneros JM**, Brand S, Bouma BE, Tearney GJ, Compton CC, Nishioka NS. Diagnosis of specialized intestinal metaplasia by optical coherence tomography. *Gastroenterology* 2001; **120**: 7-12 [PMID: 11208708 DOI: 10.1053/gast.2001.20911]

14 **Isenberg G**, Sivak MV Jr, Chak A, Wong RC, Willis JE, Wolf B, Rowland DY, Das A, Rollins A. Accuracy of endoscopic optical coherence tomography in the detection of dysplasia in Barrett's esophagus: a prospective, double-blinded study. *Gastrointest Endosc* 2005; **62**: 825-831 [PMID: 16301020 DOI: 10.1016/j.gie.2005.07.048]

15 **Evans JA**, Poneros JM, Bouma BE, Bressner J, Halpern EF, Shishkov M, Lauwers GY, Mino-Kenudson M, Nishioka NS, Tearney GJ. Optical coherence tomography to identify intramucosal carcinoma and high-grade dysplasia in Barrett's esophagus. *Clin Gastroenterol Hepatol* 2006; **4**: 38-43 [PMID: 16431303 DOI: 10.1053/S1542-3565(05)00746-9]

16 **Chen Y**, Aguirre AD, Hsiung PL, Desai S, Herz PR, Pedrosa M, Huang Q, Figueiredo M, Huang SW, Koski A, Schmitt JM, Fujimoto JG, Mashimo H. Ultrahigh resolution optical coherence tomography of Barrett's esophagus: preliminary descriptive clinical study correlating images with histology. *Endoscopy* 2007; **39**: 599-605 [PMID: 17611914 DOI: 10.1055/s-2007-966648]

17 **Cobb MJ**, Hwang JH, Upton MP, Chen Y, Oelschlager BK, Wood DE, Kimmey MB, Li X. Imaging of subsquamous Barrett's epithelium with ultrahigh-resolution optical coherence tomography: a histologic correlation study. *Gastrointest Endosc* 2010; **71**: 223-230 [PMID: 19846077 DOI: 10.1016/j.gie.2009.07.005]

18 **Adler DC**, Zhou C, Tsai TH, Lee HC, Becker L, Schmitt JM, Huang Q, Fujimoto JG, Mashimo H. Three-dimensional optical coherence tomography of Barrett's esophagus and buried glands beneath neosquamous epithelium following radiofrequency ablation. *Endoscopy* 2009; **41**: 773-776 [PMID: 19746317 DOI: 10.1055/s-0029-1215045]

19 **Vakoc BJ**, Shishko M, Yun SH, Oh WY, Suter MJ, Desjardins AE, Evans JA, Nishioka NS, Tearney GJ, Bouma BE. Comprehensive esophageal microscopy by using optical frequency-domain imaging (with video). *Gastrointest Endosc* 2007; **65**: 898-905 [PMID: 17383652 DOI: 10.1016/j.gie.2006.08.009]

20 **Trindade AJ**, George BJ, Berkowitz J, Sejpal DV, McKinley MJ. Volumetric laser endomicroscopy can target neoplasia not detected by conventional endoscopic measures in long segment Barrett's esophagus. *Endosc Int Open* 2016; **4**: E318-E322 [PMID: 27004250 DOI: 10.1055/s-0042-101409]

21 **Jain D**, Fatima S, Jain S, Singhal S. Volumetric Laser Endomicroscopy for Barrett's Esophagus - Looking at the Fine Print. *J Gastrointestin Liver Dis* 2017; **26**: 291-297 [PMID: 28922442 DOI: 10.15403/jgld.2014.1121.263.jai]

22 **Kohli DR**, Schubert ML, Zfass AM, Shah TU. Performance characteristics of optical coherence tomography in assessment of Barrett's esophagus and esophageal cancer: systematic review. *Dis Esophagus* 2017; **30**: 1-8 [PMID: 28881898 DOI: 10.1093/dote/dox049]

23 **Castela J**, Serrano M, Ferro SM, Pereira DV, Chaves P, Pereira AD. Buried Barrett's Esophagus with High-Grade Dysplasia after Radiofrequency Ablation. *Clin Endosc* 2019; **52**: 269-272 [PMID:

30300980 DOI: 10.5946/ce.2018.124]
24 **Spechler SJ**. Buried (but not dead) Barrett's metaplasia: tales from the crypts. *Gastrointest Endosc* 2012; **76**: 41-43 [PMID: 22726464 DOI: 10.1016/j.gie.2012.02.053]
25 **Zhou C**, Tsai TH, Lee HC, Kirtane T, Figueiredo M, Tao YK, Ahsen OO, Adler DC, Schmitt JM, Huang Q, Fujimoto JG, Mashimo H. Characterization of buried glands before and after radiofrequency ablation by using 3-dimensional optical coherence tomography (with videos). *Gastrointest Endosc* 2012; **76**: 32-40 [PMID: 22482920 DOI: 10.1016/j.gie.2012.02.003]
26 **Swager AF**, Boerwinkel DF, de Bruin DM, Faber DJ, van Leeuwen TG, Weusten BL, Meijer SL, Bergman JJ, Curvers WL. Detection of buried Barrett's glands after radiofrequency ablation with volumetric laser endomicroscopy. *Gastrointest Endosc* 2016; **83**: 80-88 [PMID: 26124075 DOI: 10.1016/j.gie.2015.05.028]
27 **Benjamin T**, Shakya S, Thota PN. Feasibility of volumetric laser endomicroscopy in Barrett's esophagus with dysplasia and in post-ablation surveillance. *J Gastrointestin Liver Dis* 2016; **25**: 407-408 [PMID: 27689210 DOI: 10.15403/jgld.2014.1121.253.brt]
28 **Standish BA**, Yang VX, Munce NR, Wong Kee Song LM, Gardiner G, Lin A, Mao YI, Vitkin A, Marcon NE, Wilson BC. Doppler optical coherence tomography monitoring of microvascular tissue response during photodynamic therapy in an animal model of Barrett's esophagus. *Gastrointest Endosc* 2007; **66**: 326-333 [PMID: 17643708 DOI: 10.1016/j.gie.2007.02.040]
29 **Li H**, Standish BA, Mariampillai A, Munce NR, Mao Y, Chiu S, Marcon NE, Wilson BC, Vitkin A, Yang VX. Feasibility of interstitial Doppler optical coherence tomography for in vivo detection of microvascular changes during photodynamic therapy. *Lasers Surg Med* 2006; **38**: 754-761 [PMID: 16927368 DOI: 10.1002/Lsm.20387]
30 **Lee HC**, Ahsen OO, Liang K, Wang Z, Figueiredo M, Giacomelli MG, Potsaid B, Huang Q, Mashimo H, Fujimoto JG. Endoscopic optical coherence tomography angiography microvascular features associated with dysplasia in Barrett's esophagus (with video). *Gastrointest Endosc* 2017; **86**: 476-484.e3 [PMID: 28167119 DOI: 10.1016/j.gie.2017.01.034]
31 **Beaudette K**, Baac HW, Madore WJ, Villiger M, Godbout N, Bouma BE, Boudoux C. Laser tissue coagulation and concurrent optical coherence tomography through a double-clad fiber coupler. *Biomed Opt Express* 2015; **6**: 1293-1303 [PMID: 25909013 DOI: 10.1364/BOE.6.001293]
32 **Uno K**, Koike T, Shimosegawa T. Recent development of optical coherence tomography for preoperative diagnosis of esophageal malignancies. *World J Gastrointest Endosc* 2015; **7**: 872-880 [PMID: 26240688 DOI: 10.4253/wjge.v7.i9.872]
33 **de Groof AJ**, Struyvenberg MR, Fockens KN, van der Putten J, van der Sommen F, Boers TG, Zinger S, Bisschops R, de With PH, Pouw RE, Curvers WL, Schoon EJ, Bergman JJGHM. Deep learning algorithm detection of Barrett's neoplasia with high accuracy during live endoscopic procedures: a pilot study (with video). *Gastrointest Endosc* 2020; **91**: 1242-1250 [PMID: 31926965 DOI: 10.1016/j.gie.2019.12.048]
34 **Trindade AJ**, McKinley MJ, Fan C, Leggett CL, Kahn A, Pleskow DK. Endoscopic Surveillance of Barrett's Esophagus Using Volumetric Laser Endomicroscopy With Artificial Intelligence Image Enhancement. *Gastroenterology* 2019; **157**: 303-305 [PMID: 31078625 DOI: 10.1053/j.gastro.2019.04.048]
35 **Tsai TH**, Fujimoto JG, Mashimo H. Endoscopic Optical Coherence Tomography for Clinical Gastroenterology. *Diagnostics (Basel)* 2014; **4**: 57-93 [PMID: 26852678 DOI: 10.3390/diagnostics4020057]
36 **Kahn A**, Kamboj AK, Muppa P, Sawas T, Lutzke LS, Buras MR, Golafshar MA, Katzka DA, Iyer PG, Smyrk TC, Wang KK, Leggett CL. Staging of T1 esophageal adenocarcinoma with volumetric laser endomicroscopy: a feasibility study. *Endosc Int Open* 2019; **7**: E462-E470 [PMID: 30931378 DOI: 10.1055/a-0838-5326]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 157-167

DOI: 10.37126/aige.v2.i4.157

ISSN 2689-7164 (online)

MINIREVIEWS

Artificial intelligence and colonoscopy − enhancements and improvements

Byung Soo Yoo, Steve M D'Souza, Kevin Houston, Ankit Patel, James Lau, Alsiddig Elmahdi, Parth J Parekh, David Johnson

ORCID number: Byung Soo Yoo 0000-0002-8501-7922; Steve M D'Souza 0000-0003-3772-2616; Kevin Houston 0000-0002-8441-0132; Ankit Patel 0000-0002-6981-3427; James Lau 0000-0001-9606-6829; Alsiddig Elmahdi 0000-0003-4808-7232; Parth J Parekh 0000-0003-4750-775X; David Johnson 0000-0002-8737-0711.

Author contributions: Johnson DA, Parekh PJ, D'Souza SM and Yoo BS contributed to the construction of the project; all authors wrote and edited the manuscript.

Conflict-of-interest statement: Authors have nothing to disclose.

Manuscript source: Invited manuscript

Byung Soo Yoo, Steve M D'Souza, Kevin Houston, Ankit Patel, James Lau, Alsiddig Elmahdi, Department of Medicine, Eastern Virginia Medical School, Norfolk, VA 23507, United States

Parth J Parekh, David Johnson, Division of Gastroenterology, Department of Internal Medicine, Eastern Virginia Medical School, Norfolk, VA 23505, United States

Corresponding author: David Johnson, MD MACG, FASGE, MACP, Professor of Medicine, Chief, Division of Gastroenterology Department of Internal Medicine, Eastern Virginia Medical School, 886 Kempsville Road Suite 114, Norfolk, VA 23505, United States. dajevms@aol.com

Abstract

Artificial intelligence is a technology that processes and analyzes information with reproducibility and accuracy. Its application in medicine, especially in the field of gastroenterology, has great potential to facilitate in diagnosis of various disease states. Currently, the role of artificial intelligence as it pertains to colonoscopy revolves around enhanced polyp detection and characterization. The aim of this article is to review the current and potential future applications of artificial intelligence for enhanced quality of detection for colorectal neoplasia.

Key Words: Artificial intelligence; Colon polyp; Adenoma detection rate; Dysplasia; Inflammatory bowel disease; Colon preparation

Core Tip: The application of artificial intelligence (AI) in medicine and gastroenterology has demonstrated to date, broad utility in both disease diagnostics and management. The utility of AI in colonoscopy has recently demonstrated enhanced polyp detection and characterization, assessment for mucosal healing and identification of dysplasia associated with inflammatory bowel disease, as well as assessment of the quality of bowel preparation for colonoscopy.

Citation: Yoo BS, D'Souza SM, Houston K, Patel A, Lau J, Elmahdi A, Parekh PJ, Johnson D. Artificial intelligence and colonoscopy − enhancements and improvements. *Artif Intell*

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: United States

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C, C
Grade D (Fair): 0
Grade E (Poor): 0

Received: June 5, 2021
Peer-review started: June 5, 2021
First decision: June 18, 2021
Revised: June 21, 2021
Accepted: July 23, 2021
Article in press: July 23, 2021
Published online: August 28, 2021

P-Reviewer: Morya AK, Zhao Y
S-Editor: Liu M
L-Editor: A
P-Editor: Wang LYT

Gastrointest Endosc 2021; 2(4): 157-167
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/157.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.157

INTRODUCTION

Although artificial intelligence (AI) was first conceptually presented as a means for machines to mechanize human actions and cognitive thinking approximately 70 years ago, the current applications are exponentially broad[1,2]. This technology is predicated on the fact that AI is able to exhibit certain facets of human intelligence which is derived from techniques known as machine learning (ML) and deep learning (DL)[3]. Machine learning involves automatically building mathematical algorithms from data sets and forming decisions with or without human supervision[3,4]. When an algorithm is able to learn predictive models, it can use new inputs to form outputs [3,5,6]. These models can be combined to form artificial neural networks (ANN) which mimic the neural network of a brain. Each algorithm assumes the role of a neuron and when grouped together form a network that interacts with different neurons[5,6]. ANN have pathways from inputs to outputs with hidden layers in between to help make the inner nodes more efficient and improve the overall network[3]. DL is a domain in which AI process a vast amount of data and self-creates algorithms that interconnect the nodes of ANN with interplay in the hidden neural layers[3,6]. Researchers have been using DL to form computer aided diagnosis systems (CADS) to aid in polyp detection and characterization[7]. Two major CAD systems have been developed so far: CADe (termed for computer-aided detection) and CADx (termed for computer-aided characterization). CADe uses white-light endoscopy for image analysis with the ultimate goal to increase the number of adenomas found in each colonoscopy thereby increasing adenoma detection rate (ADR) and reducing the rate of missed polyps[8]. CADx is designed to characterize polyps found during colonoscopy, thereby improving the accuracy of optical biopsies and reducing unnecessary polypectomy for non-neoplastic lesions[8]. It predominantly uses magnifying narrow band imaging (mNBI) but could also incorporate a variety of other techniques including white-light endoscopy, magnifying chromoendoscopy, confocal laser endomicroscopy, spectroscopy, and autofluorescence endoscopy[8]. In addition, AI technology is being applied to evaluate the quality of bowel preparation for colonoscopy. In this review, we outline the role of AI in polyp detection and characterization of dysplastic and/or neoplastic lesions. We also provide the current data on utility of AI in evaluation of bowel preparation and future directions of AI in colonoscopy.

POLYP DETECTION AND CADS

Polyps are abnormal tissue growths that arise in the colon that carry malignant potential[9]. Polyps are detected during colonoscopy but can sometimes be missed due to a variety of factors *e.g.*, age of patient, diminutive polyp size, failure to reach cecum, quality of bowel preparation, and experience of endoscopist[10,11]. The ADR is the frequency to detect one or more adenomatous polyps during screening colonoscopy and is a universal quality metric with the strongest association to the development of interval cancers[11-13]. Owing to growing concerns of increasing rates of colon cancer in adults, CADS have been developed and utilized to aid in polyp detection and ultimately increase ADR[9,14-18].

Multiple research groups have created automated computer vision methods to help analyze and detect polyps during colonoscopy[15-19] (Figure 1). One of the first groups to use CADe to help detect polyps relied mainly on still images from videos for analysis and polyp detection[20]. Their CADe used 24 videos containing a total of 31 polyps which were detectable in at least 1 frame[20]. The study demonstrated a sensitivity and specificity for polyp detection of 70.4% and 72.4%, respectively[20]. Another group created a model using DL which used 546 short videos and 73 full length videos to create the software and train it with positive and negative polyp containing videos[21]. The sensitivity and specificity were 90% and 63.3% respectively, showing that the model could potentially be used in a clinical setting to help minimize polyp miss rates during colonoscopy[21]. A recent, prospective multicenter trial

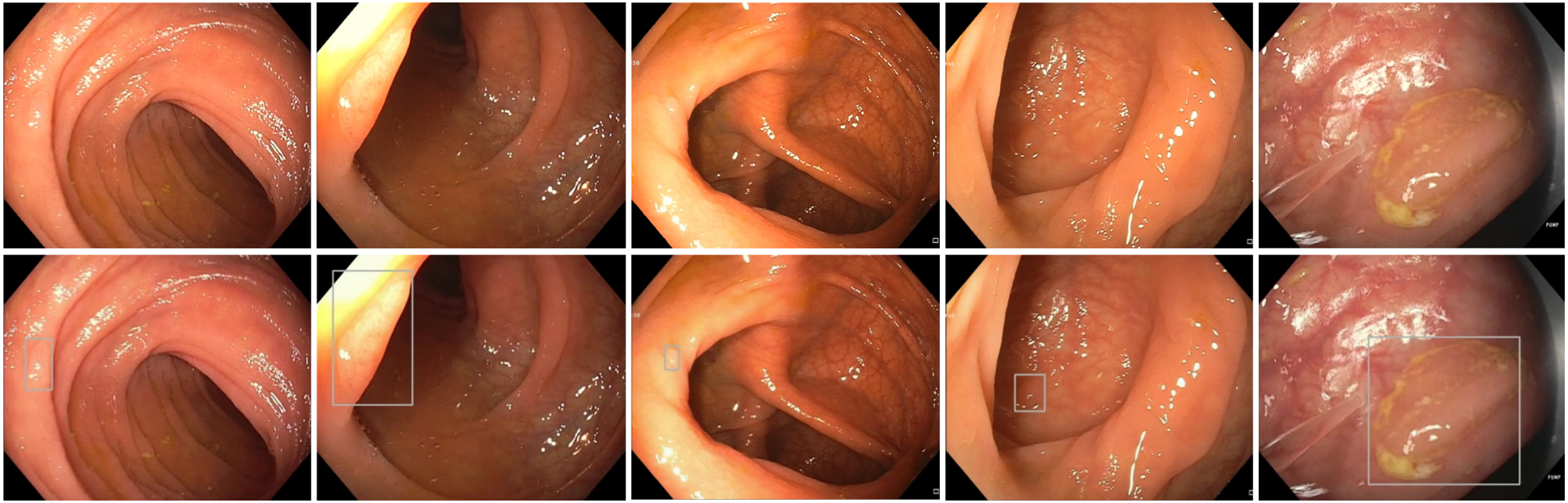

Figure 1 Polyp detection without artificial intelligence (top) and with artificial intelligence (bottom).

comparing a CADe system to trained endoscopists and found that endoscopists (with a baseline ADR ≥ 35%) and CADe had a diagnostic accuracy of 98.2% and 96.5% respectively[22]. This led the authors to conclude that CADe was non-inferior to expert endoscopists[22].

As CADe systems proved to enhance polyp detection, researchers then focused on the role of AI on improving ADR. A prospective, randomized, controlled study evaluated 1058 patients undergoing colonoscopy with or without an automatic polyp detection system (APDS) found a relative ADR increase of 43.3% (29.1% *vs* 20.3%) using the APDS compared to standard colonoscopy[23]. This increase was most prevalent amongst diminutive adenomas which suggests that smaller adenomas are more likely to be missed compared to larger adenomas[23]. To expand upon the previous study, a double-blinded, randomized, controlled trial was performed with a sham group to control for operational bias[24]. There was a 21.4% relative increase of ADR in the CADe group (34% *vs* 28%) when compared to controls[24]. They found the delta to be higher amongst endoscopists with lower baseline ADR than compared to those with a higher baseline ADR[24]. A recent meta-analysis which included 6 randomized controlled trials comparing AI-assisted-colonoscopy to non-AI-assisted-colonoscopy totaling 5058 patients showed a significantly higher ADR within the AI group compared to the control (33.7% *vs* 22.9%, respectively)[25]. The study also

showed an overall increase in detecting proximal colon adenomas with the AI-assisted-group compared to the control group (23.4% *vs* 14.5%, respectively)[25]. This is important because currently colorectal cancer (CRC) screening with colonoscopy alone is not effective at reducing proximal colon cancers and their mortality[25,26]. Thus while improving ADR is vital to preventing CRC, particularly in the proximal colon, the use of AI alongside endoscopists can be an ideal starting point. The CADe systems could be used as second observers, as second observers have been shown to increase ADR[27].

POLYP CHARACTERIZATION AND AI

Worldwide, CRC is the third most common cancer diagnosed in men and second in women[28]. Overall incidence of CRC in the United States has decreased due to lower smoking rates, early colonoscopy screenings, and early identification of patient-specific risk factors, but recent studies have reported a global increase in incidence of CRC in the younger population[29,30]. Thus, the latest endoscopic research is aimed towards techniques to better identify polyps and allow for real-time polyp histologic characterization which provides vital information for early intervention through endoscopic or surgical resection[31].

Studies evaluating AI and histologic assessment with optical biopsy have been a targeted focus- in particular for a "resect and discard" strategy for diminutive polyps < 6 mm, thereby avoiding the costs of pathology for low risk lesions[32] (Figure 2) top.

Several studies have found the range of sensitivity and specificity for polyp detection and characterization to be 70%-98% and 63%-98%, respectively[33]. An optical biopsy allows for differentiation of polyp type based on certain features. For example, NBI is an image-enhanced type of endoscopy that is used to identify microstructures and capillaries of the mucosal epithelium and allow for prediction of histologic features of colorectal polyps. Use of this advanced imaging technique often requires expertise to differentiate hyperplastic polyps from neoplastic polyps with high accuracy. AI systems offer a standardization of polyp characterization that overcomes the expertise or training differences across endoscopists[34]. Analysis of a CAD system with a deep neural network for analyzing NBI of diminutive polyps found that the AI system could identify neoplastic or hyperplastic polyps with 96.3% sensitivity and 78.1% specificity[34]. The system was compared to both novice (in-training) and expert endoscopists and it was notable that over half of the novice endoscopists classified polyps with a negative-predictive value of ranging from 73%-84%, compared to 91.5% of the system. The system also had a shorter time-to-classification compared to both expert and novice endoscopists ($P < 0.05$)[34]. Other groups have had similar results showing promise for AI-identification. One study compared images of 225 polyps as evaluated by a CAD system compared to diagnosis by endoscopists[35]. The polyps were classified using the Kudo and NBI international colorectal endoscopic classifications which found of the 225 polyps, 142 were dysplastic and 83 were non-dysplastic after endoscopy. The results of the CAD system correctly classified 205 polyps (91.1% of the total) and correctly delineated 131/142 (92%) as dysplastic and 74/83 (89%) as non-dysplastic[35]. There were no statistically significant differences in histologic prediction between the CAD system and endoscopic assessment, thus they concluded that a computer vision system based on characterization of the polyp surface could accurately predict polyp histology[35].

AI IN INFLAMMATORY BOWEL DISEASE

Inflammatory bowel disease (IBD), which includes Crohn's disease (CD) and ulcerative colitis (UC), is a chronic inflammatory gastrointestinal tract disorder that remains a global concern as incidence in developing countries continues to grow[36]. Studies with AI and large datasets of endoscopic images have shown that AI can improve the way to diagnose IBD, evaluate the severity of disease, and follow-up treatments and provide follow-up[37]. Initial diagnosis of IBD through endoscopic evaluation remains a challenge due to wide ranging clinical manifestations of IBD and overlap across subtypes. Key endoscopic features of IBD include ulceration or erosions, and AI has shown its role in better predicting the need for further evaluation [38]. Aoki *et al*[38] have demonstrated that a deep convolutional neural network (DCNN) can be trained to detect erosions and ulcerations seen on wireless capsule endoscopy. Their system evaluated 10440 images in 233 s and demonstrated an area

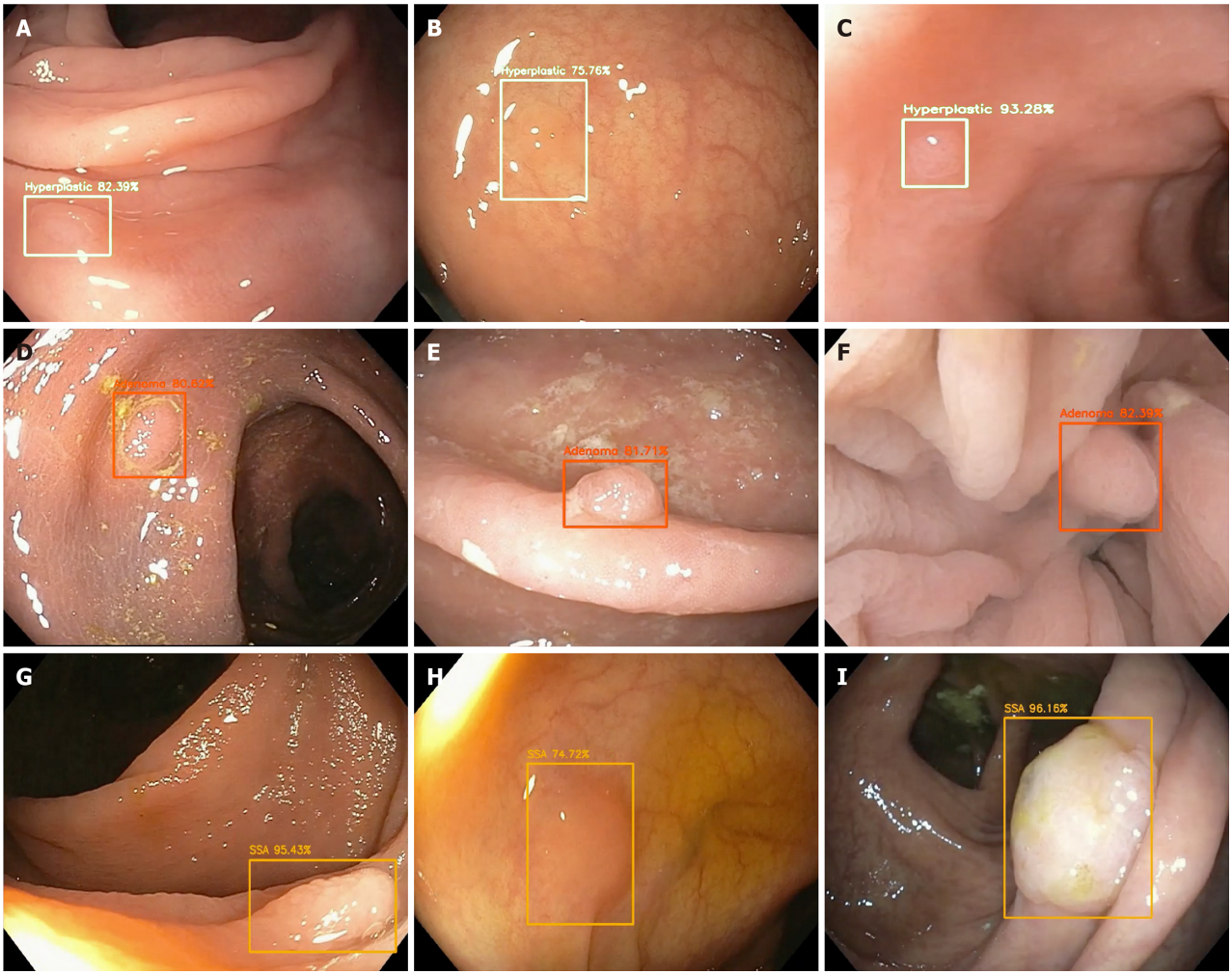

Figure 2 Optical pathology of detected polyps with associated probability utilizing artificial intelligence. A-C: Hyperplastic polyps; D-F: Adenomas; G-I: Sessile serrated adenomas.

under the curve for detection of erosions and ulcerations at 0.958 (95% confidence interval: 0.947-0.968) and sensitivity, specificity and accuracy of 88%, 90% and 90%, respectively[38]. Tong *et al*[39] studied 6399 patients with UC, CD, or intestinal tuberculosis (ITB) who underwent colonoscopies. The colonoscopic images were then translated in the form of free texts and Random Forest (RF) and CNN were utilized to distinguish the three diseases. Diagnostic sensitivity and specificity of RF in UC/CD/ITB were 0.89/0.84, 0.83/0.82, and 0.72/0.77, respectively and that of CNN were 0.99/0.97, 0.87/0.83, and 0.52/0.81, respectively[39]. The studies showed that AI can be employed to discern and diagnose IBD although real-time diagnostic utility remains an area to develop[39].

Determining disease severity and activity in IBD can be done using endoscopic inflammation indices, and histologic scores. However, there can be certain flaws to using these methodologies such as intra-observer and inter-observer variability[40]. Studies using AI have been done to help control some of these factors. Bossuyt *et al*[41] developed a red density (RD) system, which was specific for endoscopic and histologic disease activity in UC patients, to help mitigate the observer bias by endoscopists. The study had 29 UC patients compared against 6 control patients using the RD score gained during colonoscopy[41]. The RD score was linked to the Robart's Histologic Index in a multiple regression analysis and was found to be correlated with the RHI (r = 0.65, P < 0.00002) from the patients with UC[41]. The RD score from the control patients was also correlated with the RHI, Mayo endoscopic subscores (r = 0.76, P < 0.0001) and UC Endoscopic Index of Severity scores (r = 0.74, P < 0.0001), showing it correlated well with the validated tests[41]. A study done by Takenaka *et al*[40] used their algorithm, the deep neural network for evaluation of UC (DNUC), in 875 UC patients. The DNUC was developed using 40785 images from colonoscopies and 6885 biopsy results from 2012 UC patients[40]. The DNUC was able to identify patients in endoscopic remission with 90.1% accuracy and a kappa coefficient of 0.798 and

identify patients in histologic remission with 92.2% accuracy and a kappa coefficient of 0.895 between the biopsy result and the DNUC[40]. The researchers concluded that it could be used to identify patients in remission and potentially avoid mucosal biopsy and analysis[40]. Stidham *et al*[42] created a 159-layer CNN using 16514 images from 3082 UC patients to help categorize patients groups in remission (Mayo subscore 0 or 1) to moderate to severe (Mayo subscore 2 or 3). The CNN had a positive predictive value of 0.87, sensitivity 83% and specificity of 96%[42]. The CNN was compared against human reviewers when assigned the Mayo scores, with a kappa coefficient of 0.84 for the CNN *vs* 0.86 for the human reviewers[42]. This shows that the AI is effectively able to help categorize patients into their respective severity stages[42].

Patients with IBD are at increased risk for CRC and it is important for these patients to undergo frequent surveillance. Guidelines differ depending on the medical society, but overall recommended intervals are from 1-5 years[43]. IBD surveillance guidelines and whether AI has a role in CRC detection has yet to be directly studied. A large reason for the lack of studies of AI and IBD has been due to IBD being an exclusion criterion for many of the early colonoscopic AI studies. A single study by Uttam *et al*[44] was one of the first to look at IBD and cancer risk, utilizing a three-dimensional nanoscale nuclear architecture mapping (nanoNAM). By analyzing 103 patients with IBD that were undergoing colonoscopy, their system measured for submicroscopic alterations in the intrinsic nuclear structure within epithelial cells and compared findings to histologic biopsies after 3 years. They found that their nanoNAM could identify colonic neoplasia with an AUC of 0.87, sensitivity of .81, and specificity of 0.82[44]. Additional studies on AI in IBD surveillance could help personalize surveillance strategies or guidelines for patients.

AI SYSTEMS IN PRACTICE

The most recent developments in clinical practice have been with the approval of several different devices: EndoBRAIN (Olympus Corporation, Tokyo, Japan), GI Genius (Cosmo Pharmaceuticals N.V., Dublin, Ireland), and WavSTAT4 (SpectraScience, Inc., San Diego, CA)[33,45,46]. EndoBRAIN is an AI-based system that is able to analyze pathologic features present on endoscopic imaging, and was developed and approved as a class II medical device[33]. In a multi-center study to determine the diagnostic accuracy of EndoBRAIN, their system was trained using 69142 endocytoscopic images taken from patients that had undergone endoscopy and the EndoBRAIN was compared against 30 endoscopists (20 trainees, 10 experts) with primary outcome of assessing neoplastic *vs* non-neoplastic lesions. Their results found that EndoBRAIN distinguished neoplastic from non-neoplastic lesions with 96.9% sensitivity, 94.3% specificity, which was higher than trainees and comparable to experts[33].

GI Genius has been approved by the FDA as an AI device to detect colonic lesions. GI Genius was compared to experienced endoscopists for colorectal polyp detection[45]. This system was trained on a data-set using white-light endoscopy videos in a randomized controlled trial and primarily looked at reaction time on a lesion as the primary endpoint. Results demonstrated that the AI system held a faster reaction time when compared with endoscopists in 82% of cases[45].

Lastly, laser-induced fluorescence spectroscopy using a WavSTAT4 optical biopsy system was evaluated for efficacy in accurately assessing the histology of colorectal polyps with the end goal of reducing time, costs, and risks of resecting diminutive colorectal polyps[46]. The overall accuracy of predicting polyp histology was 84.7%, sensitivity of 81.8%, specificity of 85.2%, and negative predictive value of 96.1%. This suggests that the system is accurate enough to allow distal colorectal polyps to be left in place and nearly reaches the American Society for Gastrointestinal Endoscopy threshold for resecting and discarding without pathologic assessment[46].

REAL-TIME EVALUATION FOR INVASIVE CANCER

AI prediction of invasive cancers through the utilization of real-time identification of colorectal polyps has the potential to improve CRC screening by limiting misses and improving outcomes, especially in geographic regions with less access to highly trained endoscopists.

Advanced imaging techniques during endoscopy (without AI) to provide a real-time prediction of lesion pathology and depth of invasion has been widely used. For example, a study assessed white-light endoscopy, mNBI, magnifying chromoen-

doscopy, and probe-based confocal laser endomicroscopy in real-time, in order to evaluate and classify the depth of invasion for colorectal lesions[31]. Of the 22 colorectal lesions, 7 were adenomas, 10 were intramucosal cancers, and 5 had deep submucosal invasion or deeper involvement. Sensitivity and specificity of white light endoscopy and mNBI were both 60% and 94%, respectively. Magnifying chromoendoscopy and probe-based confocal laser endomicroscopy were both 80 and 94%, respectively[31].

With data showing reliability of advanced imaging techniques in real-time for information to establish a diagnosis and drive intervention pursuits, integration of AI systems with these advanced imaging techniques has been a growing research focus. A recent review assessed 5 retrospective studies with wide ranging sensitivities ranging from 67.5%-88.2% sensitivity and 77.9%-98.9% specificity in finding invasive cancers[47]. The prediction of cancer invasion was made using magnified NBI, confocal laser endomicroscopy, white light endoscopy, or endocytoscopy. As the numbers reflect, more studies are needed to better evaluate how AI can provide more stable reliability in evaluation for invasive cancers[47].

COLON PREPARATION AND AI

Bowel preparation significantly impacts the diagnostic accuracy of colonoscopies. Inadequate colon preparation impairs visualization of the mucosa, thus causing missed lesions, extended operative time, and increased need for repeat colonoscopies [48,49]. Approximately 10%-25% of all colonoscopies are inadequately prepared[50-52]. In addition, studies have shown that suboptimal bowel preparation can result in an adenoma miss rate ranging from 35%-42%[51]. A recent prospective study discovered that variable bowel preparation quality did not have a measurable effect on their AI algorithm's ability to accurately identify colonic polyps. However, the applicability of these findings is limited by the study's small sample size of 50[50]. Therefore, the ability of AI to accurately identify polyps in suboptimal conditions remains unknown.

Currently several scales, the most validated and reliable of which is the Boston Preparation Scale (BBPS), are used to assess bowel preparation quality[52]. Scores ranging from 0-3 are individually given to the right, transverse, and left colon during colonoscope withdrawal. A bowel preparation that fails to have a total BPPS score of ≥ 2 would mandate a repeat colonoscopy before the recommended 10-year interval (assuming a normal colon)[48,52]. Despite BBPS being deemed the most reliable scale, it cannot accurately account for variability in bowel preparation throughout the entire colon or gradients in adequacy of cleansing. Although BBPS takes into consideration the 3 colonic segments, regions of the same segment can be variably cleansed[49,52]. Therefore, 1 score cannot accurately represent one-third of the colon. This limitation is further exacerbated by the scale's susceptibility to subjectivity, as individual experiences can shape how physicians interpret data[49].

Most studies indicating the efficacy of AI in detecting colonic polyps utilized still images and videos of ideally prepared colons to train and test their AI software[53]. A DCNN known as ENDOANGEL (Wuhan EndoAngel Medical Technology Company, Wuhan, China) provided real-time and objective BPPS scores during the colonoscopy. ENDOANGEL circumvents subjective bias *via* DL using thousands of images scored by different endoscopists[49]. Additionally, the DCNN simultaneously calculates a real-time BPPS score every 30 s throughout the colonoscopy and provides a cumulative ratio of the different stages, thus providing an accurate assessment of preparation quality throughout the colon[49,52]. Through DL and frequent scoring, ENDOANGEL proved to be far more effective than endoscopists at accurate BPPS scoring (93.33% *vs* 75.91%)[49].

Overall, poor bowel preparation quality significantly increases ADR[51]. Although previous applications of CADe and CADx have been used to optimize endoscopic image quality and mucosal visualization, ENDOANGEL, is the first utilization of AI to provide objective, real-time assessments of bowel preparation quality throughout the entire colon[49,54,55]. Another laboratory group has since independently released promising results regarding use of their AI to assess bowel preparation, indicating AI's potential to improve the preparatory-phase of colonoscopy[56].

FUTURE DIRECTIONS

A significant problem in the advanced imaging trials is that these are done by experts and accordingly, there is good inter-observer performance characteristics. These results are not the same when evaluated by lesser experienced providers[57]. Application of AI as a formidable tool seems logical and promising to mitigate the costs and learning curves for application of these newer techniques across the broad and variable ranges of providers.

Although the current CADS provide promising results, a larger data sets for training the systems can provide improvements in sensitivity and specificity in addition to minimizing false positives and false negatives. The larger training data also increases the burden of annotations, however, this can be overcome by an annotation software which incorporates a DL module. The precise effects of AI once it is widely available in clinical practices are yet to be determined, but the evidence based on EndoBRAIN, GI Genius, and WavSTAT4 are hopeful that significant benefits in training gastroenterologist and diagnosing a polyp can be expected.

Additional areas of future study include better detection of various polyps (adenomatous, non-adenomatous, dysplastic), evaluation of lesion size and morphology, and distinguishing invasive involvement. Additionally, further study is necessary to evaluate the adequacy of large polyp resection (*i.e.*, margins free of adenomatous change). Much of the early data to date have used AI systems which are based on algorithms using still-images and videos[58]. Larger-scale studies can help us better understand real-time use of AI to show how it compares to endoscopists. Due to the novelty of AI systems in the clinical setting, study methods utilizing AI have also largely been done in a non-blinded manner, which may interfere with how the endoscopists perform the procedure, leading to a component of observation bias.

Finally, the future of AI lies in simplifying the tool for utilization by many endoscopists as well as achieving the goal of treatment. One way to overcome the complexity is incorporating the CADS into the colonoscope and display instead of existing as a separate entity that needs to be installed. In addition, an improved model for distinguishing polyps and invasion can further facilitate treatment process for patients.

CONCLUSION

AI is widely applied and utilized in endoscopy and continues to be researched to augment the accuracy of screening and differentiation of neoplastic *vs* non-neoplastic lesions. Although this wide applicability and active investigations are encouraging, further work is needed to solidify the integration of AI into everyday practice. Real-time diagnosis using AI remains technically challenging, however, these recent studies exemplify promising advancements for enhanced quality assessment and management of colonic disease.

REFERENCES

1 **McCarthy J**, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth summer research project on artificial intelligence. [cited 10 May 2021]. Available from: http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

2 **Kaplan A**, Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? *Bus Horiz* 2019; **62**: 15-25 [DOI: 10.1016/j.bushor.2018.08.004]

3 **Yang YJ**, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; **25**: 1666-1683 [PMID: 31011253 DOI: 10.3748/wjg.v25.i14.1666]

4 **Deo RC**. Machine learning in medicine. *Circulation* 2015; **132**: 1920-1930 [PMID: 26572668 DOI: 10.1161/CIRCULATIONAHA.115.001593]

5 **Lee JG**, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N. Deep learning in medical imaging: general overview. *Korean J Radiol* 2017; **18**: 570-584 [PMID: 28670152 DOI: 10.3348/kjr.2017.18.4.570]

6 **Min JK**, Kwak MS, Cha JM. Overview of deep learning in gastrointestinal endoscopy. *Gut Liver* 2019; **13**: 388-393 [PMID: 30630221 DOI: 10.5009/gnl18384]

7 **Hoerter N**, Gross SA, Liang PS. Artificial intelligence and polyp detection. *Curr Treat Options Gastroenterol* 2020; epub ahead of print [PMID: 31960282 DOI: 10.1007/s11938-020-00274-2]

8 **Vinsard DG**, Mori Y, Misawa M, Kudo SE, Rastogi A, Bagci U, Rex DK, Wallace MB. Quality assurance of computer-aided detection and diagnosis in colonoscopy. *Gastrointest Endosc* 2019; **90**: 55-63 [PMID: 30926431 DOI: 10.1016/j.gie.2019.03.019]

9 **Ribeiro E**, Uhl A, Wimmer G, Häfner M. Exploring deep learning and transfer learning for colonic polyp classification. *Comput Math Methods Med* 2016; **2016**: 6584725 [PMID: 27847543 DOI: 10.1155/2016/6584725]

10 **Ahn SB**, Han DS, Bae JH, Byun TJ, Kim JP, Eun CS. The miss rate for colorectal adenoma determined by quality-adjusted, back-to-back colonoscopies. *Gut Liver* 2012; **6**: 64-70 [PMID: 22375173 DOI: 10.5009/gnl.2012.6.1.64]

11 **Liem B**, Gupta N. Adenoma detection rate: the perfect colonoscopy quality measure or is there more? *Transl Gastroenterol Hepatol* 2018; **3**: 19 [PMID: 29682626 DOI: 10.21037/tgh.2018.03.04]

12 **Rex DK**, Schoenfeld PS, Cohen J, Pike IM, Adler DG, Fennerty MB, Lieb JG 2nd, Park WG, Rizk MK, Sawhney MS, Shaheen NJ, Wani S, Weinberg DS. Quality indicators for colonoscopy. *Gastrointest Endosc* 2015; **81**: 31-53 [PMID: 25480100 DOI: 10.1016/j.gie.2014.07.058]

13 **Corley DA**, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 1298-1306 [PMID: 24693890 DOI: 10.1056/NEJMoa1309086]

14 **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of real-time computer-aided detection of colorectal neoplasia in a randomized trial. *Gastroenterology* 2020; **159**: 512-520.e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

15 **Liu WN**, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, Huang J. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; **26**: 13-19 [PMID: 31898644 DOI: 10.4103/sjg.SJG_377_19]

16 **Pannala R**, Krishnan K, Melson J, Parsi MA, Schulman AR, Sullivan S, Trikudanathan G, Trindade AJ, Watson RR, Maple JT, Lichtenstein DR. Artificial intelligence in gastrointestinal endoscopy. *VideoGIE* 2020; **5**: 598-613 [PMID: 33319126 DOI: 10.1016/j.vgie.2020.08.013]

17 **Tajbakhsh N**, Gurudu SR, Liang J. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Trans Med Imaging* 2016; **35**: 630-644 [PMID: 26462083 DOI: 10.1109/TMI.2015.2487997]

18 **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**: 164-179 [PMID: 25952076 DOI: 10.1016/j.cmpb.2015.04.002]

19 **Lee J**, Wallace MB. State of the Art: The impact of artificial intelligence in endoscopy 2020. *Curr Gastroenterol Rep* 2021; **23**: 7 [PMID: 33855659 DOI: 10.1007/s11894-021-00810-9]

20 **Fernández-Esparrach G**, Bernal J, López-Cerón M, Córdova H, Sánchez-Montes C, Rodríguez de Miguel C, Sánchez FJ. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy* 2016; **48**: 837-842 [PMID: 27285900 DOI: 10.1055/s-0042-108434]

21 **Misawa M**, Kudo SE, Mori Y, Cho T, Kataoka S, Yamauchi A, Ogawa Y, Maeda Y, Takeda K, Ichimasa K, Nakamura H, Yagawa Y, Toyoshima N, Ogata N, Kudo T, Hisayuki T, Hayashi T, Wakamura K, Baba T, Ishida F, Itoh H, Roth H, Oda M, Mori K. Artificial intelligence-assisted polyp detection for colonoscopy: Initial experience. *Gastroenterology* 2018; **154**: 2027-2029 [PMID: 29653147 DOI: 10.1053/j.gastro.2018.04.003]

22 **Sinonquel P**, Eelbode T, Hassan C, Antonelli G, Filosofi F, Neumann H, Demedts I, Roelandt P, Maes F, Bisschops R. Real-time unblinding for validation of a new CADe tool for colorectal polyp detection. *Gut* 2021; **70**: 641-643 [PMID: 33046559 DOI: 10.1136/gutjnl-2020-322491]

23 **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

24 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

25 **Ashat M**, Klair JS, Singh D, Murali AR, Krishnamoorthi R. Impact of real-time use of artificial intelligence in improving adenoma detection during colonoscopy: A systematic review and meta-analysis. *Endosc Int Open* 2021; **9**: E513-E521 [PMID: 33816771 DOI: 10.1055/a-1341-0457]

26 **Nakagawa-Senda H**, Hori M, Matsuda T, Ito H. Prognostic impact of tumor location in colon cancer: the Monitoring of Cancer Incidence in Japan (MCIJ) project. *BMC Cancer* 2019; **19**: 431 [PMID: 31072372 DOI: 10.1186/s12885-019-5644-y]

27 **Aslanian HR**, Shieh FK, Chan FW, Ciarleglio MM, Deng Y, Rogart JN, Jamidar PA, Siddiqui UD. Nurse observation during colonoscopy increases polyp detection: a randomized prospective study. *Am J Gastroenterol* 2013; **108**: 166-172 [PMID: 23381064 DOI: 10.1038/ajg.2012.237]

28 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]

29 **Aran V**, Victorino AP, Thuler LC, Ferreira CG. Colorectal cancer: epidemiology, disease mechanisms and interventions to reduce onset and mortality. *Clin Colorectal Cancer* 2016; **15**: 195-203 [PMID: 26964802 DOI: 10.1016/j.clcc.2016.02.008]

30 **Siegel RL**, Jakubowski CD, Fedewa SA, Davis A, Azad NS. Colorectal cancer in the young: epidemiology, prevention, management. *Am Soc Clin Oncol Educ Book* 2020; **40**: 1-14 [PMID: 32315236 DOI: 10.1200/EDBK_279901]

31 **Abe S**, Saito Y, Oono Y, Tanaka Y, Sakamoto T, Yamada M, Nakajima T, Matsuda T, Ikematsu H, Yano T, Sekine S, Kojima M, Yamagishi H, Kato H. Pilot study on probe-based confocal laser endomicroscopy for colorectal neoplasms: an initial experience in Japan. *Int J Colorectal Dis* 2018; **33**: 1071-1078 [PMID: 29700599 DOI: 10.1007/s00384-018-3059-x]

32 **Rex DK**. Can we do resect and discard with artificial intelligence-assisted colon polyp "optical biopsy? *Tech Innov Gastrointest Endosc* 2020; **22**: 52-55

33 **Kudo SE**, Misawa M, Mori Y, Hotta K, Ohtsuka K, Ikematsu H, Saito Y, Takeda K, Nakamura H, Ichimasa K, Ishigaki T, Toyoshima N, Kudo T, Hayashi T, Wakamura K, Baba T, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Artificial intelligence-assisted system improves endoscopic identification of colorectal neoplasms. *Clin Gastroenterol Hepatol* 2020; **18**: 1874-1881 [PMID: 31525512 DOI: 10.1016/j.cgh.2019.09.009]

34 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate classification of diminutive colorectal polyps using computer-aided analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: 29042219 DOI: 10.1053/j.gastro.2017.10.010]

35 **Sánchez-Montes C**, Sánchez FJ, Bernal J, Córdova H, López-Cerón M, Cuatrecasas M, Rodríguez de Miguel C, García-Rodríguez A, Garcés-Durán R, Pellisé M, Llach J, Fernández-Esparrach G. Computer-aided prediction of polyp histology on white light colonoscopy using surface pattern analysis. *Endoscopy* 2019; **51**: 261-265 [PMID: 30360010 DOI: 10.1055/a-0732-5250]

36 **Ng SC**, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, Panaccione R, Ghosh S, Wu JCY, Chan FKL, Sung JJY, Kaplan GG. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 2017; **390**: 2769-2778 [PMID: 29050646 DOI: 10.1016/S0140-6736(17)32448-0]

37 **Gubatan J**, Levitte S, Patel A, Balabanis T, Wei MT, Sinha SR. Artificial intelligence applications in inflammatory bowel disease: emerging technologies and future directions. *World J Gastroenterol* 2021; **27**: 1920-1935 [PMID: 34007130 DOI: 10.3748/wjg.v27.i17.1920]

38 **Aoki T**, Yamada A, Aoyama K, Saito H, Tsuboi A, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Tanaka S, Koike K, Tada T. Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest Endosc* 2019; **89**: 357-363.e2 [PMID: 30670179 DOI: 10.1016/j.gie.2018.10.027]

39 **Tong Y**, Lu K, Yang Y, Li J, Lin Y, Wu D, Yang A, Li Y, Yu S, Qian J. Can natural language processing help differentiate inflammatory intestinal diseases in China? *BMC Med Inform Decis Mak* 2020; **20**: 248 [PMID: 32993636 DOI: 10.1186/s12911-020-01277-w]

40 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and validation of a deep neural network for accurate evaluation of endoscopic images from patients with ulcerative colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: 32060000 DOI: 10.1053/j.gastro.2020.02.012]

41 **Bossuyt P**, Nakase H, Vermeire S, de Hertogh G, Eelbode T, Ferrante M, Hasegawa T, Willekens H, Ikemoto Y, Makino T, Bisschops R. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut* 2020; **69**: 1778-1786 [PMID: 31915237 DOI: 10.1136/gutjnl-2019-320056]

42 **Stidham RW**, Liu W, Bishu S, Rice MD, Higgins PDR, Zhu J, Nallamothu BK, Waljee AK. Performance of a deep learning model *vs* human reviewers in grading endoscopic disease severity of patients with ulcerative colitis. *JAMA Netw Open* 2019; **2**: e193963 [PMID: 31099869 DOI: 10.1001/jamanetworkopen.2019.3963]

43 **Clarke WT**, Feuerstein JD. Colorectal cancer surveillance in inflammatory bowel disease: Practice guidelines and recent developments. *World J Gastroenterol* 2019; **25**: 4148-4157 [PMID: 31435169 DOI: 10.3748/wjg.v25.i30.4148]

44 **Uttam S**, Hashash JG, LaFace J, Binion D, Regueiro M, Hartman DJ, Brand RE, Liu Y. Three-dimensional nanoscale nuclear architecture mapping of rectal biopsies detects colorectal neoplasia in patients with inflammatory bowel disease. *Cancer Prev Res (Phila)* 2019; **12**: 527-538 [PMID: 31164345 DOI: 10.1158/1940-6207.CAPR-19-0024]

45 **Hassan C**, Wallace MB, Sharma P, Maselli R, Craviotto V, Spadaccini M, Repici A. New artificial intelligence system: first validation study *vs* experienced endoscopists for colorectal polyp detection. *Gut* 2020; **69**: 799-800 [PMID: 31615835 DOI: 10.1136/gutjnl-2019-319914]

46 **Rath T**, Tontini GE, Vieth M, Nägel A, Neurath MF, Neumann H. In vivo real-time assessment of colorectal polyp histology using an optical biopsy forceps system based on laser-induced fluorescence spectroscopy. *Endoscopy* 2016; **48**: 557-562 [PMID: 27009081 DOI: 10.1055/s-0042-102251]

47 **Mori Y**, Kudo SE, Misawa M, Takeda K, Kudo T, Itoh H, Oda M, Mori K. How far will clinical application of AI applications advance for colorectal cancer diagnosis? *J Anus Rectum Colon* 2020; **4**: 47-50 [PMID: 32346642 DOI: 10.23922/jarc.2019-045]

48 **Lai EJ**, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. *Gastrointest Endosc* 2009; **69**: 620-625 [PMID: 19136102 DOI: 10.1016/j.gie.2008.05.057]

49 **Zhou J**, Wu L, Wan X, Shen L, Liu J, Zhang J, Jiang X, Wang Z, Yu S, Kang J, Li M, Hu S, Hu X, Gong D, Chen D, Yao L, Zhu Y, Yu H. A novel artificial intelligence system for the assessment of

bowel preparation (with video). *Gastrointest Endosc* 2020; **91**: 428-435.e2 [PMID: 31783029 DOI: 10.1016/j.gie.2019.11.026]

50 **Millien VO**, Mansour NM. Bowel Preparation for Colonoscopy in 2020: A look at the past, present, and future. *Curr Gastroenterol Rep* 2020; **22**: 28 [PMID: 32377915 DOI: 10.1007/s11894-020-00764-4]

51 **Lebwohl B**, Kastrinos F, Glick M, Rosenbaum AJ, Wang T, Neugut AI. The impact of suboptimal bowel preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. *Gastrointest Endosc* 2011; **73**: 1207-1214 [PMID: 21481857 DOI: 10.1016/j.gie.2011.01.051]

52 **Kastenberg D**, Bertiger G, Brogadir S. Bowel preparation quality scales for colonoscopy. *World J Gastroenterol* 2018; **24**: 2833-2843 [PMID: 30018478 DOI: 10.3748/wjg.v24.i26.2833]

53 **Becq A**, Chandnani M, Bharadwaj S, Baran B, Ernest-Suarez K, Gabr M, Glissen-Brown J, Sawhney M, Pleskow DK, Berzin TM. Effectiveness of a deep-learning polyp detection system in prospectively collected colonoscopy videos with variable bowel preparation quality. *J Clin Gastroenterol* 2020; **54**: 554-557 [PMID: 31789758 DOI: 10.1097/MCG.0000000000001272]

54 **Byrne MF**, Shahidi N, Rex DK. Will computer-aided detection and diagnosis revolutionize colonoscopy? *Gastroenterology* 2017; **153**: 1460-1464.e1 [PMID: 29100847 DOI: 10.1053/j.gastro.2017.10.026]

55 **Stanek SR**, Tavanapong W, Wong J, Oh J, Nawarathna RD, Muthukudage J, de Groen PC. SAPPHIRE: a toolkit for building efficient stream programs for medical video analysis. *Comput Methods Programs Biomed* 2013; **112**: 407-421 [PMID: 24001925 DOI: 10.1016/j.cmpb.2013.07.028]

56 **Thakkar S**, Carleton NM, Rao B, Syed A. Use of artificial intelligence-based analytics from live colonoscopies to optimize the quality of the colonoscopy examination in real time: Proof of concept. *Gastroenterology* 2020; **158**: 1219-1221.e2 [PMID: 31945357 DOI: 10.1053/j.gastro.2019.12.035]

57 **Picot J**, Rose M, Cooper K, Pickett K, Lord J, Harris P, Whyte S, Böhning D, Shepherd J. Virtual chromoendoscopy for the real-time assessment of colorectal polyps in vivo: a systematic review and economic evaluation. *Health Technol Assess* 2017; **21**: 1-308 [PMID: 29271339 DOI: 10.3310/hta21790]

58 **Wu J**, Chen J, Cai J. Application of artificial intelligence in gastrointestinal endoscopy. *J Clin Gastroenterol* 2021; **55**: 110-120 [PMID: 32925304 DOI: 10.1097/MCG.0000000000001423]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 168-178

DOI: 10.37126/aige.v2.i4.168

ISSN 2689-7164 (online)

MINIREVIEWS

Impact of endoscopic ultrasound elastography in pancreatic lesion evaluation

Cosmas Rinaldi Adithya Lesmana, Maria Satya Paramitha

ORCID number: Cosmas Rinaldi Adithya Lesmana 0000-0002-8218-5971; Maria Satya Paramitha 0000-0003-0614-0517.

Author contributions: Lesmana CRA build the idea and wrote the main manuscript; Paramitha MS involved in writing the manuscript.

Conflict-of-interest statement: Authors have no conflict of interest.

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin:

Cosmas Rinaldi Adithya Lesmana, Maria Satya Paramitha, Internal Medicine, Hepatobiliary Division, Dr. Cipto Mangunkusumo National General Hospital, Universitas Indonesia, Jakarta 10430, Daerah Khusus Ibukota, Indonesia

Cosmas Rinaldi Adithya Lesmana, Digestive Disease & GI Oncology Center, Medistra Hospital, Jakarta 12950, Daerah Khusus Ibukota, Indonesia

Corresponding author: Cosmas Rinaldi Adithya Lesmana, FACG, FACP, MD, Senior Lecturer, Internal Medicine, Hepatobiliary Division, Dr. Cipto Mangunkusumo National General Hospital, Universitas Indonesia, Jl. Diponegoro 71 Jakarta, Jakarta 10430, Daerah Khusus Ibukota, Indonesia. medicaldr2001id@yahoo.com

Abstract

Pancreatic malignancy still becomes a major global problem and is considered as one of the most lethal cancers in the field of gastroenterology. Most patients come in the late stage of the disease due to organ's location, and until now the treatment result is still far away from satisfaction. Early detection is still the main key for good, prolonged survival. However, discerning from other types of tumor sometimes is not easy. Endoscopic ultrasound (EUS) is still the best tool for pancreatic assessment, whereas fine-needle aspiration biopsy (FNAB) is considered as the cornerstone for further management of pancreatic malignancy. Several conditions have become a concern for EUS-FNAB procedure, such as risk of bleeding, pancreatitis, and even needle track-seeding. Recently, an artificial intelligence innovation, such as EUS elastography has been developed to improve diagnostic accuracy in pancreatic lesions evaluation. Studies have shown the promising results of EUS elastography in improving diagnostic accuracy, as well as discerning from other tumor types. However, more studies are still needed with further considerations, such as adequate operator training, expertise, availability, and its cost-effectiveness in comparison to other imaging options.

Key Words: Pancreatic malignancy; Pancreatic lesion; Endoscopic ultrasound; Fine needle aspiration biopsy; Elastography

Indonesia

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: May 25, 2021
Peer-review started: May 25, 2021
First decision: June 18, 2021
Revised: June 20, 2021
Accepted: July 6, 2021
Article in press: July 6, 2021
Published online: August 28, 2021

P-Reviewer: Altonbary AY
S-Editor: Liu M
L-Editor: A
P-Editor: Wang LYT

Core Tip: The application of endoscopic ultrasound (EUS) elastography is one of the most potential roles of artificial intelligence in pancreaticobiliary disorders. EUS elastography becomes a promising method to evaluate pancreatic lesions by providing information of tissue elasticity, which may correlate with malignant characteristics. Incomplete elastographic delineation, especially in large tumor size, as well as compelling intra-/inter-observer variability also still become limitations in performing adequate EUS elastography examination on pancreatic lesions.

Citation: Lesmana CRA, Paramitha MS. Impact of endoscopic ultrasound elastography in pancreatic lesion evaluation. *Artif Intell Gastrointest Endosc* 2021; 2(4): 168-178
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/168.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.168

INTRODUCTION

Pancreatic malignancy is still considered as the most lethal cancer in the field of gastroenterology. Based on Global Cancer Observatory database 2020, it is still holding the 12th rank of the most common malignancies all over the world. The mortality rate related to pancreatic cancer has increased more than double within 27 years. The survival rate has also been considered far from satisfaction with regards to the standard treatment development. In Asian population, the incidence and mortality related to pancreatic cancer are also quite high (47.1% and 48.1%, respectively)[1]. Most of the patients are diagnosed at the late stage due to organ's location, non-specific clinical manifestation in early stages, and the absence of simple screening test with high accuracy for early stages of the disease.

In the evaluation of pancreatic cancer, imaging has been proven to play a central and critical role. Imaging modalities are expected to be able to detect and characterize the tumor mass, evaluate local and vascular involvement, evaluate lymphatic and perineural invasion, and find any metastases. Evolution of diagnostic imaging examination such as abdominal computed tomography (CT) scan and magnetic resonance imaging (MRI) have shown good accuracy for detecting pancreatic lesion. A single-center retrospective study in 140 subjects showed higher sensitivity (89.5% *vs* 81.4%) and specificity (63.4% *vs* 43%) in MRI compared to CT-scan for evaluating pancreatic adenocarcinoma. This study also showed that only 14% of the patients were diagnosed in the early stage at the time of diagnosis. Nevertheless, in the setting of small size of tumor mass, uncooperative patients for MRI evaluation, availability of MRI, lack of clinicians' familiarity with the device, and high cost of performing MRI still become the limitations in clinical practice. Additionally, from the same study, the highest diagnostic accuracy was shown by endoscopic ultrasound (EUS) (sensitivity 97.5%, specificity 90.3%). In the new era of the old instrument development, EUS has become the cornerstone in pancreatic malignancy, as it has a high sensitivity for small tumor size (< 2 cm), evaluation of staging (including the presence of lymph nodes, ascites, liver metastasis, and vascular involvement), and to perform direct tissue sampling[2,3]. However, in the conditions of uncertain malignant condition, normal tumor markers level, and possibility of needle tract seeding, a dilemmatic condition on whether the lesion should be punctured or not may arise[3-5]. Learning from the non-invasive tool development, such as elasticity evaluation, has opened a better insight for utilizing EUS, not only for diagnostic purpose, but also for therapeutic purpose.

PRINCIPLE OF ENDOSCOPIC ULTRASOUND ELASTOGRAPHY

The concept of utilizing combination of elastography (EG) and ultrasonography in diagnosing pancreatic disorders has been proposed as a way to overcome the diagnostic problem of solid pancreatic lesions (Figure 1). A prospective study conducted by Uchida *et al*[6] showed that real-time tissue EG and transcutaneous ultrasonography can provide real-time visualization and information of pancreatic tissue elasticity. Combination of sonic and ultrasound waves will cause less compression in fibrotic and stiff tissue, in comparison to softer and healthy tissue. This characteristic may overcome the limitation of conventional EUS, especially in patients

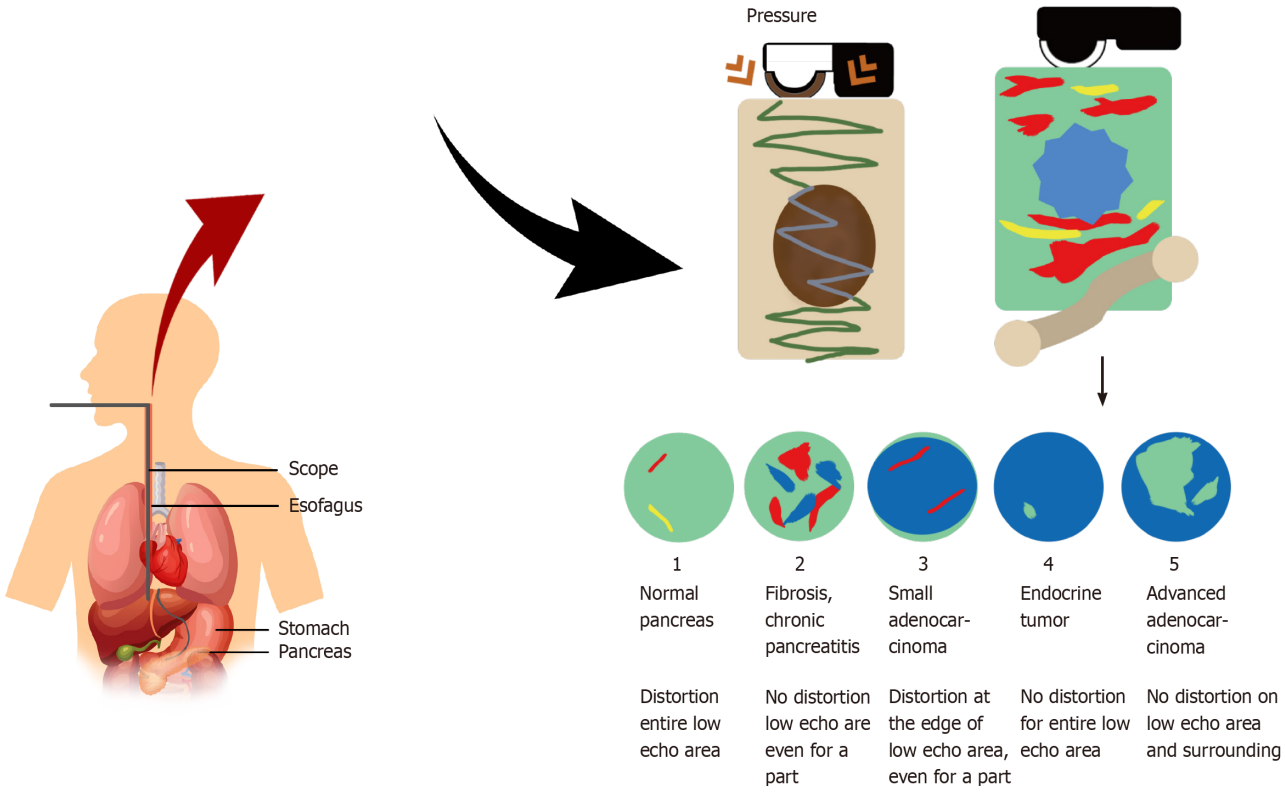

Figure 1 Basic principles of endoscopic ultrasound elastography in pancreatic lesion evaluation. The possibility of combining endoscopy and ultrasonography in evaluating pancreatic lesion through the principle of strain elastography, in which, tissues with higher elasticity will exhibit more deformation after a form of pressure is being applied. The degree of displacement will then be represented as colour pattern analysis to determine the possible diagnosis (Red = Soft tissue; Green = Intermediate tissue; Blue = Hard tissue).

with coexistent chronic pancreatitis or "pseudotumoral" pancreatitis[7]. As one of the most recent approaches in gastrointestinal endoscopy, EUS real-time tissue EG has more diagnostic potentials compared to EUS with only a B-mode imaging ability. In general, EUS EG provides information of tissue elasticity through differences in deformation and displacement among tissue areas, as well as different amount of tissue distortion attained from spatial differentiation. Tissue consistency may correlate with malignancy characteristics, in which malignant tissues have harder consistency than benign tissues[8].

Reported for the first time in 2006 for evaluating pancreatic tissues, EUS EG has been continuously developed for tissue elasticity assessment. Two methods have been differently proposed and compared for each diagnostic performance, *i.e.*, strain and shear-wave EG. Generally, strain elastograms are produced by internal physiological pulsations from respiratory contractions. Estimation of the target tissue's stiffness is conducted with semiquantitative real-time elastography (RTE) using strain histogram (SH), and quantitative strain ration (SR) histogram EG. In particular, SR is a semi-quantitative method to calculate relative tissue stiffness by dividing mean strain of reference area and mean strain in lesion of interest. Meanwhile, the global hardness of a lesion is expressed by the mean histogram value (numerical values from SH)[3,9]. There are three major important principles when RTE is applied for tissue elasticity evaluation, *i.e.*, the stress compression, the region of interest (ROI), and the tissue displacement. Semi-quantitative SH EUS EG uses the manual method through tissue compression effect or pressure application, which will create color-based results. Quantitative strain elastograms or SH needs to calculate the ratio; however, this can be a combined assessment. This software methods usually will be incorporated to the echoendoscope for pancreatic tissue assessment[3,8]. In a healthy pancreatic tissue, the internal structure is isoechoic with soft elastogram. In elderly, the consistency of pancreatic tissue is remarkably harder, but not as hardened as the histogram result of chronic pancreatitis. In acute pancreatitis, softer consistency can be observed in the necrotic zones. Significantly higher stiffness (often unequivocal) can be found in ductal adenocarcinoma. The hue color-based parameter, where it is used for tissue elasticity evaluation, consists of red, green, and blue color. Soft tissue appears as red color, whereas intermediate tissue appears as green color, and blue color will represent hard

tissue. However, perception errors or variability of interpretation between endosonographers may occur in the characterization of hue color-based parameter[8,9].

On the other hand, shear-wave EG is a quantitative tissue elasticity assessment, where it has been mostly used for liver, breast, prostate, rectum, and lymph node. In shear wave EG, focused ultrasound from the probe to target tissue is emitted and evaluation of target tissue's stiffness is performed afterwards by measuring the shear wave's propagation speed. An exploratory study of EUS shear-wave measurement (EUS-SWM) in the assessment and treatment of autoimmune pancreatitis showed approximately 97.6% success rate with no significant difference of success rate in the head, body, and tail of the pancreas (P = 0.4997)[10]. Another preliminary study also demonstrated similarly high success rate (96.8%) without any adverse events. In addition, the elastic value with unique reliability index of the velocity of shear wave measurement also allows more objective and repeated measurement with EUS-SWM [11]. However, compared to strain EG, varying results with EUS-SWM are still found from previous study by Carlsen *et al*[12] This study also showed that target diameter had the most significant effect for all methods of shear-wave EG measurement, while target depth only affected shear-wave velocity measurement in targets with hard consistency.

ENDOSCOPIC ULTRASOUND EG IN PANCREATIC LESION EVALUATION

Throughout the years, evidences related to the use of EUS EG in pancreatic lesion evaluation keep emerging (Table 1). A pioneer study by Giovannini *et al*[13], 2006 showed the impact of endosonoelastography examination for pancreatic masses evaluation in 49 patients, where the sensitivity and specificity in diagnosing malignant lesion were 100% and 67%, respectively. In this study, there were two misdiagnosed cases (neuroendocrine tumor and benign fibromyoblastic tumor of surgically resected pancreas). The sensitivity and specificity of endosonoelastography in assessing malignant lymph node invasion in this study were 100% and 50%, respectively. As mentioned in the previous section, the first experimental study for real-time tissue EG for pancreatic tissue assessment was investigated by Uchida *et al*[6], 2009, in which a linear probe, with B-mode and EG mode, was used to visualize the object. The color-based (blue for hard and red for soft) was used in the ROI. In pancreatic cancer, the lesion was identified with blue color, which was subsequently confirmed through histopathologic examination result. Combination of B-mode and EG mode increased the diagnosis accuracy of pancreatic cancer from 73.3% to 100%, corrected by operator. The sensitivity and specificity between operator and another reviewer showed the same results for EG mode evaluation (64.3% *vs* 60.7% and 88% *vs* 88%). In the case of pancreatic endocrine tumor, the diagnosis accuracy also increased from 66.7% to 100% [6]. In 2009, a prospective study by Iglesias-Garcia *et al*[14], where the EG pattern was compared to histological specimen, showed the blue color pattern supported the malignant pancreatic lesions, whereas the green color pattern excluded malignant lesions. The sensitivity and specificity of EG diagnosis in malignant pancreatic lesions were 100% and 85.5%, respectively. This study concluded that the overall diagnostic accuracy of EUS EG for malignancy was 94%. Further concordance analysis by two endosonographers yielded agreement of elastographic pattern by both of them in 93.1% of the cases. This study also addressed the possibility of EUS EG in tackling the limitation of EUS-guided fine needle aspiration (EUS-FNA). One of the major drawbacks of EUS-FNA was interposition of malignant tissue and vascular structures, which may contribute to false negative results. EUS EG can overcome this limitation by assessing tissue elasticity and discerning hardness between normal and malignant tissues[14].

In contrast to previous evidences, a prospective study by Hirche *et al*[4] showed that EUS-EG had low sensitivity (41%), specificity (53%), and accuracy (45%) in predicting malignant pancreatic lesion. A subgroup analysis in ductal adenocarcinoma also demonstrated poor sensitivity (50%). Moderate intraobserver and interobserver reproducibilities were also demonstrated from the findings. However, in this study, the sample size was considered small. Additionally, some patients were diagnosed with cystic lesion tumor, suggesting that presence of fluid might interfere the elastographic pattern. On the other hand, larger tumor size was causing the inaccurate distance between the EUS probe and the mucosal wall. Incomplete border delineation by EUS- EG was also shown in lesions with a larger diameter, leading to insufficient display of surrounding pancreatic parenchyma[4]. In another small prospective single-center study by Janssen *et al*[15], three groups were classified as normal pancreas,

Table 1 Summary of the studies utilizing endoscopic ultrasound elastography for evaluating pancreatic lesions

Ref.	Population of the study	Key findings
Giovannini *et al* [13], 2006	24 patients with pancreatic masses.	Sensitivity 100% and specificity 67% in diagnosing malignant lesions. Sensitivity 100% and specificity 50% in diagnosing malignant invasion of lymph nodes.
Uchida *et al*[6], 2009	Phase 1: pancreatic cancer (5 subjects), endocrine tumor (2 subjects), chronic pancreatitis (5 subjects), intraductal papillary mucinous neoplasm. Phase 2: 53 consecutive subjects with pancreatic lesions visible on B-mode images.	Diagnostic performance of real-time tissue elastography mode for diagnosing malignancy: Operator *vs* another reviewer Sensitivity: 64.3% *vs* 60.7%. Specificity: 88% *vs* 88%. Positive predictive value: 85.7% *vs* 85%. Negative predictive value: 68.8% *vs* 66.7%.
Iglesias-Garcia *et al*[14], 2009	130 consecutive patients with solid pancreatic masses *vs* 20 subjects with normal pancreases.	Diagnostic performance of EUS elastography in diagnosing malignancy Sensitivity: 100%. Specificity: 85.5%. Positive predictive value: 90.7%. Negative predictive value: 100%. Overall accuracy: 94%.
Hirche *et al*[4], 2008	70 patients with unclassified solid pancreatic lesions *vs* 10 subjects with healthy pancreas.	Diagnostic performance of EUS elastography in predicting the nature of pancreatic lesions Sensitivity: 41%. Specificity: 53%. Accuracy: 45%.
Janssen *et al*[15], 2007	20 patients with chronic pancreatitis *vs* 33 patients with focal pancreatic lesions *vs* 20 subjects with normal pancreas.	Diagnostic performance of EUS elastography in diagnosing chronic pancreatitis Sensitivity: 65.9%. Specificity: 56.9%. Accuracy: 60.2%. Diagnostic performance of EUS elastography in diagnosing focal pancreatic lesions Sensitivity: 93.8%. Specificity: 65.4%. Accuracy: 73.5%. Diagnostic performance of EUS elastography in differentiating pancreatic adenocarcinoma and inflammatory pancreatic masses
Li *et al*[16], 2013	Meta-analysis of 10 studies with 781 patients.	Diagnostic performance of EUS elastography in differentiating pancreatic adenocarcinoma and inflammatory pancreatic masses AUC: 0.8227. In studies with color pattern as the diagnostic standard Sensitivity: 99%. Specificity: 76%. Positive likelihood ratio: 3.36. Negative likelihood ratio: 0.03. Diagnostic odds ratio: 129.96. In studies with hue histogram as the diagnostic standard

		Sensitivity: 92%. Specificity: 68%. Positive likelihood ratio: 2.84. Negative likelihood ratio: 0.12. Diagnostic odds ratio: 24.69.
Xu *et al*[17], 2013	Meta-analysis of 9 studies.	Diagnostic performance of EUS elastography in differentiating benign and malignant pancreatic masses In studies with qualitative color pattern as the diagnostic standard Sensitivity: 99%. Specificity: 74%. AUROC: 0.9624. In studies with quantitative hue histogram value as the diagnostic standard Sensitivity: 85%-93%. Specificity: 64%-76%.
Mei *et al*[18], 2013	Meta-analysis of 13 studies with 1044 patients.	Diagnostic performance of EUS elastography in differentiating benign and malignant solid pancreatic masses Sensitivity: 95%. Specificity: 67%. Diagnostic odds ratio: 42.28.
Altonbary *et al*[19], 2019	97 patients with malignant lesions *vs* 19 patients with benign lesions	Diagnostic performance of combined elasticity score and strain ratio in differentiating benign and malignant pancreatic lesions (cut-off point: 7.75) Sensitivity: 99%. Specificity: 94.6%. Positive predictive value: 98%. Negative predictive value: 98.5%. Accuracy: 97%.
Ignee *et al*[20], 2018	218 patients with solid pancreatic lesions sized ≤ 15 mm and a definite diagnosis.	Diagnostic performance of EUS elastography with high stiffness of the lesion in diagnosing malignancy Sensitivity: 84%. Specificity: 67%. Positive predictive value: 56%. Negative predictive value: 89%. Diagnostic performance of EUS elastography in diagnosing pancreatic ductal adenocarcinoma Sensitivity: 96%. Specificity: 64%. Positive predictive value: 45%. Negative predictive value: 98%.

EUS: Endoscopic ultrasound.

chronic pancreatitis, and focal pancreatic lesions. The elastographic pattern classification (homogenous, different colors, and honeycomb pattern) and elastographic colors classification (blue, green/yellow, and red) were combinedly used. In normal pancreas group, all showed homogenous green color interfered with blue clouds' color. Whereas, in chronic pancreatitis group showed hard (blue) with honeycomb pattern. In pancreatic focal lesions' group, examination showed that almost all patients had blue/green honeycomb pattern. Only one patient which has tumorlike due to

chronic pancreatitis showed blue/green honeycomb pattern. The sensitivity and specificity for group with chronic pancreatitis were 65.9% and 56.9%, respectively, with diagnostic accuracy of 60.2%; while the sensitivity and specificity in group with focal pancreatic lesions were 93.8% and 65.4%, respectively, with slightly higher diagnostic accuracy (73.5%). The findings from this study also addressed the limitation of EUS EG in distinguishing the elastographic patterns of chronic pancreatitis and malignant tumors due to the corresponding amount of fibrous pattern of chronic pancreatitis, which can also be found in desmoplastic pancreatic carcinomas and microcystic adenomas[15]. Another meta-analysis, which evaluated the use of EUS EG in discernment of pancreatic adenocarcinoma and inflammatory masses, indicated slightly better diagnostic performance in studies with color pattern as the diagnostic standard (sensitivity 99%, specificity 76%) compared to studies with hue histogram as the diagnostic standard (sensitivity 92%, specificity 68%)[16]. In differentiating benign and malignant pancreatic masses, better diagnostic performance was also demonstrated by studies using qualitative color pattern as the diagnostic standard (sensitivity 99%, specificity 74%) in comparison to studies using hue histogram as the diagnostic standard (sensitivity 85%-93%, specificity 64%-76%). This meta-analysis also acknowledged the difficulties in distinguishing neuroendocrine tumors and adenocarcinomas due to their similar hardness[17]. Regardless of the low specificity, EUS EG can still be considered as a complementary diagnostic method. A meta-analysis by Mei *et al*[18] showed high pooled sensitivity (95%) with acceptable pooled specificity (67%) and moderate accuracy (summary Receiver Operating Characteristic: 90.46%) of EUS EG in diagnosing solid pancreatic masses. Improvement of diagnostic accuracy may be achieved with application of more meticulous computer-aided diagnosis method for EUS-EG[18]. Recent findings from a single center retrospective study by Altonbary *et al*[19] also demonstrated promising results of EUS EG with combination of elasticity score and strain ratio in discerning solid pancreatic lesions (sensitivity 99%, specificity 94.6%, and accuracy 97%). Moderately well diagnostic performance in ruling out malignancy was also demonstrated by a multicenter study conducted in 218 patients with small (< 15 mm) solid pancreatic lesions (sensitivity 84%). Higher sensitivity (96%) was shown when EUS EG was used in diagnosing Pancreatic Ductal Adenocarcinoma (PDAC)[20].

CLINICAL DILEMMA IN PANCREATIC LESION EVALUATION AND IMPACT OF EUS EG INNOVATION STUDY

Several conditions have been considered as clinical dilemma, such as small pancreatic lesion which also can be found incidentally, pseudo-tumoral in chronic pancreatitis, negative FNA biopsy (FNAB) results, and possibility of needle tract tumor seeding[3-5]. It has been known that pancreatic cancer is mostly dominated by PDAC, a highly aggressive tumor with very poor prognosis and high mortality rate. It has been reported that Negative Predictive Value (NPV) of FNAB result can vary, ranging from 16% to 85%. In the case of negative biopsy, patients with suspicion of PDAC should be referred immediately for surgical approach consideration. Spier *et al*[21] published a small retrospective EUS-FNA study in patients who had suspected pancreatic lesions with negative biopsy results. The study found that 30.9% of patients with negative/non-diagnostic FNA results were later diagnosed with pancreatic cancer (mean time 66 d to 360 d after FNA procedure)[21]. RTE has been proposed as a supplementary method to improve diagnostic performance of EUS-FNA, especially in terms of available rapid on-site tissue evaluation by a cytopathologist[22,23]. A retrospective study in 54 subjects with solid pancreatic lesions highlighted the benefit of combining RTE and EUS-FNA (sensitivity 94.4%, specificity 93.4%, and accuracy 100%) compared to the diagnostic performance of RTE alone (sensitivity 86.9%, specificity 75%, and accuracy 85.1%)[22].

Possibility of tumor seeding has become a challenging issue as it will impact on faster disease progression, patient's clinical-based management, and patient's survival after surgery or non-surgical biliary drainage procedure in patients with bile duct obstruction. There has been a debate on whether this tract seeding issue should be underestimated or overestimated, since most of the studies use retrospective study design. Small sample size and no clear tumor dissemination finding also become issues on the studies of needle tract seeding related to EUS-FNA[5]. The first reported case of EUS-FNA-related tumor dissemination was delivered in 2003, in which peritoneal dissemination occurred in intraductal papillary mucinous tumor (T1N0M0)[5]. Approximately 80% of all needle tract seeding cases following EUS-FNA happened

in pancreatic cancer and pancreatic cystic tumors located in the body or tail of pancreas. In most of the cases, 22-G FNA needle was used, even though the relationship between needle size or number of needle passes and the risk of tumor seeding is still unclear. The range of interval from EUS-FNA procedure to diagnosis of needle tract seeding is 3-48 mo[24].

EG EUS multicenter study by Ignee *et al*[20] in small solid pancreatic lesions showed that sensitivity and specificity were 84% and 67%, respectively, with 56% of positive predictive value and 89% of NPV. In PDAC cases, sensitivity and specificity were 96% and 64%, respectively. Based on this study, it is clear that early detection in less than 15 mm pancreatic lesion might prevent the delay for surgery management even though PDAC tends to be found in larger lesions (> 15 mm)[20]. Another prospective study was conducted by Dawwas *et al*[25] in patients underwent quantitative EUS EG procedure for differentiating pancreatic malignant lesion with pancreatic inflammatory lesion. The examination results were compared to histology or cytology results with follow-up imaging study. The sensitivity and specificity with quantitative EUS EG were 100% and 95.7%, respectively. This study has shown the important value of EUS EG in reducing the need of biopsy as the EUS-FNAB procedure still carries potentially harmful risks, such as pancreatitis and bleeding[25]. In 2018, Dong *et al*[26] reported the role of combination strategy using B-mode ultrasound, contrast-enhanced ultrasound (CEUS), and EUS EG in small case series of isolated pancreatic tuberculosis (PTB) cases. These findings were then compared with the clinical findings of PDAC cases. In PTB cases, common bile duct and pancreatic duct dilatation are considered to be rare findings, however, it is common to find multiple peripancreatic lymph nodes enlargement. The PTB lesion was showing less demarcation, whereas clear demarcation was found in PDAC cases. It might be difficult to differentiate PTB from PDAC cases by using the tissue stiffness result from EUS elastrography alone, however, with CEUS combination, PTB lesion showed hyperenhancement, whereas in PDAC cases showed hypoenhancement. In addition, peripancreatic pseudocysts were more commonly observed in PTB cases. This non-invasive strategy can be an accurate diagnosis tool with or without biopsy as a clinical-based approach in patients with PTB. Consequently, it can also avoid unnecessary surgical management[26].

A former retrospective analysis study by Iordache *et al*[27] in 50 consecutive patients with negative results of EUS-FNA who sequentially underwent EUS EG and CE-EUS, found that EUS EG has similar results with CE-EUS in diagnosing possibility of pancreatic malignancy. However, combination of both methods showed excellent specificity (100%). Another interesting finding from this study is the excellent specificity (100%) exhibited by CEH-EUS in patients with soft/mixed or hard (low strain) appearance from EG. Excellent specificity was shown by CEH-EUS for distinguishing chronic pancreatitis in soft/mixed (high strain) appearance; while in hard appearance, CEH-EUS exhibited outstanding specificity (100%) and sensitivity (88.89%) for distinguishing pancreatic cancer. These results suggested that hard hypovascular masses can indicate the presence of pancreatic adenocarcinoma or other malignant masses, whereas soft hyper-/isovascular masses can indicate the presence of chronic pseudotumoral pancreatitis or other benign masses[27]. Another prospective multi-center study by Costache *et al*[28] about clinical impact of combination between SH EUS EG and CE-EUS in patients with pancreatic masses, showed that combined CE-EUS with SH EUS EG had similar sensitivity. However, higher specificity (81.48%) was found in the combination method for diagnosing pancreatic carcinoma in comparison to SH EUS EG with several cut-offs (80; 60; 40; 33). Meanwhile, the specificity of single method was ranging from 29.63% to 62.96% based on several cut-offs. The overall diagnostic accuracy in combination method reached 93.81% for pancreatic cancer, whereas in the single method only ranged from 79.38 % to 80.41%. Overall, this study indicated that combination of CE-EUS and semi-quantitative EUS EG can be utilized as a supplementary modality for distinguishing benign and malignant pancreatic masses and for continuous follow-up evaluation of patients during neo-adjuvant chemotherapy and/or anti-angiogenic therapy administration[28]. A case series study by Jafri *et al*[29] showed the potential of EUS EG as a complementary method along with conventional EUS for targeting the FNA procedure in patients with suspected pancreatic masses. Also, in this case series, subjects with low risk of malignancy from EUS and EG examinations did not develop any interval cancer during the mean period of 2-year follow-up[29].

CURRENT STATUS AND LIMITATIONS

According to most studies on EUS EG, it has been shown that EUS EG has a big role in managing pancreatic lesions. This method can be a primary choice for diagnosis evaluation in patients who have coagulation disorders or history of anticoagulation drugs consumption, who are not suitable yet for chemotherapy, and who have the possibility for direct surgical approach due to the needle tract seeding risk during FNA procedures. In targeting unclear demarcation and pancreatic lesion image, EUS EG can also be an additional tool. However, it cannot be used for pancreatic cystic mass tumor evaluation. Studies to differentiate between malignant and benign pancreatic mass lesion have not shown any strong evidence yet as some studies were only performed with small sample size, and some only used retrospective study analysis.

The main objectives of performing EG for the pancreas are to ensure that the elastogram is sufficiently meticulous to represent the histological structures and to be reproducible adequately. These objectives, however, are hampered by the small size of the pancreas, the depth of its anatomical location in the center of the body, the technical difficulties in extracting biopsy specimens, and the strong influence of aortic pulsation to pancreas. In addition, EG is an operator-independent modality[30]. Other pitfalls of EUS EG are the difficulty in controlling tissue compression by the EUS transducer, the presence of motion artifacts due to respiratory movement, as well as the careful selection of ROI from its surrounding soft tissues[31].

Overall, the application of EUS EG is one of the most potential roles of artificial intelligence (AI) in pancreaticobiliary disorders. In general, AI refers to the capacity of a computer to imitate the cognitive intelligence or the learning capability of human being in order to perform tasks appropriately. In medicine, AI consists of machine learning and deep learning, which are often utilized reciprocally[32]. A cross-sectional feasibility study in Denmark established the importance of AI in distinguishing pancreatic cancer from chronic pancreatitis through the application of neural network analysis of dynamic sequences of EUS EG. In this study, the sensitivity, specificity, and accuracy were 91.4%, 87.9%, and 89.7%, respectively. In addition, the application of multilayer perceptron neural networks with high training performance was able to reach an accuracy as high as 97%[33]. Another prospective and multicenter study in 258 patients by Săftoiu *et al*[34] also highlighted the efficacy of AI in EUS EG. The utilization of multilayer perceptron as an artificial neural network demonstrated moderately high diagnostic performance (sensitivity 87.59%, specificity 82.94%, AUROC 0.94, training accuracy 91.14%, and testing accuracy 84.27%) in diagnosing focal pancreatic lesions.

CONCLUSION

EUS EG is a promising method to improve the diagnostic accuracy as well as helping to decide which type of management is probably more suitable for patients with pancreatic mass lesion. However, it would still need more studies with further considerations, such as adequate operator training, expertise, availability, and its cost-effectiveness in comparison to other imaging options. Integrating clinical data into artificial intelligence techniques concomitantly with real-time imaging results is potentially favorable for faster and more accurate clinical-decision making in pancreatic lesion evaluation.

ACKNOWLEDGEMENTS

I would like to thank to Professor Ho Khek Yu, Past President of Asian EUS Group (AEG), and Professor Laurentius A Lesmana, Chair of Digestive Disease & GI Oncology Center, Medistra Hospital, Jakarta who had given big support and contributions in the development of endoscopic ultrasound (EUS) in Indonesia.

REFERENCES

1 **World Health Organization International Agency for Research on Cancer (IARC)**. GLOBOCAN 2020: Pancreas. [homepage on the internet]; 2020 [cited 2021 May 22]. Available from:

https://gco.iarc.fr/today/data/factsheets/cancers/13-Pancreas-fact-sheet.pdf

2 **Costache MI**, Costache CA, Dumitrescu CI, Tica AA, Popescu M, Baluta EA, Anghel AC, Saftoiu A, Dumitrescu D. Which is the Best Imaging Method in Pancreatic Adenocarcinoma Diagnosis and Staging - CT, MRI or EUS? *Curr Health Sci J* 2017; **43**: 132-136 [PMID: 30595868 DOI: 10.12865/CHSJ.43.02.05]

3 **Kitano M**, Yoshida T, Itonaga M, Tamura T, Hatamaru K, Yamashita Y. Impact of endoscopic ultrasonography on diagnosis of pancreatic cancer. *J Gastroenterol* 2019; **54**: 19-32 [PMID: 30406288 DOI: 10.1007/s00535-018-1519-2]

4 **Hirche TO**, Ignee A, Barreiros AP, Schreiber-Dietrich D, Jungblut S, Ott M, Hirche H, Dietrich CF. Indications and limitations of endoscopic ultrasound elastography for evaluation of focal pancreatic lesions. *Endoscopy* 2008; **40**: 910-917 [PMID: 19009483 DOI: 10.1055/s-2008-1077726]

5 **Minaga K**, Takenaka M, Katanuma A, Kitano M, Yamashita Y, Kamata K, Yamao K, Watanabe T, Maguchi H, Kudo M. Needle Tract Seeding: An Overlooked Rare Complication of Endoscopic Ultrasound-Guided Fine-Needle Aspiration. *Oncology* 2017; **93** Suppl 1: 107-112 [PMID: 29258068 DOI: 10.1159/000481235]

6 **Uchida H**, Hirooka Y, Itoh A, Kawashima H, Hara K, Nonogaki K, Kasugai T, Ohno E, Ohmiya N, Niwa Y, Katano Y, Ishigami M, Goto H. Feasibility of tissue elastography using transcutaneous ultrasonography for the diagnosis of pancreatic diseases. *Pancreas* 2009; **38**: 17-22 [PMID: 18695627 DOI: 10.1097/MPA.0b013e318184db78]

7 **Gill KR**, Wallace MB. EUS elastography for pancreatic mass lesions: between image and FNA? *Gastrointest Endosc* 2008; **68**: 1095-1097 [PMID: 19028217 DOI: 10.1016/j.gie.2008.05.001]

8 **Dietrich CF**, Bibby E, Jenssen C, Saftoiu A, Iglesias-Garcia J, Havre RF. EUS elastography: How to do it? *Endosc Ultrasound* 2018; **7**: 20-28 [PMID: 29451165 DOI: 10.4103/eus.eus_49_17]

9 **Costache MI**, Dumitrescu D, Săftoiu A. Technique of qualitative and semiquantitative EUS elastography in pancreatic examination. *Endosc Ultrasound* 2017; **6**: S111-S114 [PMID: 29387705 DOI: 10.4103/eus.eus_75_17]

10 **Ohno E**, Hirooka Y, Kawashima H, Ishikawa T, Tanaka H, Sakai D, Ishizu Y, Kuzuya T, Nakamura M, Honda T. Feasibility and usefulness of endoscopic ultrasonography-guided shear-wave measurement for assessment of autoimmune pancreatitis activity: a prospective exploratory study. *J Med Ultrason (2001)* 2019; **46**: 425-433 [PMID: 30993580 DOI: 10.1007/s10396-019-00944-4]

11 **Ohno E**, Hirooka Y, Kawashima H, Ishikawa T. Feasibility of EUS-guided shear-wave measurement: A preliminary clinical study. *Endosc Ultrasound* 2019; **8**: 215-216 [PMID: 30924448 DOI: 10.4103/eus.eus_6_19]

12 **Carlsen JF**, Pedersen MR, Ewertsen C, Săftoiu A, Lönn L, Rafaelsen SR, Nielsen MB. A comparative study of strain and shear-wave elastography in an elasticity phantom. *AJR Am J Roentgenol* 2015; **204**: W236-W242 [PMID: 25714307 DOI: 10.2214/AJR.14.13076]

13 **Giovannini M**, Hookey LC, Bories E, Pesenti C, Monges G, Delpero JR. Endoscopic ultrasound elastography: the first step towards virtual biopsy? *Endoscopy* 2006; **38**: 344-348 [PMID: 16680632 DOI: 10.1055/s-2006-925158]

14 **Iglesias-Garcia J**, Larino-Noia J, Abdulkader I, Forteza J, Dominguez-Munoz JE. EUS elastography for the characterization of solid pancreatic masses. *Gastrointest Endosc* 2009; **70**: 1101-1108 [PMID: 19647248 DOI: 10.1016/j.gie.2009.05.011]

15 **Janssen J**, Schlörer E, Greiner L. EUS elastography of the pancreas: feasibility and pattern description of the normal pancreas, chronic pancreatitis, and focal pancreatic lesions. *Gastrointest Endosc* 2007; **65**: 971-978 [PMID: 17531630 DOI: 10.1016/j.gie.2006.12.057]

16 **Li X**, Xu W, Shi J, Lin Y, Zeng X. Endoscopic ultrasound elastography for differentiating between pancreatic adenocarcinoma and inflammatory masses: a meta-analysis. *World J Gastroenterol* 2013; **19**: 6284-6291 [PMID: 24115828 DOI: 10.3748/wjg.v19.i37.6284]

17 **Xu W**, Shi J, Li X, Zeng X, Lin Y. Endoscopic ultrasound elastography for differentiation of benign and malignant pancreatic masses: a systemic review and meta-analysis. *Eur J Gastroenterol Hepatol* 2013; **25**: 218-224 [PMID: 23169307 DOI: 10.1097/MEG.0b013e32835a7f7c]

18 **Mei M**, Ni J, Liu D, Jin P, Sun L. EUS elastography for diagnosis of solid pancreatic masses: a meta-analysis. *Gastrointest Endosc* 2013; **77**: 578-589 [PMID: 23199646 DOI: 10.1016/j.gie.2012.09.035]

19 **Altonbary AY**, Hakim H, El-Shamy AM. Diagnostic Efficacy of Endoscopic Ultrasound Elastography in Differentiating Solid Pancreatic Lesions: A Single-Center Experience. *Clin Endosc* 2019; **52**: 360-364 [PMID: 30625265 DOI: 10.5946/ce.2018.160]

20 **Ignee A**, Jenssen C, Arcidiacono PG, Hocke M, Möller K, Saftoiu A, Will U, Fusaroli P, Iglesias-Garcia J, Ponnudurai R, Petrone MC, Braden B, Burmester E, Dong Y, Atkinson NS, Dietrich CF. Endoscopic ultrasound elastography of small solid pancreatic lesions: a multicenter study. *Endoscopy* 2018; **50**: 1071-1079 [PMID: 29689572 DOI: 10.1055/a-0588-4941]

21 **Spier BJ**, Johnson EA, Gopal DV, Frick T, Einstein MM, Byrne S, Koscik RL, Liou JI, Broxmeyer T, Selvaggi SM, Pfau PR. Predictors of malignancy and recommended follow-up in patients with negative endoscopic ultrasound-guided fine-needle aspiration of suspected pancreatic lesions. *Can J Gastroenterol* 2009; **23**: 279-286 [PMID: 19373422 DOI: 10.1155/2009/870323]

22 **Facciorusso A**, Martina M, Buccino RV, Nacchiero MC, Muscatiello N. Diagnostic accuracy of fine-needle aspiration of solid pancreatic lesions guided by endoscopic ultrasound elastography. *Ann Gastroenterol* 2018; **31**: 513-518 [PMID: 29991898 DOI: 10.20524/aog.2018.0271]

23 **Yang F**, Liu E, Sun S. Rapid on-site evaluation (ROSE) with EUS-FNA: The ROSE looks beautiful. *Endosc Ultrasound* 2019; **8**: 283-287 [PMID: 31603143 DOI: 10.4103/eus.eus_65_19]

24 **Hirooka Y**, Goto H, Itoh A, Hashimoto S, Niwa K, Ishikawa H, Okada N, Itoh T, Kawashima H. Case of intraductal papillary mucinous tumor in which endosonography-guided fine-needle aspiration biopsy caused dissemination. *J Gastroenterol Hepatol* 2003; **18**: 1323-1324 [PMID: 14535994 DOI: 10.1046/j.1440-1746.2003.03040.x]

25 **Dawwas MF**, Taha H, Leeds JS, Nayar MK, Oppong KW. Diagnostic accuracy of quantitative EUS elastography for discriminating malignant from benign solid pancreatic masses: a prospective, single-center study. *Gastrointest Endosc* 2012; **76**: 953-961 [PMID: 22854060 DOI: 10.1016/j.gie.2012.05.034]

26 **Dong Y**, Jürgensen C, Puri R, D'Onofrio M, Hocke M, Wang WP, Atkinson N, Sharma M, Dietrich CF. Ultrasound imaging features of isolated pancreatic tuberculosis. *Endosc Ultrasound* 2018; **7**: 119-127 [PMID: 28721972 DOI: 10.4103/2303-9027.210901]

27 **Iordache S**, Costache MI, Popescu CF, Streba CT, Cazacu S, Săftoiu A. Clinical impact of EUS elastography followed by contrast-enhanced EUS in patients with focal pancreatic masses and negative EUS-guided FNA. *Med Ultrason* 2016; **18**: 18-24 [PMID: 26962549 DOI: 10.11152/mu.2013.2066.181.ich]

28 **Costache MI**, Cazacu IM, Dietrich CF, Petrone MC, Arcidiacono PG, Giovannini M, Bories E, Garcia JI, Siyu S, Santo E, Popescu CF, Constantin A, Bhutani MS, Saftoiu A. Clinical impact of strain histogram EUS elastography and contrast-enhanced EUS for the differential diagnosis of focal pancreatic masses: A prospective multicentric study. *Endosc Ultrasound* 2020; **9**: 116-121 [PMID: 32295969 DOI: 10.4103/eus.eus_69_19]

29 **Jafri M**, Sachdev AH, Khanna L, Gress FG. The Role of Real Time Endoscopic Ultrasound Guided Elastography for Targeting EUS-FNA of Suspicious Pancreatic Masses: A Review of the Literature and A Single Center Experience. *JOP* 2016; **17**: 516-524 [PMID: 28912670]

30 **Kawada N**, Tanaka S. Elastography for the pancreas: Current status and future perspective. *World J Gastroenterol* 2016; **22**: 3712-3724 [PMID: 27076756 DOI: 10.3748/wjg.v22.i14.3712]

31 **Lee TH**, Cha SW, Cho YD. EUS elastography: advances in diagnostic EUS of the pancreas. *Korean J Radiol* 2012; **13** Suppl 1: S12-S16 [PMID: 22563282 DOI: 10.3348/kjr.2012.13.S1.S12]

32 **Goyal H**, Mann R, Gandhi Z, Perisetti A, Zhang Z, Sharma N, Saligram S, Inamdar S, Tharian B. Application of artificial intelligence in pancreaticobiliary diseases. *Ther Adv Gastrointest Endosc* 2021; **14**: 2631774521993059 [PMID: 33644756 DOI: 10.1177/2631774521993059]

33 **Săftoiu A**, Vilmann P, Gorunescu F, Gheonea DI, Gorunescu M, Ciurea T, Popescu GL, Iordache A, Hassan H, Iordache S. Neural network analysis of dynamic sequences of EUS elastography used for the differential diagnosis of chronic pancreatitis and pancreatic cancer. *Gastrointest Endosc* 2008; **68**: 1086-1094 [PMID: 18656186 DOI: 10.1016/j.gie.2008.04.031]

34 **Săftoiu A**, Vilmann P, Gorunescu F, Janssen J, Hocke M, Larsen M, Iglesias-Garcia J, Arcidiacono P, Will U, Giovannini M, Dietrich CF, Havre R, Gheorghe C, McKay C, Gheonea DI, Ciurea T; European EUS Elastography Multicentric Study Group. Efficacy of an artificial neural network-based approach to endoscopic ultrasound elastography in diagnosis of focal pancreatic masses. *Clin Gastroenterol Hepatol* 2012; **10**: 84-90.e1 [PMID: 21963957 DOI: 10.1016/j.cgh.2011.09.014]

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 179-184

DOI: 10.37126/aige.v2.i4.179

ISSN 2689-7164 (online)

MINIREVIEWS

Artificial intelligence as a means to improve recognition of gastrointestinal angiodysplasia in video capsule endoscopy

Gerald A Cox II, Christian S Jackson, Kenneth J Vega

ORCID number: Gerald A Cox II 0000-0001-9116-5764; Christian S Jackson 0000-0003-4229-4206; Kenneth J Vega 0000-0002-2432-6123.

Author contributions: Cox II GA, Jackson CS, and Vega KJ planned the minireview and reviewed all data included, wrote the minireview and revised it for intellectual content; Vega KJ is the editorial guarantor; All authors approved the final submitted version.

Conflict-of-interest statement: No conflict of interest exists for all authors of this manuscript.

Manuscript source: Invited manuscript

Gerald A Cox II, Department of Medicine, Loma Linda University Medical Center, Loma Linda, CA 92354, United States

Christian S Jackson, Gastroenterology Section, VA Loma Linda Healthcare System, Loma Linda, CA 92357, United States

Kenneth J Vega, Division of Gastroenterology and Hepatology, Augusta University - Medical College of Georgia, Augusta, GA 30912, United States

Corresponding author: Kenneth J Vega, MD, Director, Professor, Division of Gastroenterology and Hepatology, Augusta University - Medical College of Georgia, 1120 15th Street, AD 2226, Augusta, GA 30912, United States. kvega@augusta.edu

Abstract

Gastrointestinal angiodysplasia (GIAD) is defined as the pathological process where blood vessels, typically venules and capillaries, become engorged, tortuous and thin walled – which then form arteriovenous connections within the mucosal and submucosal layers of the gastrointestinal tract. GIADs are a significant cause of gastrointestinal bleeding and are the main cause for suspected small bowel bleeding. To make the diagnosis, gastroenterologists rely on the use of video capsule endoscopy (VCE) to "target" GIAD. However, the use of VCE can be cumbersome secondary to reader fatigue, suboptimal preparation, and difficulty in distinguishing images. The human eye is imperfect. The same capsule study read by two different readers are noted to have miss rates like other forms of endoscopy. Artificial intelligence (AI) has been a means to bridge the gap between human imperfection and recognition of GIAD. The use of AI in VCE have shown that detection has improved, however the other burdens and limitations still need to be addressed. The use of AI for the diagnosis of GIAD shows promise and the changes needed to enhance the current practice of VCE are near.

Key Words: Artificial intelligence; Video capsule endoscopy; Gastrointestinal angiodysplasia; Detection; Bleeding; Small bowel

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: United States

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: June 3, 2021
Peer-review started: June 3, 2021
First decision: June 23, 2021
Revised: July 7, 2021
Accepted: August 13, 2021
Article in press: August 13, 2021
Published online: August 28, 2021

P-Reviewer: Balaban DV
S-Editor: Liu M
L-Editor: A
P-Editor: Ma YJ

Core Tip: Video capsule endoscopy (VCE) is the primary modality to diagnose gastrointestinal angiodysplasias (GIADs). Typically, gastroenterologists rely on VCE to make a diagnosis of GIAD prior to referral for deep enteroscopy. However, VCE analysis can be cumbersome secondary to reader fatigue, suboptimal preparation, and difficulty in distinguishing images. Use of artificial intelligence in VCE has shown improved GIAD detection, however limitations exist that still need to be addressed. The use of artificial intelligence for GIAD diagnosis shows promise and changes needed to enhance current VCE practices are near.

Citation: Cox II GA, Jackson CS, Vega KJ. Artificial intelligence as a means to improve recognition of gastrointestinal angiodysplasia in video capsule endoscopy. *Artif Intell Gastrointest Endosc* 2021; 2(4): 179-184
URL: https://www.wjgnet.com/2689-7164/full/v2/i4/179.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.179

INTRODUCTION

Gastrointestinal angiodysplasia (GIAD) is defined as the pathological process where blood vessels, typically venules and capillaries, become engorged, tortuous and thin walled – which then form arteriovenous connections within the mucosal and submucosal layers of the gastrointestinal (GI) tract[1]. GIADs are found throughout the GI tract, but they most often occur in the small intestine (80% jejunum, 57% duodenum), stomach (22.8%) and less frequently the ascending colon (11.4%)[2]. The gold standard in diagnosis of GIAD has been endoscopy, with the addition of video capsule endoscopy (VCE) in 2001. The technology of VCE radically improved the diagnostic yield of GIADs as well as other small bowel diseases. VCE provided a means to target lesions in the small bowel and has played a role in the development of balloon enteroscopy for advanced diagnoses and treatment options. Although, VCE improved the diagnostic yield of GIADs, as well other as small bowel diseases, there are several challenges which a reader continues to face. First, review of these images has been an arduous process, which can last from 30-40 min to over an hour. The abnormalities that are of interest may only present in a couple of frames that last a minute or less. Second, the long reading time may lead to reader fatigue and a reduction in diagnostic accuracy. To address these issues, there have been several advances made to VCE technology such as a Quick-view algorithm, suspected blood indicator and adaptive frame rate technology. None of these technologic advances have improved diagnostic accuracy[3-5]. Despite these limitations, VCE is still the widely used technology to diagnose GIAD and has become a growing focus for the use of artificial intelligence (AI) to improve the identification of GIAD. We discuss the implementation of computer software known as AI, machine programs capable of learning and simulating patterns like the human brain.

TYPES OF AI

Several layers exist within AI and have been utilized throughout the field of gastroenterology, especially endoscopy. One aspect is machine learning (ML), a discipline where large, complex data sets are used to predict outcomes and identify patterns using various algorithms[6]. These algorithms are often trained to differentiate data sets or characteristics such as color, size and shape, which help to distinguish between lesions within the GI tract. Beyond ML, two other types of AI exist, artificial neural networks (ANNs) and convolutional neural networks (CNNs). ANNs utilize the patterns observed within data sets to perform complex task of cross comparison at various points of calculation. Therefore, numerous computed data sets can be collected at any stage and compared to provide one outcome. This simulates the intelligence and neurobiological processes of the human brain, as the computer continues to learn to perform new task through automated analysis. CNNs use real time or still images to distinguish between normal and abnormal, then further investigate abnormal objects to identify a diagnosis with relatively highly accuracy and efficacy (Table 1).

Table 1 Artificial intelligence methods for gastrointestinal angiodysplasia detection[17]

Artificial intelligence	Description	Function	Advantages	Disadvantages
Machine learning	Ability of a computer program to learn	Discern logic-based rules from input and output data	Automation of tasks	Requires high-quality data likely to have some causal link
	Algorithm workflow improves performance		Detect patterns between input and output data	
Artificial neural network	Use of weighted/graded signals to perceive data	Adaptive learning	Mapping performance between input and output data	Requires labeled data
	Use of computational communication		Adaptive learning capability	Requires large volumes of data
Convolutional neural network	Image detection	Computer vision	Highly accurate image recognition and classification	Highly dependent on a training modelor models
	Interpretation through three-dimensional convolutional layers			Limited by image rotation or orientation

CNNs have become one of the most commonly used AI modalities, particularly in VCE, which has significantly aided in the detection of GIADs. The use of AI, particularly CNNs, has created a new era in capsule endoscopy (CE) capable of improving lesion detection rates, reducing capsule reading time, as well as reducing reviewer fatigue. This shift towards computer-aided diagnostic tools in clinical practice may represent a future of common practice. Further investigation with AI in computer-aided diagnosis of GIAD leans heavily towards CE. Three of the most popular areas of CNN implementation include newly developed algorithms, single-shot multibox detection (SSD) and region of interest (ROI) color contrast analysis.

MODALITIES WHERE AI CAN BE USED WHEN DETECTING GIAD

In 2019, Leenhardt *et al*[7] analyzed 2946 still frames with vascular lesions utilizing CNN, where two data sets were used to create a trained algorithm for GIAD detection. The first dataset, also termed the "training and learning phase," consisted of a CNN analysis of 2946 still frame images containing vascular lesions for characteristic analysis of abnormal lesions based on size, shape, color, pattern, and contour. This helped the CNN distinguish GIADs within a still frame. The second data set utilized the learned features from the previous data set, which were applied to new images to detect and located GIAD within a still frame. The primary and secondary endpoints were the sensitivity and specificity of the computer aided diagnosis (CADx) algorithm. These values were 100% and 96% respectively[7].

Similarly, Hwang *et al*[8] developed their own CNN-based AI model bases on a collection of still images later classified as ulcerative or hemorrhagic, which were augmented by rotating each image by 90 degrees 3 times and flipping each rotated image horizontally. As a result, a collection of 30224 abnormal images (11776 hemorrhagic lesions and 18,448 ulcerative lesions) and 30224 normal images were used to train their CNN model by observing similar outcomes in the Leenhardt *et al*[7] study. However, Hwang *et al*[8] went a step further in developing their own CNN based on VGGnet, a CNN that incorporates more convolution filters or layers when screening an image to improve its accuracy of image recognition[9]. Using two training protocols, Hwang *et al*[8] developed a binary model, trained to detect any pathological images as abnormal without distinguishing the types of lesions, and a combined model, trained to detect distinctive hemorrhage or ulcerative lesions.

Another type of CNN is called SSD which is very similar to CNNs described above. However, in this instance, an expert endoscopist will demarcate a rectangular box around a lesion within an image making it much faster to provide a unifying framework for both training and interpretation[10]. Tsuboi *et al*[11] incorporated this technique with 2237 still images of small-bowel GIAD captured by VCE and placed a bounding box where GIAD were found. Through this method, Tsuboi *et al*[11] were able to test their ability to detect GIAD using an area under the receiver operating characteristic (ROC) curve for the probability score, as well as sensitivity, specificity, PPV and NPV of their CNN's detection rate for GIAD and accurately distinguish their

location within an image. Lin *et al*[12] delved deeper into this approach by combining SSD with RetinaNet, a CNN that mimics VGGnet described above, with the enhanced ability to find shortcuts when comparing images in order to limit the number of layers used when training. Otani *et al*[13] was able to analyze and characterize images of erosions and ulcers, GIAD and tumors, then compared the ROC, sensitivity, specificity, and accuracy of their AI detection system for each lesion image.

Another prevalent area of CNN performance is color contrast analysis. Since color is one of the most relevant features in diagnosing GIAD, Noya *et al*[14] used color to detect potential regions of GIAE within an image. This is done in 4 categorized steps: Image preprocessing (contrast enhancement), selection of potential ROI (geometric outline of colored pixels making up the angiodysplastic lesion), feature extraction and selection (labeling a ROI based on color, texture and geometric pattern) and classification of a ROI (recognizing patterns of potential angiodysplasia lesions as pathological vs. non-pathological). Comparably, Iakovidis and Koulaouzidis[15] use color-based pattern recognition to separate pathological vs. normal lesions from 137 still images, which they placed into four categories: vascular, inflammatory, lymphangiectatic, and polypoid. Iakovidis and Koulaouzidis[15] used a 4-step categorization process, like Noya *et al*[14] above, however, they differ with the introduction of salient point saturation (SPS), an automated extraction algorithm which selects salient points in a digital image based on changes in observed color intensity[16].

OUTCOMES OF AI IN DETECTING GIAD

The effects of AI computer-aided diagnosis in GIAD are producing promising results that individual practitioners may hope to incorporate into their practices. The diagnostic yield of GIADs using AI leans heavily on VCE with the use of CNNs. Newly developed algorithms, such as SSD and ROI color contrast analysis have been areas of particular focus in medical literature. Each modality of these CNN implementing tools stands on their own, as very limited research compares these techniques by using the same data set or still images for a head-to-head comparison.

The diagnostic performance of a CADx algorithm for the detection of GIAD using VCE, assess its diagnostic precision as a means for a segmental approach in localizing lesions. Leenhardt *et al*[7] found a sensitivity of 100% [95% confidence interval (CI), 100%-100%]. Secondary endpoints revealed a specificity of 96.0% (95%CI: 93.78%-98.22%), a positive predictive value of 96.15% (95%CI: 93.97%-98.33%), a negative predictive value of 100.0% (95%CI: 100%-100%) and a kappa coefficient of reproducibility at 1.0[7]. Only "clean" images were used in their data set, which meant that images with poor preparation quality or the presence of bubbles would not be included. This is a limitation to the study, which the authors point to. In comparison, the algorithm of Hwang *et al*[8] combined (all images trained separately as hemorrhagic or ulcerative) *vs* binary training (all images trained without segregation) approach in the development of an automated CNN, demonstrated that combined training revealed higher sensitivity (97.61% *vs* 95.07%, $P < 0.001$). Although, accuracy classifying GIADs as small bowel lesions was 100% in both the combined and binary training models.

The use of SSD by Tsuboi *et al*[11] to automatically detect GIAD in VCE images focuses on diagnostic accuracy utilizing t-test analysis. The study reported a ROC curve for CNN detection of GIAD at 0.999. The cut-off value for the probability score was 0.36, exhibiting a sensitivity, specificity, positive predictive value, and negative predictive value of their CNN at 98.8%, 98.4%, 75.4%, and 99.9% respectively at this value[11]. Otani *et al*[13] enhanced CNN by combination of SSD with RetinaNet detection of vascular lesions displayed an AUC 0.950 (95%CI: 0.923-0.978) among the internal cohort (images obtained for training) and 0.884 (95%CI: 0.874-0.893) among the external cohort (randomly obtained imaged for cross-validation). This is an observable difference compared to Tsuboi *et al*[11] study, although still relatively high in automated lesion detection.

Color contrast has been used as well. Iakovidis and Koulaouzidis[15] assessed the validity of color-based pattern recognition in the classification of pathologic lesions with the addition of SPS, including p0 GIAD (low probability of bleeding), p1 GIAD (intermediate probability of bleeding) and p2 GIAD (high probability of bleeding). Classification per type of GIAD revealed an AUC of 69.9 ± 15.8 (P0 GIAD), 97.5 ± 2.4 (P1 GIAD), and 79.6 ± 13.1 (P3 GIAD) respectively[15]. Noya *et al*[14] used the combination of a color-based, texture, statistical and morphological features analysis for GIAD detection. Utilization of this method led to a sensitivity of 89.51% and a

specificity of 96.8%, as well as an AUC 82.33% ± 10.43% detection of GIAD[14].

CONCLUSION

GIADs are a significant cause of GI bleeding and are the main cause for suspected small bowel bleeding. To make the diagnosis, gastroenterologists rely on the use of VCE to "target" GIAD. However, the use of VCE can be cumbersome secondary to reader fatigue, suboptimal preparation, and difficulty in distinguishing images. Humans are imperfect. The human eye is imperfect. The same capsule read by two different readers are noted to have miss rates like other forms of endoscopy. The use of AI in VCE have shown that detection has improved, however the other burdens and limitations still need to be addressed. AI used for the diagnosis of GIAD shows promise and the changes needed to enhance the current practice of VCE are near.

REFERENCES

1 **Jackson CS**, Strong R. Gastrointestinal Angiodysplasia: Diagnosis and Management. *Gastrointest Endosc Clin N Am* 2017; **27**: 51-62 [PMID: 27908518 DOI: 10.1016/j.giec.2016.08.012]

2 **Bollinger E**, Raines D, Saitta P. Distribution of bleeding gastrointestinal angioectasias in a Western population. *World J Gastroenterol* 2012; **18**: 6235-6239 [PMID: 23180943 DOI: 10.3748/wjg.v18.i43.6235]

3 **Buscaglia JM**, Giday SA, Kantsevoy SV, Clarke JO, Magno P, Yong E, Mullin GE. Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study. *Clin Gastroenterol Hepatol* 2008; **6**: 298-301 [PMID: 18255353 DOI: 10.1016/j.cgh.2007.12.029]

4 **Xavier S**, Monteiro S, Magalhães J, Rosa B, Moreira MJ, Cotter J. Capsule endoscopy with PillCamSB2 versus PillCamSB3: has the improvement in technology resulted in a step forward? *Rev Esp Enferm Dig* 2018; **110**: 155-159 [PMID: 29278000 DOI: 10.17235/reed.2017.5071/2017]

5 **Shiotani A**, Honda K, Kawakami M, Kimura Y, Yamanaka Y, Fujita M, Matsumoto H, Tarumi K, Manabe N, Haruma K. Analysis of small-bowel capsule endoscopy reading by using Quickview mode: training assistants for reading may produce a high diagnostic yield and save time for physicians. *J Clin Gastroenterol* 2012; **46**: e92-e95 [PMID: 22495816 DOI: 10.1097/MCG.0b013e31824fff94]

6 **Handelman GS**, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med* 2018; **284**: 603-619 [PMID: 30102808 DOI: 10.1111/joim.12822]

7 **Leenhardt R**, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, Becq A, Marteau P, Histace A, Dray X; CAD-CAP Database Working Group. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc* 2019; **89**: 189-194 [PMID: 30017868 DOI: 10.1016/j.gie.2018.06.036]

8 **Hwang Y**, Lee HH, Park C, Tama BA, Kim JS, Cheung DY, Chung WC, Cho YS, Lee KM, Choi MG, Lee S, Lee BI. Improved classification and localization approach to small bowel capsule endoscopy using convolutional neural network. *Dig Endosc* 2021; **33**: 598-607 [PMID: 32640059 DOI: 10.1111/den.13787]

9 **Somonyan K**, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014 Preprint. Available from: arXiv:1409.1556

10 **Liu W**, Dragomir A, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC. SSD: Single Shot MultiBox Detector. In: Leibe B, Matas J, Sebe N, Welling M. Computer Vision – ECCV 2016. ECCV 2016. Berlin, Heidelberg: Springer, 2016 [DOI: 10.1007/978-3-319-46448-0_2]

11 **Tsuboi A**, Oka S, Aoyama K, Saito H, Aoki T, Yamada A, Matsuda T, Fujishiro M, Ishihara S, Nakahori M, Koike K, Tanaka S, Tada T. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig Endosc* 2020; **32**: 382-390 [PMID: 31392767 DOI: 10.1111/den.13507]

12 **Lin TY**, Goyal P, Girshick R, He K, Dollar P. Focal Loss for Dense Object Detection. *IEEE Trans Pattern Anal Mach Intell* 2020; **42**: 318-327 [PMID: 30040631 DOI: 10.1109/TPAMI.2018.2858826]

13 **Otani K**, Nakada A, Kurose Y, Niikura R, Yamada A, Aoki T, Nakanishi H, Doyama H, Hasatani K, Sumiyoshi T, Kitsuregawa M, Harada T, Koike K. Automatic detection of different types of small-bowel lesions on capsule endoscopy images using a newly developed deep convolutional neural network. *Endoscopy* 2020; **52**: 786-791 [PMID: 32557474 DOI: 10.1055/a-1167-8157]

14 **Noya F**, Alvarez-Gonzalez MA, Benitez R. Automated angiodysplasia detection from wireless capsule endoscopy. *Annu Int Conf IEEE Eng Med Biol Soc* 2017; **2017**: 3158-3161 [PMID: 29060568 DOI: 10.1109/EMBC.2017.8037527]

15 **Iakovidis DK**, Koulaouzidis A. Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software. *Gastrointest Endosc* 2014; **80**: 877-

883 [PMID: 25088924 DOI: 10.1016/j.gie.2014.06.026]

16 **Bay H**, Ess A, Tuytelaars T, Van Gool L. SURF: Speeded-Up Robust Features. In: Computer Vision – ECCV 2006. Berlin, Heidelberg: Springer, 2006: 404-417 [DOI: 10.1007/11744023_32]

17 **Chang AC**. Chapter 1 - Basic Concepts of Artificial Intelligence. In: Intelligence-Based Medicine: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare. Academic Press, Elsevier B.V., 2020: 7-22 [DOI: 10.1016/B978-0-12-823337-5.00001-9]

AIGE Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 August 28; 2(4): 185-197

DOI: 10.37126/aige.v2.i4.185

ISSN 2689-7164 (online)

MINIREVIEWS

Early gastrointestinal cancer: The application of artificial intelligence

Hang Yang, Bing Hu

ORCID number: Hang Yang 0000-0002-7235-8162; Bing Hu 0000-0002-9898-8656.

Author contributions: All authors participated in the work; Yang H contributed to the design and draft of the manuscript; Hu B contributed to the design and review of the manuscript.

Supported by the 1·3·5 project for disciplines of excellence Clinical Research Incubation Project, West China Hospital, Sichuan University, China, No. 20HXFH016.

Conflict-of-interest statement: The authors declare that they have no competing interests.

Manuscript source: Unsolicited manuscript

Hang Yang, Bing Hu, Department of Gastroenterology, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, China

Corresponding author: Bing Hu, MD, Professor, Department of Gastroenterology, West China Hospital, Sichuan University, No. 37 Guo Xue Xiang, Wu Hou District, Chengdu 610041, Sichuan Province, China. hubingnj@163.com

Abstract

Early gastrointestinal (GI) cancer has been the core of clinical endoscopic work. Its early detection and treatment are tightly associated with patients' prognoses. As a novel technology, artificial intelligence has been improved and applied in the field of endoscopy. Studies on detection, diagnosis, risk, and prognosis evaluation of diseases in the GI tract have been in development, including precancerous lesions, adenoma, early GI cancers, and advanced GI cancers. In this review, research on esophagus, stomach, and colon was concluded, and associated with the process from precancerous lesions to early GI cancer, such as from Barrett's esophagus to early esophageal cancer, from dysplasia to early gastric cancer, and from adenoma to early colonic cancer. A status quo of research on early GI cancers and artificial intelligence was provided.

Key Words: Artificial intelligence; Early esophageal cancer; Early gastric cancer; Early colonic cancer

Core Tip: Diagnosis and management of early gastrointestinal (GI) cancer is one of the cores of clinical practice. Endoscopy is the indispensable tool for standard surveillance and management. Artificial intelligence is a novel technology used in some fields of cancer including early GI cancer. Therefore, we provide an overview and introduce how artificial intelligence can be applied to endoscopy on early GI cancer mainly including esophagus, stomach, and colon from the point of view of the clinical diagnosis and management guidelines. Studies with quality control on the diagnosis and management of early GI cancer and their precancerous lesions have also been concluded.

Citation: Yang H, Hu B. Early gastrointestinal cancer: The application of artificial intelligence. *Artif Intell Gastrointest Endosc* 2021; 2(4): 185-197

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C, C, C, C
Grade D (Fair): D, D
Grade E (Poor): 0

Received: June 11, 2021
Peer-review started: June 11, 2021
First decision: June 24, 2021
Revised: June 25, 2021
Accepted: August 18, 2021
Article in press: August 18, 2021
Published online: August 28, 2021

P-Reviewer: Balakrishnan DS, Lalmuanawma S, Tanabe S, Vijh S
S-Editor: Liu M
L-Editor: Filipodia
P-Editor: Xing YX

URL: https://www.wjgnet.com/2689-7164/full/v2/i4/185.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i4.185

INTRODUCTION

Artificial intelligence (AI) is essentially a process of learning human thinking and transferring human experience. Recognizing images based on artificial neural networks/convolutional neural networks (CNNs) is one of the novel and main fields of AI. Computer-aided diagnosis (CAD) systems are designed to interpret medical images using advances in AI from method learning to deep learning (DL) and includes mainly three groups (CADe, CADx, and CADm)[1].

AI has been widely involved in cancer[2]. In regard to digestive cancer, it has been utilized to find more intelligent ways to facilitate detection, diagnosis, risk evaluation, and prognosis. For instance, radiomics machine learning signature for diagnosing hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules was also validated in a multicenter retrospective cohort, which could enhance clinicians' decisions[3].

In the aspect of pancreatic cancer, it continues to be one of the deadliest malignancies with less than 10% overall survival rate. Survival rates will increase if pancreatic cancer can be detected at an early stage[4]. Intraductal papillary mucinous neoplasms are precursor lesions of pancreatic adenocarcinoma. A DL model was shown to be a more accurate and objective method to diagnose malignancies of intraductal papillary mucinous neoplasms in comparison to human diagnosis and conventional endoscopic ultrasonography (EUS) images[5]. Pancreatic cystic lesions are also precursors of pancreatic cancer. Radiomics utilizing quantitative image analysis to extract features in conjunction with machine learning and AI methods helped differentiate benign pancreatic cystic lesions from malignant ones[6]. An artificial neural network was trained to help predict pancreatic ductal adenocarcinoma based on gene expression[7]. An AI-assisted CAD system using DL analysis of EUS images was efficient to help detect pancreatic ductal carcinoma[8]. The artificial neural network model could accurately predict the survival of pancreatic adenocarcinoma patients as a useful objective decision tool in complex treatment decisions[9].

In this review, we concluded the application and research of AI based on endoscopic examination related to early gastrointestinal (GI) cancer mainly including esophagus, stomach, and colon. The progression of carcinogenesis from Barrett's esophagus (BE) to early esophageal cancer (EEC), from dysplasia to early gastric cancer (EGC), and from adenoma to early colonic cancer (ECC) were reviewed in detailed as well as related AI research on the histopathology and invasion depth detection of these GI cancer.

LITERATURE SEARCH

This review was aimed to make a qualitative only review of the application of AI on early GI cancer. We searched the PubMed database for articles that were published in the last 5 years using the term combinations of AI/DL and EEC, esophageal squamous cell carcinoma (ESCC), esophageal adenocarcinoma (EAC), EGC, and ECC for early GI cancer, and term combinations of AI/DL and precancerous lesions [BE/dysplasia/chronic atrophic gastritis (CAG)/gastric intestinal metaplasia/*Helicobacter pylori*/adenoma/polyp/inflammatory bowel diseases] for precancerous lesions of early GI cancer. Endoscopic-related results were qualitatively concluded in Table 1.

SEARCH RESULTS

Initially, a total of 424 articles were identified. After manually screening and reading, 22 studies were tabulated in Table 1, and 2 prospective studies on detecting adenoma were also added in Table 1. Meanwhile, 13 studies on precancerous lesions of early GI cancer were showed in the review. The flowchart was presented in Figure 1.

Table 1 Early gastrointestinal cancer and artificial intelligence

Ref.	Target disease	Prospective/ retrospective	AI	Endoscopy image	Training dataset	Validation dataset	Sensitivity	Specificity	Accuracy[1] /AUC
[1]	Diagnosing ESCC and EAC	Retrospective	CNNs (SSD)	WLI and NBI	8428 images	1118 images	98%	95%	98%[1]
[2]	Diagnosing ESCC	Retrospective	CAD (SegNet)	NBI/videos	6473 images	6671 images	98.04%	95.03%	0.989
[3]	Detecting EEC and BE	Retrospective	CAD (ResNet-UNet)	WLI	494364 images	1704 images	90%	88%	89%[1]
[4]	Detecting E/J cancers	Retrospective	CNNs (SSD)	WLI and NBI	3443 images	232 images	94%	42%	66%[1]
[5]	Detecting ESCC	Retrospective	DCNNs-CAD	NBI	2428 images	187 images	97.80%	85.40%	91.4%[1]
[6]	Diagnosing BE and EAC	Retrospective	CAD (ResNet)	WLI and NBI	148/100	Leave-one patient-out cross validation	97%(WLI)/94%(NBI)	88% (WLI)/80%(NBI)	
[7]	Diagnosing ESCC	Retrospective	CAD (FCN)	ME-NBI		3-fold cross-validation			
[8]	Detecting EAC	Retrospective	CNNs (SSD)	WLI		100 images	96%	92%	
[9]	Detecting EGC	Retrospective	CNNs	WLI	348943 images	9650 images	80.00%	94.80%	
[10]	Diagnosing EGC	Retrospective	CNNs	WLI	21217 images	1091 images	36.8	91.20%	
[11]	Diagnosing EGC	Retrospective	CNNs (Inception-v3)	ME-NBI	1702 images	170 images	91.18%	90.64%	90.91%[1]
[12]	Diagnosing EGC	Retrospective	CNNs (VGG16)	WLI	896 t1a-EGC and 809 t1b-EGC	5-fold cross-validation			Detection (0.981) Depth prediction (0.851)
[13]	Detecting EGC	Retrospective	CNNs (VGG16 and ResNet-50)	WLI/NBI/BLI	3170 images		94.00%	91.00%	92.5%[1]
[14]	Diagnosing EGC	Retrospective	CNNs (ResNet-50)	WLI	790 images	203 images	76.47%	95.56%	89.16%[1]
[15]	Detecting EGC	Retrospective	CNNs (SSD)	WLI	13584 images	2940 images	58.40%	87.30%	0.76
[16]	Classifying EGC	Retrospective	CNNs (Inception-ResNet-v2)	WLI	5017 images	5-fold cross-validation			0.85
[17]	Diagnosing EGC	Retrospective	CNNs (ResNet-50)	ME-NBI	4460 images	1114 images	98%	100%	98.7%[1]
[18]	Detecting and localizing colonic adenoma	Representative	CNNs (VGG16,19, ResNet50)	WLI and NBI	8641 images/9 videos, 11 videos	Cross-validation			
[19]	Detecting ECC	Representative	CNNs	WLI	190 images	3-fold cross-validation	67.50%	89.00%	81.2%[1] /0.871
[20]	Classifying ECC	Representative	CNNs (ResNet-152)	WLI		3-fold cross-validation	95.40%	30.10%	
[21]	Detecting colonic	Prospective	Cade	1058 patients	ADR (29.1% *vs* 20.3%)				

	adenoma				
[22]	Detecting colonic adenoma	Prospective	Cade	962 patients	ADR (34% *vs* 28%)

[1]Accuracy is with "[1]" and AUC is without "[1]", *e.g.*, 100%[1] means accuracy is 100%.
ADR: Adenoma detection rates; AI: Artificial intelligence; AUC: Area under the curve; BE: Barrett's esophagus; BLI: Bright light imaging; CAD: Computer-aided diagnosis; CNN: Convolutional neural network; DCNN: Deep convolutional neural network; EAC: Esophageal adenocarcinoma; ECC: Early colonic cancer; EEC: Early esophageal cancer; EGC: Early gastric cancer; E/J: Esophagogastric junctional; ESCC: Esophageal squamous cell carcinoma; ME-NBI: Magnifying narrow band imaging; NBI: Narrow-band imaging; SSD: Single-Shot Multibox Detector; WLI: White-light imaging.

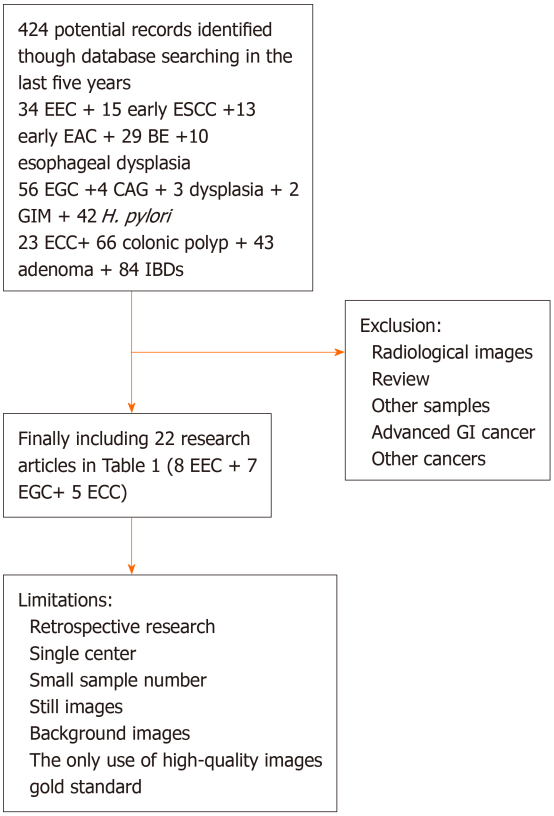

Figure 1 Flow chart of study selection and logic arrangement of review. BE: Barrett's esophagus; CAG: Chronic atrophic gastritis; EAC: Esophageal adenocarcinoma; ECC: Early colonic cancer; EEC: Early esophageal cancer; EGC: Early gastric cancer; ESCC: Esophageal squamous cell carcinoma; GI: Gastrointestinal; GIM: Gastric intestinal metaplasia; *H. pylori*: *Helicobacter pylori*; IBD: Inflammatory bowel diseases.

AI AND EEC FROM PRECANCEROUS LESIONS TO EEC

Esophageal cancer is one of most common cancers related to a considerable decline in health-related quality of life and a reduction in survival rate. ESCC and EAC are two main histological types. Many patients with ESCC have a history of heavy tobacco and alcohol use[10] as well as other risk factors including polycyclic aromatic hydrocarbons, high-temperature foods, diet, oral health, microbiome, and genetic factors [11]. Some risk factors for EAC have been considered mainly as gastroesophageal reflux disease, BE, obesity, and tobacco smoking as well as genetic variants[12]. Chronic gastroesophageal reflux disease can cause metaplasia from the native squamous cell mucosa to a specialized columnar epithelium[13]. BE and dysplasia in squamous epithelium are precancerous lesions to EAC and ESCC, respectively, and they are supposed to be as one of the main aims of early diagnosis. Endoscopic diagnosis of EEC, white-light imaging (WLI), iodine staining, narrow-band imaging (NBI), and biopsy have been widely used clinically[14].

There is also study on AI being involved in preclinical stage. For instance, the diagnostic ability of AI using DL to detect esophageal cancer including superficial and advanced squamous cell carcinoma and adenocarcinoma was characterized as highly sensitive (98%) and efficient based on WLI images. Small cancer lesions less than 10

mm in size could be detected[15].

In terms of EAC, AI using DL to diagnose superficial esophagogastric junctional adenocarcinoma showed favorable sensitivity (94%) and acceptable specificity (42%) of WLI images compared with experts[16]. A CAD using DL (CAD-DL) model was trained by two datasets based on two different kinds of images (WLI and NBI images) used to detect early EAC. The diagnosis of EAC by CAD-DL reached sensitivities/specificities of 97%/88% for WLI images and sensitivities/specificities of 94%/80% for NBI images, respectively (Augsburg dataset) and 92%/100% (another dataset) for WLI images[17]. Additionally, one research compared several AI methods including regional-based CNN (R-CNN), Fast R-CNN, Faster R-CNN, and Single-Shot Multibox Detector. Single-Shot Multibox Detector outperformed other methods achieving a sensitivity of 96% in automatically identify EAC[18].

In terms of ESCC, the endocytoscopic system (ECS) helps in virtual realization of histology. The CNN method was applied to detect ESCC with an overall sensitivity of 92.6% based on ECS images aimed at replacing biopsy-based histology[19]. NBI is currently regarded as the standard modality for diagnosing ESCC. A CNN model was applied to detect ESCC based on NBI images and showed significantly higher sensitivity (91%), specificity (51%), and accuracy (63%) than those of endoscopic experts[20]. Besides NBI and ECS, AI was also applied in magnified endoscopy (ME). The accuracy, sensitivity, and specificity of AI based on ME images were 89%, 71%, and 95% for the AI system, respectively[21]. Accuracy, sensitivity, and specificity with WLI images were 87%, 50%, and 99%, respectively. Furthermore, as endoscopic resection (ER) is often used to treat ESCC when invasion depths are diagnosed as intraepithelial–submucosal layer (tumor invasion is within 0.5 mm of the muscularis mucosae). The invasion depth of superficial ESCC was also calculated by a CNN method based on WLI and NBI images, which demonstrated higher accuracy. The diagnosis accuracy of the CNN method was higher in the intraepithelial-lamina propria and muscularis mucosa groups (91.2% and 91.4%, respectively) than that in the submucosal layer group (67.8%)[22].

Recently, there have been some application and research of AI on precursor lesions of EEC including BE and dysplasia in squamous epithelium. For instance, AI could enhance the image of volumetric laser endomicroscopy to facilitate the surveillance BE [23]. The CNN method was developed to recognized early esophageal neoplasia in BE. It could correctly detect early neoplasia with the sensitivity of 96.4%, the specificity of 94.2%, and the accuracy of 95.4%. In addition, the object detection algorithm was able to draw a localization box around areas of dysplasia with a mean average accuracy of 75.33% and sensitivity of 95.60%[24]. Another similar research demonstrated that a CAD system used five independent endoscopy datasets to detect early neoplasia in patients with BE. In dataset 4, the CAD classified images as containing neoplasms or non-dysplastic BE with 89% accuracy, 90% sensitivity, and 88% specificity. The CAD also identified the optimal site for biopsy of detected neoplasia in 97% of cases in dataset 4[25].

Moreover, AI was also applied in esophageal histopathology; attention-based deep neural networks were used to detect cancerous and precancerous esophagus tissue on histopathological slides. Classification accuracies of the proposed model were 85% for the BE-no-dysplasia class, 89% for the BE-with-dysplasia class, and 88% for the adenocarcinoma class[26].

AI AND EGC FROM CAG AND DYSPLASIA TO EGC

EGC is defined as a cancer confined to the mucosa or submucosa, regardless of lymph node metastasis (LNM). Standard WLI and image enhancement endoscopy, such as NBI and ME, have been widely used in screening and surveillance of EGC as well as EUS, which can enable the precise assessment of the risk of LNM of EGC[27]. Risk factors include *Helicobacter pylori* infection, age, high salt intake, diets low in fruit and vegetables, and genetic factors[28]. ER is a minimally invasive treatment for EGC with negligible risk of LNM[29]. Patients with CAG, intestinal metaplasia, or dysplasia are at risk for gastric adenocarcinoma and are recommended to accept the regular endoscopic surveillance. Virtual chromoendoscopy can guide biopsies for staging atrophic and metaplastic changes and can target neoplastic lesions[30]. The 5-year survival rate of EGC patients is significantly higher than that of advanced GC patients [31,32]. Early detection and treatment are always one of the top priorities.

In regard to the application of AI in EGC, there are some considerations both related on the promise such as the benefits for endoscopists and patients and limitations[33].

To detect and diagnose EGC *via* ME with NBI (ME-NBI) requires considerable experience; AI-assisted CNN CAD system based on ME-NBI images was constructed to diagnose EGC, and the overall accuracy, sensitivity, and specificity of the CNN were 98.7%, 98.0%, and 100%, respectively, in a short period of time[34]. Different deep CNN methods have been designed (such as VGG, Single-Shot Multibox Detector, and ResNet) based on different image types (such as WLI, NBI, and chromoendoscopy) and mucosal backgrounds (normal mucosa, superficial gastritis, and erosive mucosa) (shown in Table 1). There was also research on differentiating EGC from gastritis[35] and peptic ulcer[36] achieving reliable accuracy.

Moreover, training with video is considered to improve accuracy in a real clinical setting. A CNN model based on videos demonstrated a high detection rate (94.1%) with a high processing speed[37]. Furthermore, CNN-CAD was applied to diagnose the invasion depth of GC based on WLI images and distinguish EGC from advanced GC, with the sensitivity of 76.47%, specificity of 95.56%, and accuracy of 89.16%[38]. Another model was also involved in invasion depth. For instance, a CNN method (lesion-based VGG-16 model) was used to classify EGC with of sensitivity (91.0%), specificity (97.6%), and accuracy (98.1%), respectively. The prediction of invasion depth achieved sensitivity (79.2%), specificity (77.8%), and accuracy (85.1%), respectively, higher than results of non-lesion-based models, indicating a lesion-based CNN was an appropriate training method for AI in EGC[39].

In terms of histopathology, a CNN model trained with pixel-level annotated hematoxylin and eosin stained whole slide images achieved a sensitivity near 100% and an average specificity of 80.6% in diagnosing GC, aimed at alleviating the workload and increasing diagnostic accuracy[40]. Similarly, AI automatically classified GC in hematoxylin and eosin stained histopathological whole slide images from different groups and demonstrated favorable results[41,42]. Besides endoscopic images, machine learning based on radiographic-radiomic images could help predict adverse histopathological status of GC[43]. Dual-energy computed tomography based DL radiomics could improve LNM risk prediction for GC[44]

In the aspect of gastric precancerous conditions, the application of AI has also been focused. For example, atrophic gastritis, as a kind of precancerous condition was diagnosed by the pretrained CNN based on WLI images achieved an accuracy of 93% in an independent dataset, outperforming expert endoscopists[45]. The CNN method was trained by WLI images of gastric antrum in diagnosing CAG, and the diagnostic accuracy, sensitivity, and specificity were 94.2%, 94.5%, and 94.0%, respectively, which were higher than those of experts. The further detection rates of mild, moderate, and severe atrophic gastritis were 93%, 95%, and 99%, respectively[46]. *Helicobacter pylori* infection, as a dominant cause of CAG and GC, has also been detected *via* AI method based on endoscopic images, such as CNN (GoogLeNet) and CNN (ResNet-50 model), and achieved the higher accuracy and reliability in a considerably shorter time[47-49].

AI AND ECC FROM POLYPS AND ADENOMA TO ECC

ECC has been defined as a carcinoma with invasion limited to the submucosa regardless of lymph node status and according to the Royal College of Pathologists as TNM stage $T_1N_XM_0$[50]. If the dysplasia is restricted to the layer of epithelium, it is defined as low-grade or high-grade intraepithelial neoplasia. Mild or moderate dysplasia is the pathological character of low-grade intraepithelial neoplasia, and severe dysplasia is the pathological character of high-grade intraepithelial neoplasia or preinvasive carcinoma[51]. Colonic precancerous lesions include traditional serrated adenoma and sessile serrated adenoma/polyps[52,53]. The submucosal invasion in clinical practice is considered as the superficial depth of tumor invasion and further as a surrogate for nominal LNM risk. Meanwhile, it can be a general criterion to identify whether patients are eligible for local ER or surgery[54]. Curative ER is indicated for lesions confined to the mucosal layer or invading less than 1 mm into the submucosal layer[50]. Endoscopic screening is proven to decrease the risk of disease-specific morbidity and mortality[55]. Current guidelines recommend screening beginning at age 50 and continuing until age 75 with fecal immunochemical test every year, flexible sigmoidoscopy every 5 years, and/or colonoscopy every 10 years[56]. Early diagnosis and treatment are pivotal. When colon carcinoma is detected in a localized stage, the 5-year relative survival is 91.1%. However, the 5-year relative survival of colon carcinoma patients with regional metastasis or distant metastasis were 71.7% and 13.3%, respectively[57].

AI has been widely involved in the research of ECC on the aspect of detection, diagnosis, classification, invasion depth, and histopathology as well as inflammatory bowel diseases associated with inflammation-dysplasia-colon cancer pattern. Regarding the detection and diagnosis, a research trained Faster R-CNN with VGG16 based on WLI images and videos covering ECC (Tis or T_1) and precursor lesions including hyperplastic polyps, sessile serrated adenoma/polyps, traditional serrated adenoma, low-grade intraepithelial neoplasia, high-grade intraepithelial neoplasia, and submucosal invasive cancer was conducted. It showed the sensitivity and specificity were 97.3% and 99.0%, respectively[58]. Another research used two CNN methods trained by WLI images. ResNet-152 showed a higher mean area under the curve for detecting tubular adenoma + lesions (0.818), and the mean area under the curve for detecting high-grade intraepithelial neoplasia + lesions reached 0.876 by ResNet-v2[59]. Regarding the invasion depth, for deeply invasive cT_1 (SM) (hereafter, cT_{1b}) or deeper colorectal cancer (CRC), there is a 10%–15% or higher risk of lymph node metastases. Further surgical resection including lymph node dissection is required[60]. For an accurate depth of invasion diagnosis, the CNN method was used to assist in cT_{1b} diagnosis and demonstrated that cT_{1b} sensitivity, specificity, and accuracy were 67.5%, 89.0%, and 81.2%, respectively[61].

In the research of AI application in precancerous lesions such as polyps, there has been some research of AI, especially retrospective research related to polyp detection and diagnosis with high accuracy[62,63]. For example, a local-feature-prioritized automatic CADe system could detect laterally spreading tumors and sessile serrated adenoma/polyps with high sensitivity from 85.71% to 100%[64]. Besides retrospective research, AI has been designed into some associated prospective research. For instance, a multicenter randomized trial used CAD to detect colorectal neoplasia. It showed a significant increase in adenoma detection rates and adenomas detected per colonoscopy without increasing withdrawal time (54.8% *vs* 40.4%). Additionally, the detection rate of adenomas 5 mm or smaller was significantly higher in the CAD group (33.7%) than in the control group[65]. Another randomized study used CAD to detect adenomas and achieved increased adenoma detection rates (29.1% *vs* 20.3%) and the mean number of adenomas per patient (0.53 *vs* 0.31). Similarly, a higher number of diminutive adenomas were found (185 *vs* 102)[66]. In addition, inflammatory bowel diseases including Crohn's disease and ulcerative colitis are also associated precancerous lesions, and some AI methods aiding in scoring have been trained, such as DL model in grading endoscopic disease severity of patients with ulcerative colitis[67] and in predicting remission in patients with moderate to severe Crohn's disease[68].

In the aspect of histopathology, AI has been used in ECC and precancerous lesions. A systematic review has concluded that AI use in CRC pathology image analysis included gland segmentation, tumor classification, tumor microenvironment characterization, and prognosis prediction[69]. A DL approach was developed to recognize four different stages of cancerous tissue development, including normal mucosa, early preneoplastic lesion, adenoma, and cancer and obtained an overall accuracy more than 95%[70]. Prediction of LNM for early CRC is critical for determining treatment strategies after ER. An LNM prediction algorithm for submucosal invasive (T_1) CRC based on machine learning showed better LNM predictive ability than the conventional method on some datasets[71-82].

PROSPECTS AND CHALLENGES OF AI APPLICATION ON EARLY GI CANCER

Endoscopy is usually the first choice in the diagnosis and management of early GI cancer. According to the Clinical Practice Guideline, ER is now a standard treatment for early GI cancers without regional LNM. Early GI cancers can completely be removed by *en bloc* fashion (resection of a tumor in one piece without visible residual tumor) *via* endoscopic mucosal resection and/or endoscopic submucosal dissection. High-definition white light endoscopy, chromoendoscopy, and image-enhanced endoscopy such as ME-NBI can be used to assess the edge and depth of early GI cancers for delineation of resection boundaries and prediction of the possibility of LNM before the decision of ER. Histopathological evaluation can confirm the depth of cancer invasion and lymphovascular invasion[83]. From this review, we can see AI as a novel technology has been penetrated in early GI cancer detection, diagnosis, boundaries, invasion depth, lymphovascular invasion, and prognosis prediction based on endoscopic images and videos and pathological tissue slides obtained after ER.

Both high-quality endoscopy and high-quality AI model construction research are crucial to ensure better health outcomes and benefits of patients. Some AI methods have been designed to identify and assure the quality of endoscopy to improve the detection rate of early GI cancer. In upper GI tract, missed EGC rates are an important measure of quality. A deep CNN model was built to monitor blind spots, time the procedure, and automatically generate photo-documentation during esophagogastroduodenoscopy[84]. Meanwhile, in colonoscopy, poorer adenoma detection rates are associated with poorer outcomes and higher rates of post-colonoscopy colonic cancer[85]. A deep CNN model was developed for timing withdrawal phase, supervising withdrawal stability, evaluating bowel preparation, and detecting colorectal polyps[86].

In the aspect of quality control of AI studies related to endoscopy, some limitations should be concerned. Different CNN models have demonstrated high accuracies or area under the curve and 7 out of 22 more than 90%/0.9 with high sensitivities and specificities in Table 1. These limitations were concentrated on the retrospective research, the single center, the small sample number, still images, background images, the only use of high-quality images, and not all images with lesions identified by gold standard such as pathology. They may discount the reliability of the results. As most endoscopic-related algorithms are trained in a supervised manner, labeling data is important. Meanwhile, videos and large, heterogenous, and prospectively collected data are less prone to biases[87].

CONCLUSION

AI has been widely used in medicine, although most studies have remained at the preclinical stage. In this review, we provided an overview of the associated application of AI in early GI cancer including EEC, EGC, and ECC as well as their precancerous lesions. Detection, diagnosis, classification, invasion depth, and histopathology have been involved. Indeed, AI will bring benefits to patients and doctors. It will provide useful support during endoscopies to achieve more precise diagnosis of early GI cancer after more intelligent detection and biopsy with high efficiency and reduce workload to fill the lack of clinical resources in the future.

REFERENCES

1 **Ahmad OF**, Soares AS, Mazomenos E, Brandao P, Vega R, Seward E, Stoyanov D, Chand M, Lovat LB. Artificial intelligence and computer-aided diagnosis in colonoscopy: current evidence and future directions. *Lancet Gastroenterol Hepatol* 2019; **4**: 71-80 [PMID: 30527583 DOI: 10.1016/S2468-1253(18)30282-6]

2 **Huang S**, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Lett* 2020; **471**: 61-71 [PMID: 31830558 DOI: 10.1016/j.canlet.2019.12.007]

3 **Mokrane FZ**, Lu L, Vavasseur A, Otal P, Peron JM, Luk L, Yang H, Ammari S, Saenger Y, Rousseau H, Zhao B, Schwartz LH, Dercle L. Radiomics machine-learning signature for diagnosis of hepatocellular carcinoma in cirrhotic patients with indeterminate liver nodules. *Eur Radiol* 2020; **30**: 558-570 [PMID: 31444598 DOI: 10.1007/s00330-019-06347-w]

4 **Young MR**, Abrams N, Ghosh S, Rinaudo JAS, Marquez G, Srivastava S. Prediagnostic Image Data, Artificial Intelligence, and Pancreatic Cancer: A Tell-Tale Sign to Early Detection. *Pancreas* 2020; **49**: 882-886 [PMID: 32675784 DOI: 10.1097/MPA.0000000000001603]

5 **Kuwahara T**, Hara K, Mizuno N, Okuno N, Matsumoto S, Obata M, Kurita Y, Koda H, Toriyama K, Onishi S, Ishihara M, Tanaka T, Tajika M, Niwa Y. Usefulness of Deep Learning Analysis for the Diagnosis of Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Clin Transl Gastroenterol* 2019; **10**: 1-8 [PMID: 31117111 DOI: 10.14309/ctg.0000000000000045]

6 **Dalal V**, Carmicheal J, Dhaliwal A, Jain M, Kaur S, Batra SK. Radiomics in stratification of pancreatic cystic lesions: Machine learning in action. *Cancer Lett* 2020; **469**: 228-237 [PMID: 31629933 DOI: 10.1016/j.canlet.2019.10.023]

7 **Almeida PP**, Cardoso CP, de Freitas LM. PDAC-ANN: an artificial neural network to predict pancreatic ductal adenocarcinoma based on gene expression. *BMC Cancer* 2020; **20**: 82 [PMID: 32005189 DOI: 10.1186/s12885-020-6533-0]

8 **Tonozuka R**, Itoi T, Nagata N, Kojima H, Sofuni A, Tsuchiya T, Ishii K, Tanaka R, Nagakawa Y, Mukai S. Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study. *J Hepatobiliary Pancreat Sci* 2021; **28**: 95-104 [PMID: 32910528 DOI: 10.1002/jhbp.825]

9 **Walczak S**, Velanovich V. An Evaluation of Artificial Neural Networks in Predicting Pancreatic

Cancer Survival. *J Gastrointest Surg* 2017; **21**: 1606-1612 [PMID: 28776157 DOI: 10.1007/s11605-017-3518-7]

10 **Prabhu A**, Obi KO, Rubenstein JH. The synergistic effects of alcohol and tobacco consumption on the risk of esophageal squamous cell carcinoma: a meta-analysis. *Am J Gastroenterol* 2014; **109**: 822-827 [PMID: 24751582 DOI: 10.1038/ajg.2014.71]

11 **Abnet CC**, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* 2018; **154**: 360-373 [PMID: 28823862 DOI: 10.1053/j.gastro.2017.08.023]

12 **Coleman HG**, Xie SH, Lagergren J. The Epidemiology of Esophageal Adenocarcinoma. *Gastroenterology* 2018; **154**: 390-405 [PMID: 28780073 DOI: 10.1053/j.gastro.2017.07.046]

13 **Spechler SJ**, Souza RF. Barrett's esophagus. *N Engl J Med* 2014; **371**: 836-845 [PMID: 25162890 DOI: 10.1056/NEJMra1314704]

14 **di Pietro M**, Canto MI, Fitzgerald RC. Endoscopic Management of Early Adenocarcinoma and Squamous Cell Carcinoma of the Esophagus: Screening, Diagnosis, and Therapy. *Gastroenterology* 2018; **154**: 421-436 [PMID: 28778650 DOI: 10.1053/j.gastro.2017.07.041]

15 **Horie Y**, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019; **89**: 25-32 [PMID: 30120958 DOI: 10.1016/j.gie.2018.07.037]

16 **Iwagami H**, Ishihara R, Aoyama K, Fukuda H, Shimamoto Y, Kono M, Nakahira H, Matsuura N, Shichijo S, Kanesaka T, Kanzaki H, Ishii T, Nakatani Y, Tada T. Artificial intelligence for the detection of esophageal and esophagogastric junctional adenocarcinoma. *J Gastroenterol Hepatol* 2021; **36**: 131-136 [PMID: 32511793 DOI: 10.1111/jgh.15136]

17 **Ebigbo A**, Mendel R, Probst A, Manzeneder J, Souza LA Jr, Papa JP, Palm C, Messmann H. Computer-aided diagnosis using deep learning in the evaluation of early oesophageal adenocarcinoma. *Gut* 2019; **68**: 1143-1145 [PMID: 30510110 DOI: 10.1136/gutjnl-2018-317573]

18 **Ghatwary N**, Zolgharni M, Ye X. Early esophageal adenocarcinoma detection using deep learning methods. *Int J Comput Assist Radiol Surg* 2019; **14**: 611-621 [PMID: 30666547 DOI: 10.1007/s11548-019-01914-4]

19 **Kumagai Y**, Takubo K, Kawada K, Aoyama K, Endo Y, Ozawa T, Hirasawa T, Yoshio T, Ishihara S, Fujishiro M, Tamaru JI, Mochiki E, Ishida H, Tada T. Diagnosis using deep-learning artificial intelligence based on the endocytoscopic observation of the esophagus. *Esophagus* 2019; **16**: 180-187 [PMID: 30547352 DOI: 10.1007/s10388-018-0651-7]

20 **Fukuda H**, Ishihara R, Kato Y, Matsunaga T, Nishida T, Yamada T, Ogiyama H, Horie M, Kinoshita K, Tada T. Comparison of performances of artificial intelligence versus expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video). *Gastrointest Endosc* 2020; **92**: 848-855 [PMID: 32505685 DOI: 10.1016/j.gie.2020.05.043]

21 **Shimamoto Y**, Ishihara R, Kato Y, Shoji A, Inoue T, Matsueda K, Miyake M, Waki K, Kono M, Fukuda H, Matsuura N, Nagaike K, Aoi K, Yamamoto K, Nakahara M, Nishihara A, Tada T. Real-time assessment of video images for esophageal squamous cell carcinoma invasion depth using artificial intelligence. *J Gastroenterol* 2020; **55**: 1037-1045 [PMID: 32778959 DOI: 10.1007/s00535-020-01716-5]

22 **Tokai Y**, Yoshio T, Aoyama K, Horie Y, Yoshimizu S, Horiuchi Y, Ishiyama A, Tsuchida T, Hirasawa T, Sakakibara Y, Yamada T, Yamaguchi S, Fujisaki J, Tada T. Application of artificial intelligence using convolutional neural networks in determining the invasion depth of esophageal squamous cell carcinoma. *Esophagus* 2020; **17**: 250-256 [PMID: 31980977 DOI: 10.1007/s10388-020-00716-x]

23 **Trindade AJ**, McKinley MJ, Fan C, Leggett CL, Kahn A, Pleskow DK. Endoscopic Surveillance of Barrett's Esophagus Using Volumetric Laser Endomicroscopy With Artificial Intelligence Image Enhancement. *Gastroenterology* 2019; **157**: 303-305 [PMID: 31078625 DOI: 10.1053/j.gastro.2019.04.048]

24 **Hashimoto R**, Requa J, Dao T, Ninh A, Tran E, Mai D, Lugo M, El-Hage Chehade N, Chang KJ, Karnes WE, Samarasena JB. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). *Gastrointest Endosc* 2020; **91**: 1264-1271.e1 [PMID: 31930967 DOI: 10.1016/j.gie.2019.12.049]

25 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929.e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

26 **Tomita N**, Abdollahi B, Wei J, Ren B, Suriawinata A, Hassanpour S. Attention-Based Deep Neural Networks for Detection of Cancerous and Precancerous Esophagus Tissue on Histopathological Slides. *JAMA Netw Open* 2019; **2**: e1914645 [PMID: 31693124 DOI: 10.1001/jamanetworkopen.2019.14645]

27 **Sumiyama K**. Past and current trends in endoscopic diagnosis for early stage gastric cancer in Japan. *Gastric Cancer* 2017; **20**: 20-27 [PMID: 27734273 DOI: 10.1007/s10120-016-0659-4]

28 **Kumar S**, Metz DC, Ellenberg S, Kaplan DE, Goldberg DS. Risk Factors and Incidence of Gastric Cancer After Detection of Helicobacter pylori Infection: A Large Cohort Study. *Gastroenterology* 2020; **158**: 527-536.e7 [PMID: 31654635 DOI: 10.1053/j.gastro.2019.10.019]

29 **Hatta W**, Gotoda T, Koike T, Masamune A. History and future perspectives in Japanese guidelines for endoscopic resection of early gastric cancer. *Dig Endosc* 2020; **32**: 180-190 [PMID: 31529716 DOI: 10.1111/den.13531]

30 **Pimentel-Nunes P**, Libânio D, Marcos-Pinto R, Areia M, Leja M, Esposito G, Garrido M, Kikuste I, Megraud F, Matysiak-Budnik T, Annibale B, Dumonceau JM, Barros R, Fléjou JF, Carneiro F, van Hooft JE, Kuipers EJ, Dinis-Ribeiro M. Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European Society of Gastrointestinal Endoscopy (ESGE), European Helicobacter and Microbiota Study Group (EHMSG), European Society of Pathology (ESP), and Sociedade Portuguesa de Endoscopia Digestiva (SPED) guideline update 2019. *Endoscopy* 2019; **51**: 365-388 [PMID: 30841008 DOI: 10.1055/a-0859-1883]

31 **Suzuki H**, Oda I, Abe S, Sekiguchi M, Mori G, Nonaka S, Yoshinaga S, Saito Y. High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection. *Gastric Cancer* 2016; **19**: 198-205 [PMID: 25616808 DOI: 10.1007/s10120-015-0469-0]

32 **Shi Y**, Xu X, Zhao Y, Qian F, Tang B, Hao Y, Luo H, Chen J, Yu P. Long-term oncologic outcomes of a randomized controlled trial comparing laparoscopic versus open gastrectomy with D2 lymph node dissection for advanced gastric cancer. *Surgery* 2019; **165**: 1211-1216 [PMID: 30772006 DOI: 10.1016/j.surg.2019.01.003]

33 **Mori Y**, Berzin TM, Kudo SE. Artificial intelligence for early gastric cancer: early promise and the path ahead. *Gastrointest Endosc* 2019; **89**: 816-817 [PMID: 30902205 DOI: 10.1016/j.gie.2018.12.019]

34 **Ueyama H**, Kato Y, Akazawa Y, Yatagai N, Komori H, Takeda T, Matsumoto K, Ueda K, Hojo M, Yao T, Nagahara A, Tada T. Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. *J Gastroenterol Hepatol* 2021; **36**: 482-489 [PMID: 32681536 DOI: 10.1111/jgh.15190]

35 **Horiuchi Y**, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Convolutional Neural Network for Differentiating Gastric Cancer from Gastritis Using Magnified Endoscopy with Narrow Band Imaging. *Dig Dis Sci* 2020; **65**: 1355-1363 [PMID: 31584138 DOI: 10.1007/s10620-019-05862-6]

36 **Namikawa K**, Hirasawa T, Nakano K, Ikenoyama Y, Ishioka M, Shiroma S, Tokai Y, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Artificial intelligence-based diagnostic system classifying gastric cancers and ulcers: comparison between the original and newly developed systems. *Endoscopy* 2020; **52**: 1077-1083 [PMID: 32503056 DOI: 10.1055/a-1194-8771]

37 **Ishioka M**, Hirasawa T, Tada T. Detecting gastric cancer from video images using convolutional neural networks. *Dig Endosc* 2019; **31**: e34-e35 [PMID: 30449050 DOI: 10.1111/den.13306]

38 **Zhu Y**, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, Zhang YQ, Chen WF, Yao LQ, Zhou PH, Li QL. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointest Endosc* 2019; **89**: 806-815.e1 [PMID: 30452913 DOI: 10.1016/j.gie.2018.11.011]

39 **Yoon HJ**, Kim S, Kim JH, Keum JS, Oh SI, Jo J, Chun J, Youn YH, Park H, Kwon IG, Choi SH, Noh SH. A Lesion-Based Convolutional Neural Network Improves Endoscopic Detection and Depth Prediction of Early Gastric Cancer. *J Clin Med* 2019; **8**: 1310 [PMID: 31454949 DOI: 10.3390/jcm8091310]

40 **Song Z**, Zou S, Zhou W, Huang Y, Shao L, Yuan J, Gou X, Jin W, Wang Z, Chen X, Ding X, Liu J, Yu C, Ku C, Liu C, Sun Z, Xu G, Wang Y, Zhang X, Wang D, Wang S, Xu W, Davis RC, Shi H. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 2020; **11**: 4294 [PMID: 32855423 DOI: 10.1038/s41467-020-18147-8]

41 **Yoshida H**, Shimazu T, Kiyuna T, Marugame A, Yamashita Y, Cosatto E, Taniguchi H, Sekine S, Ochiai A. Automated histological classification of whole-slide images of gastric biopsy specimens. *Gastric Cancer* 2018; **21**: 249-257 [PMID: 28577229 DOI: 10.1007/s10120-017-0731-8]

42 **Sharma H**, Zerbe N, Klempert I, Hellwich O, Hufnagl P. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Comput Med Imaging Graph* 2017; **61**: 2-13 [PMID: 28676295 DOI: 10.1016/j.compmedimag.2017.06.001]

43 **Li Q**, Qi L, Feng QX, Liu C, Sun SW, Zhang J, Yang G, Ge YQ, Zhang YD, Liu XS. Machine Learning-Based Computational Models Derived From Large-Scale Radiographic-Radiomic Images Can Help Predict Adverse Histopathological Status of Gastric Cancer. *Clin Transl Gastroenterol* 2019; **10**: e00079 [PMID: 31577560 DOI: 10.14309/ctg.0000000000000079]

44 **Li J**, Dong D, Fang M, Wang R, Tian J, Li H, Gao J. Dual-energy CT-based deep learning radiomics can improve lymph node metastasis risk prediction for gastric cancer. *Eur Radiol* 2020; **30**: 2324-2333 [PMID: 31953668 DOI: 10.1007/s00330-019-06621-x]

45 **Guimarães P**, Keller A, Fehlmann T, Lammert F, Casper M. Deep-learning based detection of gastric precancerous conditions. *Gut* 2020; **69**: 4-6 [PMID: 31375599 DOI: 10.1136/gutjnl-2019-319347]

46 **Zhang Y**, Li F, Yuan F, Zhang K, Huo L, Dong Z, Lang Y, Zhang Y, Wang M, Gao Z, Qin Z, Shen L. Diagnosing chronic atrophic gastritis by gastroscopy using artificial intelligence. *Dig Liver Dis* 2020; **52**: 566-572 [PMID: 32061504 DOI: 10.1016/j.dld.2019.12.146]

47 **Shichijo S**, Nomura S, Aoyama K, Nishikawa Y, Miura M, Shinagawa T, Takiyama H, Tanimoto T, Ishihara S, Matsuo K, Tada T. Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images. *EbioMedicine* 2017; **25**: 106-111 [PMID: 29056541 DOI: 10.1016/j.ebiom.2017.10.014]

48 **Zheng W**, Zhang X, Kim JJ, Zhu X, Ye G, Ye B, Wang J, Luo S, Li J, Yu T, Liu J, Hu W, Si J. High Accuracy of Convolutional Neural Network for Evaluation of Helicobacter pylori Infection Based on Endoscopic Images: Preliminary Experience. *Clin Transl Gastroenterol* 2019; **10**: e00109 [PMID: 31833862 DOI: 10.14309/ctg.0000000000000109]

49 **Bang CS**, Lee JJ, Baik GH. Artificial Intelligence for the Prediction of Helicobacter Pylori Infection in Endoscopic Images: Systematic Review and Meta-Analysis Of Diagnostic Test Accuracy. *J Med Internet Res* 2020; **22**: e21983 [PMID: 32936088 DOI: 10.2196/21983]

50 **Bianco F**, Arezzo A, Agresta F, Coco C, Faletti R, Krivocapic Z, Rotondano G, Santoro GA, Vettoretto N, De Franciscis S, Belli A, Romano GM; Italian Society of Colorectal Surgery. Practice parameters for early colon cancer management: Italian Society of Colorectal Surgery (Società Italiana di Chirurgia Colo-Rettale; SICCR) guidelines. *Tech Coloproctol* 2015; **19**: 577-585 [PMID: 26403233 DOI: 10.1007/s10151-015-1361-y]

51 **Dumoulin FL**, Hildenbrand R. Endoscopic resection techniques for colorectal neoplasia: Current developments. *World J Gastroenterol* 2019; **25**: 300-307 [PMID: 30686899 DOI: 10.3748/wjg.v25.i3.300]

52 **Murakami T**, Sakamoto N, Nagahara A. Endoscopic diagnosis of sessile serrated adenoma/polyp with and without dysplasia/carcinoma. *World J Gastroenterol* 2018; **24**: 3250-3259 [PMID: 30090005 DOI: 10.3748/wjg.v24.i29.3250]

53 **Kawasaki K**, Fujii M, Sugimoto S, Ishikawa K, Matano M, Ohta Y, Toshimitsu K, Takahashi S, Hosoe N, Sekine S, Kanai T, Sato T. Chromosome Engineering of Human Colon-Derived Organoids to Develop a Model of Traditional Serrated Adenoma. *Gastroenterology* 2020; **158**: 638-651.e8 [PMID: 31622618 DOI: 10.1053/j.gastro.2019.10.009]

54 **Itatani Y**, Kawada K, Sakai Y. Treatment of Elderly Patients with Colorectal Cancer. *Biomed Res Int* 2018; **2018**: 2176056 [PMID: 29713641 DOI: 10.1155/2018/2176056]

55 **Ladabaum U**. You Should Get Screened for Colon Cancer, Really. *JAMA Netw Open* 2019; **2**: e1910452 [PMID: 31469390 DOI: 10.1001/jamanetworkopen.2019.10452]

56 **US Preventive Services Task Force**, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW Jr, García FAR, Gillman MW, Harper DM, Kemper AR, Krist AH, Kurth AE, Landefeld CS, Mangione CM, Owens DK, Phillips WR, Phipps MG, Pignone MP, Siu AL. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 2016; **315**: 2564-2575 [PMID: 27304597 DOI: 10.1001/jama.2016.5989]

57 **Siegel RL**, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A. Colorectal cancer statistics, 2017. *CA Cancer J Clin* 2017; **67**: 177-193 [PMID: 28248415 DOI: 10.3322/caac.21395]

58 **Yamada M**, Saito Y, Imaoka H, Saiko M, Yamada S, Kondo H, Takamaru H, Sakamoto T, Sese J, Kuchiba A, Shibata T, Hamamoto R. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Sci Rep* 2019; **9**: 14465 [PMID: 31594962 DOI: 10.1038/s41598-019-50567-5]

59 **Yang YJ**, Cho BJ, Lee MJ, Kim JH, Lim H, Bang CS, Jeong HM, Hong JT, Baik GH. Automated Classification of Colorectal Neoplasms in White-Light Colonoscopy Images *via* Deep Learning. *J Clin Med* 2020; **9**: 1593 [PMID: 32456309 DOI: 10.3390/jcm9051593]

60 **Hashiguchi Y**, Muro K, Saito Y, Ito Y, Ajioka Y, Hamaguchi T, Hasegawa K, Hotta K, Ishida H, Ishiguro M, Ishihara S, Kanemitsu Y, Kinugasa Y, Murofushi K, Nakajima TE, Oka S, Tanaka T, Taniguchi H, Tsuji A, Uehara K, Ueno H, Yamanaka T, Yamazaki K, Yoshida M, Yoshino T, Itabashi M, Sakamaki K, Sano K, Shimada Y, Tanaka S, Uetake H, Yamaguchi S, Yamaguchi N, Kobayashi H, Matsuda K, Kotake K, Sugihara K; Japanese Society for Cancer of the Colon and Rectum. Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2019 for the treatment of colorectal cancer. *Int J Clin Oncol* 2020; **25**: 1-42 [PMID: 31203527 DOI: 10.1007/s10147-019-01485-z]

61 **Ito N**, Kawahira H, Nakashima H, Uesato M, Miyauchi H, Matsubara H. Endoscopic Diagnostic Support System for cT1b Colorectal Cancer Using Deep Learning. *Oncology* 2019; **96**: 44-50 [PMID: 30130758 DOI: 10.1159/000491636]

62 **Hoerter N**, Gross SA, Liang PS. Artificial Intelligence and Polyp Detection. *Curr Treat Options Gastroenterol* 2020; epub ahead of print [PMID: 31960282 DOI: 10.1007/s11938-020-00274-2]

63 **Vinsard DG**, Mori Y, Misawa M, Kudo SE, Rastogi A, Bagci U, Rex DK, Wallace MB. Quality assurance of computer-aided detection and diagnosis in colonoscopy. *Gastrointest Endosc* 2019; **90**: 55-63 [PMID: 30926431 DOI: 10.1016/j.gie.2019.03.019]

64 **Zhou G**, Xiao X, Tu M, Liu P, Yang D, Liu X, Zhang R, Li L, Lei S, Wang H, Song Y, Wang P. Computer aided detection for laterally spreading tumors and sessile serrated adenomas during colonoscopy. *PloS One* 2020; **15**: e0231880 [PMID: 32315365 DOI: 10.1371/journal.pone.0231880]

65 **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology* 2020; **159**: 512-520.e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

66 **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective andomized controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

67 **Stidham RW**, Liu W, Bishu S, Rice MD, Higgins PDR, Zhu J, Nallamothu BK, Waljee AK.

Performance of a Deep Learning Model *vs* Human Reviewers in Grading Endoscopic Disease Severity of Patients With Ulcerative Colitis. *JAMA Netw Open* 2019; **2**: e193963 [PMID: 31099869 DOI: 10.1001/jamanetworkopen.2019.3963]

68 **Waljee AK**, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, Stidham RW, Zhu J, Higgins PDR. Development and Validation of Machine Learning Models in Prediction of Remission in Patients With Moderate to Severe Crohn Disease. *JAMA Netw Open* 2019; **2**: e193721 [PMID: 31074823 DOI: 10.1001/jamanetworkopen.2019.3721]

69 **Thakur N**, Yoon H, Chong Y. Current Trends of Artificial Intelligence for Colorectal Cancer Pathology Image Analysis: A Systematic Review. *Cancers (Basel)* 2020; **12**: 1884 [PMID: 32668721 DOI: 10.3390/cancers12071884]

70 **Sena P**, Fioresi R, Faglioni F, Losi L, Faglioni G, Roncucci L. Deep learning techniques for detecting preneoplastic and neoplastic lesions in human colorectal histological images. *Oncol Lett* 2019; **18**: 6101-6107 [PMID: 31788084 DOI: 10.3892/ol.2019.10928]

71 **Takamatsu M**, Yamamoto N, Kawachi H, Chino A, Saito S, Ueno M, Ishikawa Y, Takazawa Y, Takeuchi K. Prediction of early colorectal cancer metastasis by machine learning using digital slide images. *Comput Methods Programs Biomed* 2019; **178**: 155-161 [PMID: 31416544 DOI: 10.1016/j.cmpb.2019.06.022]

72 **Guo L**, Xiao X, Wu C, Zeng X, Zhang Y, Du J, Bai S, Xie J, Zhang Z, Li Y, Wang X, Cheung O, Sharma M, Liu J, Hu B. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest Endosc* 2020; **91**: 41-51 [PMID: 31445040 DOI: 10.1016/j.gie.2019.08.018]

73 **Cai SL**, Li B, Tan WM, Niu XJ, Yu HH, Yao LQ, Zhou PH, Yan B, Zhong YS. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). *Gastrointest Endosc* 2019; **90**: 745-753.e2 [PMID: 31302091 DOI: 10.1016/j.gie.2019.06.044]

74 **Zhao YY**, Xue DX, Wang YL, Zhang R, Sun B, Cai YP, Feng H, Cai Y, Xu JM. Computer-assisted diagnosis of early esophageal squamous cell carcinoma using narrow-band imaging magnifying endoscopy. *Endoscopy* 2019; **51**: 333-341 [PMID: 30469155 DOI: 10.1055/a-0756-8754]

75 **Sakai Y**, Takemoto S, Hori K, Nishimura M, Ikematsu H, Yano T, Yokota H. Automatic detection of early gastric cancer in endoscopic images using a transferring convolutional neural network. *Annu Int Conf IEEE Eng Med Biol Soc* 2018; **2018**: 4138-4141 [PMID: 30441266 DOI: 10.1109/EMBC.2018.8513274]

76 **Zhang L**, Zhang Y, Wang L, Wang J, Liu Y. Diagnosis of gastric lesions through a deep convolutional neural network. *Dig Endosc* 2021; **33**: 788-796 [PMID: 32961597 DOI: 10.1111/den.13844]

77 **Li L**, Chen Y, Shen Z, Zhang X, Sang J, Ding Y, Yang X, Li J, Chen M, Jin C, Chen C, Yu C. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. *Gastric Cancer* 2020; **23**: 126-132 [PMID: 31332619 DOI: 10.1007/s10120-019-00992-2]

78 **Wu L**, Zhou W, Wan X, Zhang J, Shen L, Hu S, Ding Q, Mu G, Yin A, Huang X, Liu J, Jiang X, Wang Z, Deng Y, Liu M, Lin R, Ling T, Li P, Wu Q, Jin P, Chen J, Yu H. A deep neural network improves endoscopic detection of early gastric cancer without blind spots. *Endoscopy* 2019; **51**: 522-531 [PMID: 30861533 DOI: 10.1055/a-0855-3532]

79 **Ikenoyama Y**, Hirasawa T, Ishioka M, Namikawa K, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Takeuchi Y, Shichijo S, Katayama N, Fujisaki J, Tada T. Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. *Dig Endosc* 2021; **33**: 141-150 [PMID: 32282110 DOI: 10.1111/den.13688]

80 **Cho BJ**, Bang CS, Park SW, Yang YJ, Seo SI, Lim H, Shin WG, Hong JT, Yoo YT, Hong SH, Choi JH, Lee JJ, Baik GH. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019; **51**: 1121-1129 [PMID: 31443108 DOI: 10.1055/a-0981-6133]

81 **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078.e8 [PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037]

82 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind andomized study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

83 **Park CH**, Yang DH, Kim JW, Kim JH, Min YW, Lee SH, Bae JH, Chung H, Choi KD, Park JC, Lee H, Kwak MS, Kim B, Lee HJ, Lee HS, Choi M, Park DA, Lee JY, Byeon JS, Park CG, Cho JY, Lee ST, Chun HJ. Clinical Practice Guideline for Endoscopic Resection of Early Gastrointestinal Cancer. *Clin Endosc* 2020; **53**: 142-166 [PMID: 32252507 DOI: 10.5946/ce.2020.032]

84 **Lee JK**, Jensen CD, Levin TR, Zauber AG, Doubeni CA, Zhao WK, Corley DA. Accurate Identification of Colonoscopy Quality and Polyp Findings Using Natural Language Processing. *J Clin Gastroenterol* 2019; **53**: e25-e30 [PMID: 28906424 DOI: 10.1097/MCG.0000000000000929]

85 **Rutter MD**, Rees CJ. Quality in gastrointestinal endoscopy. *Endoscopy* 2014; **46**: 526-528 [PMID: 24788539 DOI: 10.1055/s-0034-1365738]

86 **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective

randomized controlled study (with videos). *Gastrointest Endosc* 2020; **91**: 415-424.e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

87 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. 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 [PMID: 29066576 DOI: 10.1136/gutjnl-2017-314547]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com

ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 October 28; 2(5): 198-210

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 5 October 28, 2021

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 5 October 28, 2021

Contents

ABOUT COVER

Associate Editor of *Artificial Intelligence in Gastrointestinal Endoscopy*, Chong-Chi Chiu, MD, Professor, Surgeon, Surgical Oncologist, Department of General Surgery, School of Medicine, College of Medicine, E-Da Cancer Hospital, I-Shou University, Kaohsiung 82445, Taiwan. chiuchongchi@gmail.com

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

PUBLICATION DATE
October 28, 2021

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 October 28; 2(5): 198-210

DOI: 10.37126/aige.v2.i5.198

ISSN 2689-7164 (online)

MINIREVIEWS

Artificial intelligence and early esophageal cancer

Ning Li, Shi-Zhu Jin

ORCID number: Ning Li 0000-0002-0913-0624; Shi-Zhu Jin 0000-0003-3613-0926.

Author contributions: Li N wrote the paper and prepared the figures and tables; Jin SZ revised the paper.

Supported by Heilongjiang Province Education Science "13th Five-Year Plan" 2020 Key Project, No. GJB1320190.

Conflict-of-interest statement: The authors declare no conflicts of interest related to this article.

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Ning Li, Shi-Zhu Jin, Department of Gastroenterology and Hepatology, The Second Affiliated Hospital of Harbin Medical University, Harbin 150086, Heilongjiang Province, China

Corresponding author: Shi-Zhu Jin, MD, Chief Physician, Professor, Department of Gastroenterology and Hepatology, The Second Affiliated Hospital of Harbin Medical University, No. 246 Xuefu Road, Nangang District, Harbin 150086, Heilongjiang Province, China. drshizhujin@hrbmu.edu.cn

Abstract

The development of esophageal cancer (EC) from early to advanced stage results in a high mortality rate and poor prognosis. Advanced EC not only poses a serious threat to the life and health of patients but also places a heavy economic burden on their families and society. Endoscopy is of great value for the diagnosis of EC, especially in the screening of Barrett's esophagus and early EC. However, at present, endoscopy has a low diagnostic rate for early tumors. In recent years, artificial intelligence (AI) has made remarkable progress in the diagnosis of digestive system tumors, providing a new model for clinicians to diagnose and treat these tumors. In this review, we aim to provide a comprehensive overview of how AI can help doctors diagnose early EC and precancerous lesions and make clinical decisions based on the predicted results. We analyze and summarize the recent research on AI and early EC. We find that based on deep learning (DL) and convolutional neural network methods, the current computer-aided diagnosis system has gradually developed from *in vitro* image analysis to real-time detection and diagnosis. Based on powerful computing and DL capabilities, the diagnostic accuracy of AI is close to or better than that of endoscopy specialists. We also analyze the shortcomings in the current AI research and corresponding improvement strategies. We believe that the application of AI-assisted endoscopy in the diagnosis of early EC and precancerous lesions will become possible after the further advancement of AI-related research.

Key Words: Artificial intelligence; Computer-aided diagnosis; Deep learning; Convolutional neural network; Barrett's esophagus; Early esophageal cancer

Core Tip: The early diagnosis and early treatment of esophageal cancer (EC) have always been a hot spot in clinical medicine research and are of great importance to the

Country/Territory of origin: China

Peer-review report's scientific quality classification
Grade A (Excellent): A
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: July 28, 2021
Peer-review started: July 28, 2021
First decision: September 12, 2021
Revised: September 23, 2021
Accepted: October 27, 2021
Article in press: October 27, 2021
Published online: October 28, 2021

P-Reviewer: Shafqat S, Viswanath YK
S-Editor: Liu M
L-Editor: Wang TQ
P-Editor: Liu M

prognosis of patients. With continuous improvements in computer technology and the arrival of the era of big data, the artificial intelligence (AI)-assisted endoscopic diagnosis of EC has also flourished. This review mainly introduces the research progress of AI-assisted endoscopy in the diagnosis of Barrett's esophagus and early EC.

Citation: Li N, Jin SZ. Artificial intelligence and early esophageal cancer. *Artif Intell Gastrointest Endosc* 2021; 2(5): 198-210
URL: https://www.wjgnet.com/2689-7164/full/v2/i5/198.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i5.198

INTRODUCTION

Barrett's esophagus (BE) is a premalignant condition characterized by the replacement of columnar epithelium with esophageal squamous epithelium. Esophageal cancer (EC) is the seventh most common cancer and the sixth leading cause of cancer-related mortality worldwide[1]. EC mainly consists of two histological types: Esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC). ESCC is the main pathological type in Asian countries, and the 5-year survival rate is less than 20% [2]. EAC is more common in Western countries, and its incidence has been on the rise globally in recent years[3]. The development of EC from early to advanced stage is accompanied by a high mortality rate and poor prognosis. Early detection and diagnosis greatly impact the prognosis of EC. The need for more efficient detection methods for early EC has led to in-depth research in the field of artificial intelligence (AI). The purpose of this review is to summarize the diagnostic value of AI for BE and early EC, which is conducive to the early treatment of patients and the reduction in mortality. In this review, we will discuss the following: (1) The utility of AI techniques in the endoscopic detection of BE; (2) the utility of AI techniques in the endoscopic detection of early EC; and (3) problems and prospects of AI-assisted endoscopic diagnosis.

ARTIFICIAL INTELLIGENCE

AI refers to the abilities of computers to imitate the cognitive function of the human mind and conduct autonomous learning. In recent years, AI has made great progress in various fields of medicine, such as radiological oncology, diabetic retinopathy, and skin cancer[4-6]. Machine learning (ML) for AI can be roughly divided into traditional learning and deep learning (DL). Traditional learning methods require artificial design, which is time-consuming and laborious. DL methods can independently extract and learn image features and extract more complex and abstract advanced features layer by layer through a multilayer system, which allows them to be truly mature and be applied in clinical practice[7]. Convolutional neural networks (CNNs) are a kind of DL method commonly used in AI-assisted image recognition. These networks contain multilayer perceptrons and imitate human brain neural circuits to carry out high generalization, abstraction, and synthesis to process information. The DL method is an end-to-end learning method without the need to design specific image features[8,9]. With the rapid development of information technology, DL has received increasing attention in the medical field. The computer-aided diagnosis (CAD) of gastrointestinal (GI) diseases has become a hot research topic. CAD is an advanced technology used to preprocess endoscopic images, extract image features, process data, and obtain diagnostic results with the help of computer algorithms and graphics processing technology[10] (Figure 1).

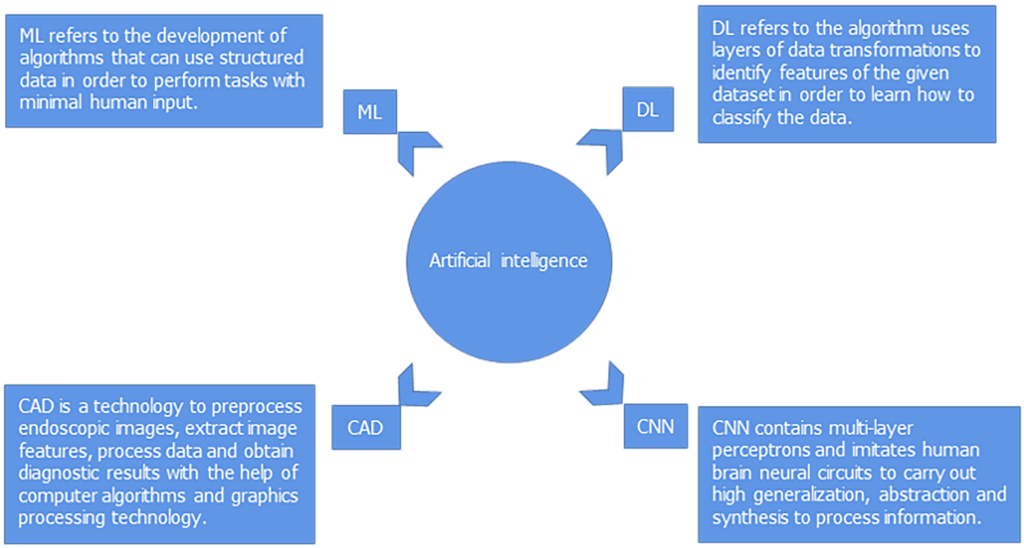

Figure 1 Diagram representation of artificial intelligence domains. ML: Machine learning; DL: Deep learning; CAD: Computer-aided diagnosis; CNN: Convolutional neural network.

APPLICATION OF ARTIFICIAL INTELLIGENCE IN BARRETT'S ESOPHAGUS

BE is a result of chronic inflammation of the esophagus and is a risk factor for the development of EAC[11]. GI societies recommend regular endoscopy for BE patients to detect dysplasia or carcinoma early[12,13]. Endoscopic surveillance currently follows the Seattle protocol: Patients with BE are required to undergo a systematic four-quadrant biopsy, in which the entire BE area is sampled at intervals of 1-2 cm using a "turn and suction technique"[14]. However, this method can be invasive, costly, time-consuming, and difficult for patients to follow[15]. Due to poor patient compliance with the Seattle protocol, the American Society of Gastrointestinal Endoscopy established a performance threshold for optical diagnosis. Random biopsies can be replaced if targeted biopsies assisted by any imaging technique have a per-patient sensitivity of 90%, negative predictive value (NPV) of 98%, and specificity of 80%[16]. However, these requirements can only be achieved by experts.

In addition, early neoplastic lesions and dysplasia are subtle, showing focal distribution, and are difficult to detect endoscopically[17]. Cases of BE progression to early tumors are rare, and endoscopic surveillance is generally carried out in community hospitals; therefore, general endoscopists may not be familiar with these lesions, and this lack of familiarity is an important reason for missed diagnosis[18,19]. In recent years, to improve the diagnosis of BE, many new endoscopic techniques have been developed, such as magnification endoscopy (ME), chromoendoscopy, confocal laser endomicroscopy, and volumetric laser endomicroscopy, most of which are expensive and take a long time for endoscopists to learn[20,21]. Differences in endoscopists' interpretations of the images can also lead to differences in diagnosis[22]. Therefore, there is an urgent need for a practical tool to improve the accuracy of endoscopists in the clinic. Moreover, the endoscopist's diagnosis may be influenced by the time of the endoscopy, psychological state, time pressure, and cumbersome procedures. However, AI has a short learning time and, unlike endoscopists, does not suffer from fatigue easily; therefore, it has good application prospects (Table 1).

Computer-aided diagnosis using white light imaging/narrow band imaging

van der Sommen *et al*[23] collected 100 images from 44 BE patients and created an ML algorithm called support vector machine (SVM), which employed specific texture and color filters to detect early neoplasia in BE. The sensitivity and specificity of the system were both 83% for the per-image analysis and 86% and 87% for the per-patient analysis, respectively.

Struyvenberg *et al*[24] developed a CAD system based on a CNN model that was first trained with 494364 images and then further trained with 690 BE neoplasia and 557 nondysplastic BE (NDBE) white light imaging (WLI) images. Next, 112 BE neoplasia and 71 NDBE narrow band imaging (NBI) zoom images were used for

Table 1 Application of artificial intelligence in endoscopic detection of Barrett's esophagus

Ref.	Target disease	Endoscopic modality	AI technology	Database	Outcomes
van der Sommen *et al*[23], 2016	Early neoplasia in BE	WLI	SVM	100 images	Per-image sensitivity 83%/specificity 83%; Per-patient sensitivity 86%/specificity: 87%
Struyvenberg *et al* [24], 2021	BE	WLI/NBI	CNN	Train 494364 images/1247 images; test 183 images/157 videos	Images: Accuracy 84%/sensitivity 88%/specificity 78%; Videos: Accuracy 83%/sensitivity 85%/specificity 83%
de Groof *et al*[25], 2020	Early neoplasia in BE	WLI	ResNet-UNet	Train 1544 images; test 160 images	Dataset 4: Accuracy 89%/sensitivity 90%/specificity 88%; Dataset 5: Accuracy 88%/sensitivity 93%/specificity 83%
de Groof *et al*[26], 2020	Barrett's neoplasia	WLI	ResNet-UNet	Train 1544 images; test 20 patients	Accuracy 90%/sensitivity 91%/specificity 89%
Hong *et al*[27], 2017	BE	Endomicroscopy	CNN	Train 236 images; test 26 images	Accuracy 80.77%
Hashimoto *et al* [28], 2020	Early neoplasia in BE	WLI/NBI	CNN	Train 1832 images; test 458 images	Accuracy 95.4%/sensitivity 96.4%/ specificity 94.2%
de Groof *et al*[29], 2019	Barrett's neoplasia	WLI	SVM	60 images	Accuracy 92%/sensitivity 95%/specificity 85%

BE: Barrett's esophagus; WLI: White light imaging; SVM: Support vector machine; NBI: Narrow band imaging; CNN: Convolutional neural network.

training and validation. Finally, 59 BE neoplasia and 98 NDBE NBI zoom videos were used for training and validation. Fourfold cross-validation was used to evaluate the detection performance of the CAD system. The results showed that the accuracy, sensitivity, and specificity of the NBI zoom image based CAD system were 84%, 88%, and 78%, respectively. Accuracy, sensitivity, and specificity of the NBI zoom videos were 83%, 85%, and 83%, respectively.

de Groof *et al*[25,26] developed a CAD system based on ResNet/U-Net model to help endoscopists detect early BE neoplasia. The system was trained with 1544 endoscopic images of BE neoplasia and NDBE and then validated on 160 images. In an *in vitro* study, the accuracy, sensitivity, and specificity of the CAD system for detecting early BE neoplasia were 89%, 90%, and 88% in dataset 4, and 88%, 93% and 83% in dataset 5, respectively. Compared with 53 nonspecialist endoscopists, the CAD system outperformed them in terms of accuracy and sensitivity. In an *in vivo* evaluation of the CAD system, endoscopic examinations were performed on ten patients with NDBE and ten patients with BE neoplasia. The images obtained by WLI were analyzed immediately by the CAD system and used to provide feedback to the endoscopist. The accuracy, sensitivity, and specificity of the CAD system were 90%, 91% and 89%, respectively. Therefore, the CAD system has a high accuracy for tumor detection and low false positive rate; thus, the CAD system can be tested in larger and multicenter trials.

Hong *et al*[27] constructed a CNN-based CAD system to distinguish intestinal metaplasia (IM), gastric metaplasia (GM), and BE neoplasia. The researchers obtained 236 endoscopic images of BE from the 2016 International Symposium on Biomedical Imaging using 155 IM, 26 GM, and 55 BE neoplasia samples as a training set. Because the number of images in the training set was insufficient, the researchers implemented image distortion to achieve data enhancement and increase the sample size of the data. Then, 26 images, including 17 IM, 4 GM, and 5 BE neoplasia images, were used as the verification set. The results showed that the accuracy of the CAD system for the classification of IM, GM, and BE neoplasia was 80.77%. Although the number of images was small, this study suggested that the CNN-structured CAD system can be applied to the classification of esophageal lesions.

Real-time recognition by computer-aided diagnosis

Hashimoto *et al*[28] collected 916 images from 70 patients with early neoplastic BE and 916 control images from 30 normal BE patients and then trained a CNN algorithm on ImageNet. The researchers analyzed 458 images using the CNN algorithm. The accuracy, sensitivity, and specificity of the system for detecting early neoplastic BE were 95.4%, 96.4%, and 94.2%, respectively.

de Groof *et al*[29] designed an ML algorithm called SVM based on WLI images from 40 BE neoplasias patients and 20 NDBEs patients. All of the images were delineated by endoscopic experts, with overlapping areas of at least four delineations marked as "sweet spots" and areas with at least one delineation marked as "soft spots". The CAD system was trained (color and texture features) and then evaluated for its performance using leave-one-out cross-validation. The accuracy, sensitivity, and specificity of the CAD system were 92%, 95%, and 85%, and the localization and labeling of soft spots were 100% and 90%, respectively. Therefore, this CAD system can detect and locate early BE neoplasia with a high accuracy on WLI images, which lays a foundation for the real-time automatic recognition of BE neoplasia in the future.

APPLICATION OF ARTIFICIAL INTELLIGENCE IN ESOPHAGEAL CANCER

EC is usually diagnosed at an advanced stage, and the main treatment is esophagectomy. Surgical treatment is a highly invasive treatment with relatively high mortality and recurrence rates and poor patient prognoses. However, if EC is detected at an early stage, the prognosis can be improved by endoscopic resection[30,31]. Therefore, the early diagnosis of EC is essential for favorable treatment. Some studies have applied certain dyes to the esophageal mucosa that can more clearly reveal the surface vasculature and neoplasia. The most commonly used dyes are acetic acid, iodine, indigo carmine, and methylene blue. However, there are limitations in terms of the cost and complexity of their application[32,33]. NBI provides a better view of intrapapillary capillary loops (IPCLs) and is used to detect superficial ESCC. However, inexperienced endoscopists are still prone to missed diagnoses[34-36]. Therefore, AI, which can outperform humans in image recognition, is expected to be used in the field of EC diagnosis (Table 2).

Detection of lesions

Ebigbo *et al*[37] created a CAD system based on CNN. In the Augsburg database, the sensitivity and specificity of the CAD system for the diagnosis of EAC in WLI images were 97% and 88%, respectively, and the sensitivity and specificity in NBI images were 94% and 80%, respectively. In the MICCAI database, the sensitivity and specificity of the CAD system for the diagnosis of EAC in WLI images were 92% and 100%, respectively. Then, Ebigbo *et al*[38] developed an artificial neural network of encoder-decoders with 101 layers of ResNet and trained the CAD system using 129 endoscopic images from the Augsburg database. The researchers evaluated 62 images using the CAD system, including 36 images of early EAC and 26 images of BE. Although the number of patients evaluated was low, real-time monitoring of EAC demonstrated good results. The sensitivity and specificity of the system were 83.7% and 100%, respectively, and the overall accuracy was 89.9%.

Horie *et al*[39] developed a CNN system using DL to correctly detect EC based on 8428 images from 384 patients with EC. The researchers used the CNN system to analyze 1118 images (47 patients with EC and 50 patients without EC). The system takes 27 s and has a sensitivity of 98%; it can detect EC lesions less than 10 mm in size. The NPV was 95%, but the positive predictive value was only 40%. This may be due to the small number of DL training sets and few images from patients with esophageal inflammation. In addition, the system can distinguish between superficial EC and advanced EC with a 98% accuracy. These results indicate that the CNN system constructed by researchers can accurately analyze a large number of endoscopic images in a short period of time, which is conducive to the early diagnosis of EC.

Cai *et al*[40] developed a CAD system using a deep neural network based on 2428 endoscopic images (746 patients) with the aim of identifying early ESCC from WLI images. Among these images, there were 1332 ESCC images and 1096 normal tissue images. The researchers evaluated the CAD system using 187 images (52 patients), and 16 endoscopic physicians reviewed the images. The results showed that the accuracy, sensitivity, and specificity of the CAD system for the early diagnosis of ESCC were 91.4%, 97.8% and 85.4%, respectively. With the help of the CAD system, the diagnostic accuracy and sensitivity of endoscopists with different seniority levels were improved, especially for those with less seniority. This result indicates that AI-assisted digestive endoscopy can reduce the rate of missed diagnosis and improve the diagnostic level of endoscopists with different experiences in early EC.

Ohmori *et al*[41] developed a CAD system based on CNN to evaluate the diagnosis of ESCC under ME and non-ME. The researchers used 7844 ME and 9591 non-ME images from ESCC and 3435 ME and 1692 non-ME images from noncancerous or

Table 2 Application of artificial intelligence in endoscopic detection of early esophageal cancer

Ref.	Target disease	Endoscopic modality	AI technology	Database	Outcomes
Ebigbo *et al* [37], 2019	EAC	WLI/NBI	CNN	248 images	Augsburg database: Sensitivity 97%/specificity 88% (WLI); Sensitivity 94%/specificity 80% (NBI); MICCAI database: Sensitivity 92%/specificity 100%
Ebigbo *et al* [38], 2020	EAC	WLI	CNN	Train 129 images; test 62 images	Accuracy 89.9%/sensitivity 83.7%/specificity 100%
Horie *et al*[39], 2019	EC	WLI/NBI	CNN	Train 8428 images; test 1118 images	Accuracy 98%/sensitivity 98%
Cai *et al*[40], 2019	ESCC	WLI	DNN	Train 2428 images; test 187 images	Accuracy 91.4%/sensitivity 97.8%/specificity 85.4%
Ohmori *et al* [41], 2020	ESCC	WLI/NBI/BLI	CNN	Train 22562 images; test 727 images	Non-ME: Accuracy 81.0%/sensitivity 90%/specificity 76% (WLI); Accuracy 77%/sensitivity 100%/specificity 63% (NBI/BLI); ME: Accuracy 77%/sensitivity 98%/specificity 56%
Liu *et al*[42], 2020	EC	WLI	CNN	Train 1017 images; test 255 images	Accuracy 85.83%/sensitivity 94.23%/specificity 94.67%
Kumagai *et al* [43], 2019	ESCC	ECS	CNN	Train 4715 images; test 1520 images	Accuracy 90.9%/sensitivity 92.6%/specificity 89.3%
Guo *et al*[44], 2020	ESCC	NBI	CNN	Train 6473 images; test 6671 images and 80 videos	Images: Sensitivity 98.04%/specificity 95.03%; videos: Non-ME sensitivity 60.8% (per frame)/100% (per lesion); ME sensitivity 96.1% (per frame)/100% (per lesion)
Tokai *et al*[46], 2020	ESCC	WLI/NBI	CNN	Train 1751 images; test 291 images	Accuracy 80.9%/sensitivity 84.1%/specificity 73.3%
Nakagawa *et al* [47], 2019	ESCC	WLI/NBI	CNN	Train 14338 images; test 914 images	Accuracy 91%/sensitivity 90.1%/specificity 95.8%
Zhao *et al*[48], 2019	ESCC	NBI	Double-labeling FCN	1350 images	Lesion level: Accuracy 89.2%; pixel level: Accuracy 93%
Everson *et al* [49]	ESCC	NBI	CNN	7046 images	Accuracy 93.7%/sensitivity 89.3%/specificity 98%
Uema *et al*[50], 2021	ESCC	NBI	CNN	Train 1777 images; test 747 images	Accuracy 84.2%
Fukuda *et al* [51], 2020	ESCC	NBI/BLI	CNN	Train 28333 images; test 144 patients	Accuracy 63%/sensitivity 91%/specificities 51% (detection); accuracy 88%/sensitivity 86%/specificities 89% (characterization)
Shimamoto *et al* [52], 2020	ESCC	WLI/NBI/BLI	CNN	Train 23977 images; test 102 videos	Non-ME: Accuracy 87%/sensitivity 50%/specificity 99%; ME: Accuracy 89%/sensitivity 71%/specificity 95%
Waki *et al*[53], 2021	ESCC	WLI/NBI/BLI	CNN	Train 18797 images; test 100 videos	Sensitivity 85.7%/specificity 40%

EAC: Esophageal adenocarcinoma; WLI: White light imaging; NBI: Narrow band imaging; CNN: Convolutional neural network; EC: Esophageal cancer; ESCC: Esophageal squamous cell carcinoma; DNN: Deep neural network; BLI: Blue laser imaging; ME: Magnification endoscopy; ECS: Endocytoscopic system; FCN: Fully convolutional network.

normal esophagi as a training set. Then, 255 non-ME WLI images, 268 non-ME-NBI/blue laser imaging (BLI) images, and 204 ME-NBI/BLI images of ESCC were used as a validation set. The accuracy, sensitivity, and specificity of the CAD system were 81%, 90%, and 76%, respectively, in non-ME WLI images. In the non-ME diagnosis of NBI/BLI images, the accuracy, sensitivity, and specificity of the CAD system were 77%, 100%, and 63%, respectively. In the diagnosis of ME, the CAD system had an accuracy of 77%, sensitivity of 98%, and specificity of 56%. In conclusion, the diagnosis of ESCC with the CAD system was not significantly different from that of experienced endoscopists.

Liu *et al*[42] developed a CNN model using the DL approach to distinguish among normal esophagi, precancerous lesions, and EC. The model consists of two subnetworks: The O-stream and the P-stream. In the application process, the O-stream is used to input the original images to extract color changes and overall features, and the P-stream is used to input the preprocessing images to lift texture changes and detail features. In total, 1017 images (normal esophagi, precancerous lesions, and EC)

were used as the training set, and 255 images (normal esophagi, precancerous lesions, and EC) were used as the validation set. The results showed that the accuracy, sensitivity, and specificity of the CNN model were 85.83%, 94.23% and 94.67%, respectively, which shows good prospects in the diagnosis of esophageal lesions.

Kumagai *et al*[43] constructed an AI model based on CNN with GoogLeNet to judge benign and malignant endocytoscopic system (ECS) images with different degrees of magnification. The AI system was trained using 4715 esophageal ECS images (1141 malignant and 3574 nonmalignant) and validated using 1520 images (27 ESCCs and 28 nonmalignant lesions). The results showed that the sensitivity of the AI system was 92.6% for the diagnosis of ESCC, the specificity was 89.3% for the diagnosis of nonmalignant lesions, and the accuracy was 90.9% for the overall diagnosis. Early EC under endoscopy usually presents as slight swelling, depression, or a color change in the mucosa, which is difficult to diagnose, especially for less experienced endoscopists. The above research results indicate that AI has good auxiliary value for the endoscopic diagnosis of early EC and its precancerous lesions and plays an important role in guiding learning for the applications of some new standards and technologies.

Scope of lesions

Guo *et al*[44] developed a CAD system based on CNN for the real-time detection of precancerous lesions and ESCC. A total of 6473 NBI images were used to train the CAD system, and endoscopic static images and dynamic videos were used to validate the CAD system. Each input endoscopic image generates an AI probabilistic heat map, where yellow indicates highly suspected cancerous lesions and blue indicates noncancerous lesions. When the CAD system is used to detect canceration, the identified tumor area is covered with color. The CAD system was used to diagnose 1480 malignant NBI images and 5191 nonmalignant NBI images with a sensitivity and specificity of 98.04% and 95.03%, respectively. In 27 non-ME and 20 ME videos of precancerous lesions and early ESCC, the sensitivities per frame were 60.8% and 96.1%, respectively, and the sensitivities per lesion were 100% and 100%. In 33 normal esophageal videos, the specificities were 99.9% per frame and 90.9% per case. The AI model can mark the location and range of lesions according to the input images, and the range is roughly the same as that marked by endoscopists. This finding indicates the feasibility and great potential of AI in the identification of a range of precancerous or early EC lesions.

Depth of lesions

The depth of EC invasion is a key factor affecting treatment decisions. In principle, endoscopic resection can be performed for intraepithelial esophageal lesions confined to the lamina propria or muscularis mucosa and/or lesions with a submucosal infiltration depth less than 200 μm. Surgical resection and chemoradiotherapy are required for lesions larger than 200 μm. Therefore, accurate determination of the depth of infiltration can avoid the impact of overtreatment on patient quality of life[45].

Tokai *et al*[46] collected 1751 ESCC images to design an AI diagnostic system using CNN techniques. The system used DI technology to evaluate the infiltration depth of ESCC. The researchers used the AI system to evaluate 55 patients (291 images) and compared them with the evaluations of 13 endoscopists. It was found that the detection rate of the AI system for ESCC was 95.5%, taking 10 s. In the images with ESCC detected, the accuracy, sensitivity, and specificity of the assessment of infiltration depth were 80.9%, 84.1%, and 73.3%, respectively, taking 6 s. Moreover, the AI system was more accurate than 12 of the 13 endoscopists. This result indicates that the AI system has great potential in detecting the infiltration depth of ESCC.

Nakagawa *et al*[47] developed a CNN-based AI system to assess the infiltration depth of ESCC. The researchers trained the AI system with images from 804 EC patients (8660 non-ME images and 5678 ME images) and then validated the system with images from 155 patients (405 non-ME images and 509 ME images). The accuracy, sensitivity, and specificity of the system were 91%, 90.1%, and 95.8%, respectively. When 16 endoscopists evaluated the same images, the accuracy, sensitivity, and specificity were 89.6%, 89.8%, and 88.3%, respectively. These results suggest that the AI system performs well in assessing the depth of ESCC infiltration, even better than endoscopists.

IPCLs are the hallmark of ESCC, and their morphologic changes correlate with the depth of tumor invasion. Zhao *et al*[48] used the ME-NBI technique to evaluate patients' esophageal conditions and established a CAD system for the automatic classification of IPCLs based on endoscopic diagnosis and histological analysis. This system uses a double-labeling fully convolutional network to evaluate 1350 images with 1383 lesions and compare them with the evaluations of endoscopists. The results showed

that the diagnostic accuracy of the system was 89.2% at the lesion level and 93% at the pixel level, which were higher than those of endoscopists.

Everson *et al*[49] developed an AI system to detect the presence and stage of early ESCC lesions. A total of 7046 ME-NBI images from 17 patients were used to train the CNN. Among these patients, ten had early ESCC, and seven had a normal esophagus. All of the imaging areas were supported by histological results. Studies have shown that the accuracy of this CNN system for distinguishing normal and abnormal IPCL patterns is 93.7%, and the sensitivity and specificity for distinguishing abnormal IPCL patterns are 89.3% and 98%, respectively. Therefore, the CNN system can relatively accurately distinguish normal and abnormal IPCL patterns and may provide guidance for decision-making regarding the clinical treatment of ESCC.

Uema *et al*[50] constructed a CNN (ResNeXt-101) model to classify ESCC microvessels. The study used 1777 ESCC images under ME-NBI as a training set and 747 ESCC images under ME-NBI as a validation set (validated by the CAD system and 8 endoscopists). The results showed that the accuracy of the CAD system for microvascular classification was 84.2%, which was higher than the average accuracy achieved by endoscopists. Therefore, this CAD system has good application potential for ESCC microvascular classification.

Dynamic images

Fukuda *et al*[51] developed a CNN-based CAD system to diagnose ESCC. The researchers used 23746 ESCC images (1544 patients) and 4587 noncancerous images (458 patients) as a training set. Video image clips from 144 patients were used as a validation set, and then 13 endoscopic specialists used the same videos for diagnoses. The accuracy, sensitivity, and specificity of the CAD system in identifying suspicious lesions were 63%, 91%, and 51%, respectively. The accuracy, sensitivity, and specificity in differentiating cancerous from noncancerous lesions were 88%, 86% and 89%, respectively. In previous studies, the diagnosis of ESCC by CAD systems was mainly based on static images, with few video images. Because video images are affected by many factors, such as distance, angle, breathing movement, and esophageal motility, using a CAD system to analyze video images is more challenging. Fukuda *et al*[51] demonstrated that compared with endoscopic experts, CAD systems are more sensitive to ESCC detection and have a significantly higher accuracy and specificity in differentiating cancer from noncancer, which will provide valuable clinical support for endoscopists in their diagnoses.

Using 23977 ESCC images (6857 WLI images and 17120 NBI/BL images) as a training set, Shimamoto *et al*[52] developed a CNN-based AI system to assess the infiltration depth of ESCC. The AI system was then validated on 102 video images, while some endoscopic specialists were invited to view the same video images for diagnoses. The study showed that the accuracy, sensitivity, and specificity of AI for ME diagnosis were 89%, 71%, and 95%, respectively, and those for non-ME diagnosis were 87%, 50%, and 99%. Compared with the diagnostic parameters of endoscopic experts, those of the AI system were mostly higher. This suggests that AI system can provide useful support during endoscopy.

Waki *et al*[53] constructed an AI system based on CNN with 17336 images of ESCC (1376 patients) and 1461 images of noncancerous/normal esophagi (196 patients). While recording the verification video, the endoscopic operator passed through the esophagus at a constant speed to simulate a situation when a lesion was missed. A total of 100 videos (50 ESCCs, 22 noncancerous esophagi, and 28 normal esophagi) were then evaluated by the AI system and 21 endoscopists. The study showed that the sensitivity and specificity of the AI system for ESCC diagnosis were 85.7% and 40%, respectively, and those of the endoscopists were 75% and 91.4%, respectively. With the help of the AI system, the diagnostic specificity of the endoscopists was almost the same, but the sensitivity was improved. Therefore, the AI system, as an auxiliary tool, plays an important role in the diagnosis of ESCC by endoscopists.

PROBLEMS AND PROSPECTS OF AI-ASSISTED ENDOSCOPIC DIAGNOSES

With continuous improvements in endoscopic technology and the diagnostic levels of AI, the combination of AI and endoscopy has become popular. Although AI has made some achievements in the diagnosis of esophageal precancerous lesions and early EC, there are still some problems.

False positive and false negative results

First, almost all AI diagnostic systems yield some false negative results. Small lesions are easily missed in clinical practice, so it is crucial to improve the detection accuracy of these easily neglected lesions. In addition, the AI diagnostic system yields false positive results, which can lead to overtreatment. Shadowed portions, color changes in the gastric antrum and pylorus, and changes in the normal tissue structure and benign lesions (scarring, local atrophy, inflammation, ectopic esophagus, and gastric mucosa) are all reasons for false positive results[54]. To solve this problem, on the one hand, a large number of high-quality endoscopic images should be accumulated for computer algorithm training and verification to produce more accurate results. On the other hand, endoscopic videos often contain more low-resolution real images, which are difficult to capture in still pictures. The use of a large number of images taken from videos as learning materials can reduce the rates of false positives and false negatives to a certain extent.

Retrospective experimental studies have a single source of learning materials, and prospective experimental studies are lacking

At present, most of the training data sets and validation data sets of AI systems have been derived from the same batch of data from the same center. Although the accuracy of AI systems has been internally verified, there is still a lack of external verification [55]. The resolution of examination images obtained by different types of endoscopes varies greatly among different devices. Therefore, future studies should try to include endoscopic image data from multiple institutions, multiple models, and multiple devices to ensure the repeatability of the research results.

In addition, most of the current studies are retrospective, and researchers tend to select clear and high-quality endoscopic images after excluding low-quality images caused by interference factors (such as bleeding, mucus secretion, and food interference), thus resulting in selection bias. This bias often causes the results of retrospective trials to outperform the actual results in clinical applications[56]. In the future, a large number of prospective studies should be carried out to continuously improve AI systems and improve their accuracy, sensitivity, and specificity for clinical trials to lay a solid foundation for the real-time clinical application of AI.

Lack of endoscopic video-assisted diagnoses

Currently, most AI systems are based on the processing of static data rather than the modeling of dynamic videos. Static images are mostly taken after the mucosal environment is well prepared and the lesion location is determined. Due to the lack of environmental impact caused by poor preparation of the mucosal environment and endoscopic movement in dynamic videos, information is missing[57]. There is a large gap between AI training sets and the actual endoscopic working environment, which affects the clinical applicability of AI to some extent. The application of video sets can better solve the above problems. Moreover, endoscopic video analysis can be used for secondary review after real-time endoscopy to quickly identify and screen esophageal diseases and reduce the number of missed diagnoses, as this type of analysis has considerable development potential in DL-assisted endoscopy in the future.

Prospects for development

CAD system based on DL technology has gained increasing attention and is closely related to the good development prospects of DL technology applied in real-time endoscopy. CAD can indicate the lesion site in real-time endoscopic examinations, provide an accurate classification, and serve as a second observer to assist in disease diagnosis. In low-resource or densely populated areas, CAD is used for population-based endoscopic screening, which can avoid missed diagnosis or misdiagnosis of diseases caused by endoscopists' lack of experience and professional knowledge or heavy work fatigue. CAD can be used to train new endoscopists who lack experience, provide them with professional knowledge training, and improve their professional skills. CAD can also be performed online to provide more professional endoscopic diagnoses in areas where experienced endoscopists are lacking, making it easier for patients to visit local hospitals.

CONCLUSION

Most of the current research is still focused on early system development and

feasibility studies, but subsequent product development has not followed. The CAD system based on DL is still in the experimental research stage. Therefore, in the future, a large number of high-quality prospective experimental studies should be carried out in combination with high-quality algorithms and frameworks with more powerful functions, higher efficiency, and better stability. With the establishment of a standardized and large sample data center, the CAD system can provide endoscopic physicians with more accurate diagnosis and treatment options, auxiliary teaching, auxiliary assessments, and telemedicine for early EC. An increasing number of patients and physicians will benefit from the progress of the CAD system.

REFERENCES

1 **Bray F**, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**: 394-424 [PMID: 30207593 DOI: 10.3322/caac.21492]

2 **Ferlay J**, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer* 2019; **144**: 1941-1953 [PMID: 30350310 DOI: 10.1002/ijc.31937]

3 **Thrift AP**. The epidemic of oesophageal carcinoma: Where are we now? *Cancer Epidemiol* 2016; **41**: 88-95 [PMID: 26851752 DOI: 10.1016/j.canep.2016.01.013]

4 **Bibault JE**, Giraud P, Burgun A. Big Data and machine learning in radiation oncology: State of the art and future prospects. *Cancer Lett* 2016; **382**: 110-117 [PMID: 27241666 DOI: 10.1016/j.canlet.2016.05.033]

5 **Esteva A**, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115-118 [PMID: 28117445 DOI: 10.1038/nature21056]

6 **Gulshan V**, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016; **316**: 2402-2410 [PMID: 27898976 DOI: 10.1001/jama.2016.17216]

7 **Le Berre C**, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, Danese S, Peyrin-Biroulet L. Application of Artificial Intelligence to Gastroenterology and Hepatology. *Gastroenterology* 2020; **158**: 76-94.e2 [PMID: 31593701 DOI: 10.1053/j.gastro.2019.08.058]

8 **Shin HC**, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging* 2016; **35**: 1285-1298 [PMID: 26886976 DOI: 10.1109/TMI.2016.2528162]

9 **Yamashita R**, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 2018; **9**: 611-629 [PMID: 29934920 DOI: 10.1007/s13244-018-0639-9]

10 **Mori Y**, Kudo SE, Mohmed HEN, Misawa M, Ogata N, Itoh H, Oda M, Mori K. Artificial intelligence and upper gastrointestinal endoscopy: Current status and future perspective. *Dig Endosc* 2019; **31**: 378-388 [PMID: 30549317 DOI: 10.1111/den.13317]

11 **Soh YSA**, Lee YY, Gotoda T, Sharma P, Ho KY; Asian Barrett's Consortium. Challenges to diagnostic standardization of Barrett's esophagus in Asia. *Dig Endosc* 2019; **31**: 609-618 [PMID: 30892742 DOI: 10.1111/den.13402]

12 **Fitzgerald RC**, di Pietro M, Ragunath K, Ang Y, Kang JY, Watson P, Trudgill N, Patel P, Kaye PV, Sanders S, O'Donovan M, Bird-Lieberman E, Bhandari P, Jankowski JA, Attwood S, Parsons SL, Loft D, Lagergren J, Moayyedi P, Lyratzopoulos G, de Caestecker J; British Society of Gastroenterology. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* 2014; **63**: 7-42 [PMID: 24165758 DOI: 10.1136/gutjnl-2013-305372]

13 **Shaheen NJ**, Falk GW, Iyer PG, Gerson LB; American College of Gastroenterology. ACG Clinical Guideline: Diagnosis and Management of Barrett's Esophagus. *Am J Gastroenterol* 2016; **111**: 30-50; quiz 51 [PMID: 26526079 DOI: 10.1038/ajg.2015.322]

14 **Wani S**, Gaddam S. Editorial: Best Practices in Surveillance of Barrett's Esophagus. *Am J Gastroenterol* 2017; **112**: 1056-1060 [PMID: 28725066 DOI: 10.1038/ajg.2017.117]

15 **Tavakkoli A**, Appelman HD, Beer DG, Madiyal C, Khodadost M, Nofz K, Metko V, Elta G, Wang T, Rubenstein JH. Use of Appropriate Surveillance for Patients With Nondysplastic Barrett's Esophagus. *Clin Gastroenterol Hepatol* 2018; **16**: 862-869.e3 [PMID: 29432922 DOI: 10.1016/j.cgh.2018.01.052]

16 **Sharma P**, Savides TJ, Canto MI, Corley DA, Falk GW, Goldblum JR, Wang KK, Wallace MB, Wolfsen HC; ASGE Technology and Standards of Practice Committee. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on imaging in Barrett's Esophagus. *Gastrointest Endosc* 2012; **76**: 252-254 [PMID: 22817781 DOI: 10.1016/j.gie.2012.05.007]

17 **Schölvinck DW**, van der Meulen K, Bergman JJGHM, Weusten BLAM. Detection of lesions in dysplastic Barrett's esophagus by community and expert endoscopists. *Endoscopy* 2017; **49**: 113-120

[PMID: 27855466 DOI: 10.1055/s-0042-118312]
18 **Hvid-Jensen F**, Pedersen L, Drewes AM, Sørensen HT, Funch-Jensen P. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med* 2011; **365**: 1375-1383 [PMID: 21995385 DOI: 10.1056/NEJMoa1103042]
19 **Sikkema M**, de Jonge PJ, Steyerberg EW, Kuipers EJ. Risk of esophageal adenocarcinoma and mortality in patients with Barrett's esophagus: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2010; **8**: 235-244 [PMID: 19850156 DOI: 10.1016/j.cgh.2009.10.010]
20 **Fleischmann C**, Messmann H. Endoscopic treatment of early esophageal squamous neoplasia. *Minerva Chir* 2018; **73**: 378-384 [PMID: 29843499 DOI: 10.23736/S0026-4733.18.07805-7]
21 **Sami SS**, Iyer PG. Recent Advances in Screening for Barrett's Esophagus. *Curr Treat Options Gastroenterol* 2018; **16**: 1-14 [PMID: 29330747 DOI: 10.1007/s11938-018-0166-2]
22 **Liu J**, Li M, Li Z, Zuo XL, Li CQ, Dong YY, Zhou CJ, Li YQ. Learning curve and interobserver agreement of confocal laser endomicroscopy for detecting precancerous or early-stage esophageal squamous cancer. *PLoS One* 2014; **9**: e99089 [PMID: 24897112 DOI: 10.1371/journal.pone.0099089]
23 **van der Sommen F**, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, Bergman JJ, de With PH, Schoon EJ. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. *Endoscopy* 2016; **48**: 617-624 [PMID: 27100718 DOI: 10.1055/s-0042-105284]
24 **Struyvenberg MR**, de Groof AJ, van der Putten J, van der Sommen F, Baldaque-Silva F, Omae M, Pouw R, Bisschops R, Vieth M, Schoon EJ, Curvers WL, de With PH, Bergman JJ. A computer-assisted algorithm for narrow-band imaging-based tissue characterization in Barrett's esophagus. *Gastrointest Endosc* 2021; **93**: 89-98 [PMID: 32504696 DOI: 10.1016/j.gie.2020.05.050]
25 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929.e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]
26 **de Groof AJ**, Struyvenberg MR, Fockens KN, van der Putten J, van der Sommen F, Boers TG, Zinger S, Bisschops R, de With PH, Pouw RE, Curvers WL, Schoon EJ, Bergman JJGHM. Deep learning algorithm detection of Barrett's neoplasia with high accuracy during live endoscopic procedures: a pilot study (with video). *Gastrointest Endosc* 2020; **91**: 1242-1250 [PMID: 31926965 DOI: 10.1016/j.gie.2019.12.048]
27 **Hong J**, Park BY, Park H. Convolutional neural network classifier for distinguishing Barrett's esophagus and neoplasia endomicroscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2017; **2017**: 2892-2895 [PMID: 29060502 DOI: 10.1109/EMBC.2017.8037461]
28 **Hashimoto R**, Requa J, Dao T, Ninh A, Tran E, Mai D, Lugo M, El-Hage Chehade N, Chang KJ, Karnes WE, Samarasena JB. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). *Gastrointest Endosc* 2020; **91**: 1264-1271.e1 [PMID: 31930967 DOI: 10.1016/j.gie.2019.12.049]
29 **de Groof J**, van der Sommen F, van der Putten J, Struyvenberg MR, Zinger S, Curvers WL, Pech O, Meining A, Neuhaus H, Bisschops R, Schoon EJ, de With PH, Bergman JJ. The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019; **7**: 538-547 [PMID: 31065371 DOI: 10.1177/2050640619837443]
30 **Naveed M**, Kubiliun N. Endoscopic Treatment of Early-Stage Esophageal Cancer. *Curr Oncol Rep* 2018; **20**: 71 [PMID: 30058019 DOI: 10.1007/s11912-018-0713-y]
31 **Yang H**, Hu B. Recent advances in early esophageal cancer: diagnosis and treatment based on endoscopy. *Postgrad Med* 2021; **133**: 665-673 [PMID: 34030580 DOI: 10.1080/00325481.2021.1934495]
32 **Shimizu Y**, Omori T, Yokoyama A, Yoshida T, Hirota J, Ono Y, Yamamoto J, Kato M, Asaka M. Endoscopic diagnosis of early squamous neoplasia of the esophagus with iodine staining: high-grade intra-epithelial neoplasia turns pink within a few minutes. *J Gastroenterol Hepatol* 2008; **23**: 546-550 [PMID: 17573830 DOI: 10.1111/j.1440-1746.2007.04990.x]
33 **Kolb JM**, Wani S. Barrett's esophagus: current standards in advanced imaging. *Transl Gastroenterol Hepatol* 2021; **6**: 14 [PMID: 33409408 DOI: 10.21037/tgh.2020.02.10]
34 **Minami H**, Isomoto H, Inoue H, Akazawa Y, Yamaguchi N, Ohnita K, Takeshima F, Hayashi T, Nakayama T, Nakao K. Significance of background coloration in endoscopic detection of early esophageal squamous cell carcinoma. *Digestion* 2014; **89**: 6-11 [PMID: 24458106 DOI: 10.1159/000356200]
35 **Nagami Y**, Tominaga K, Machida H, Nakatani M, Kameda N, Sugimori S, Okazaki H, Tanigawa T, Yamagami H, Kubo N, Shiba M, Watanabe K, Watanabe T, Iguchi H, Fujiwara Y, Ohira M, Hirakawa K, Arakawa T. Usefulness of non-magnifying narrow-band imaging in screening of early esophageal squamous cell carcinoma: a prospective comparative study using propensity score matching. *Am J Gastroenterol* 2014; **109**: 845-854 [PMID: 24751580 DOI: 10.1038/ajg.2014.94]
36 **Ishihara R**, Takeuchi Y, Chatani R, Kidu T, Inoue T, Hanaoka N, Yamamoto S, Higashino K, Uedo N, Iishi H, Tatsuta M, Tomita Y, Ishiguro S. Prospective evaluation of narrow-band imaging endoscopy for screening of esophageal squamous mucosal high-grade neoplasia in experienced and less experienced endoscopists. *Dis Esophagus* 2010; **23**: 480-486 [PMID: 20095991 DOI:

10.1111/j.1442-2050.2009.01039.x]
37 **Ebigbo A**, Mendel R, Probst A, Manzeneder J, Souza LA Jr, Papa JP, Palm C, Messmann H. Computer-aided diagnosis using deep learning in the evaluation of early oesophageal adenocarcinoma. *Gut* 2019; **68**: 1143-1145 [PMID: 30510110 DOI: 10.1136/gutjnl-2018-317573]
38 **Ebigbo A**, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA Jr, Papa J, Palm C, Messmann H. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. *Gut* 2020; **69**: 615-616 [PMID: 31541004 DOI: 10.1136/gutjnl-2019-319460]
39 **Horie Y**, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019; **89**: 25-32 [PMID: 30120958 DOI: 10.1016/j.gie.2018.07.037]
40 **Cai SL**, Li B, Tan WM, Niu XJ, Yu HH, Yao LQ, Zhou PH, Yan B, Zhong YS. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). *Gastrointest Endosc* 2019; **90**: 745-753.e2 [PMID: 31302091 DOI: 10.1016/j.gie.2019.06.044]
41 **Ohmori M**, Ishihara R, Aoyama K, Nakagawa K, Iwagami H, Matsuura N, Shichijo S, Yamamoto K, Nagaike K, Nakahara M, Inoue T, Aoi K, Okada H, Tada T. Endoscopic detection and differentiation of esophageal lesions using a deep neural network. *Gastrointest Endosc* 2020; **91**: 301-309.e1 [PMID: 31585124 DOI: 10.1016/j.gie.2019.09.034]
42 **Liu G**, Hua J, Wu Z, Meng T, Sun M, Huang P, He X, Sun W, Li X, Chen Y. Automatic classification of esophageal lesions in endoscopic images using a convolutional neural network. *Ann Transl Med* 2020; **8**: 486 [PMID: 32395530 DOI: 10.21037/atm.2020.03.24]
43 **Kumagai Y**, Takubo K, Kawada K, Aoyama K, Endo Y, Ozawa T, Hirasawa T, Yoshio T, Ishihara S, Fujishiro M, Tamaru JI, Mochiki E, Ishida H, Tada T. Diagnosis using deep-learning artificial intelligence based on the endocytoscopic observation of the esophagus. *Esophagus* 2019; **16**: 180-187 [PMID: 30547352 DOI: 10.1007/s10388-018-0651-7]
44 **Guo L**, Xiao X, Wu C, Zeng X, Zhang Y, Du J, Bai S, Xie J, Zhang Z, Li Y, Wang X, Cheung O, Sharma M, Liu J, Hu B. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest Endosc* 2020; **91**: 41-51 [PMID: 31445040 DOI: 10.1016/j.gie.2019.08.018]
45 **Kuwano H**, Nishimura Y, Oyama T, Kato H, Kitagawa Y, Kusano M, Shimada H, Takiuchi H, Toh Y, Doki Y, Naomoto Y, Matsubara H, Miyazaki T, Muto M, Yanagisawa A. Guidelines for Diagnosis and Treatment of Carcinoma of the Esophagus April 2012 edited by the Japan Esophageal Society. *Esophagus* 2015; **12**: 1-30 [PMID: 25620903 DOI: 10.1007/s10388-014-0465-1]
46 **Tokai Y**, Yoshio T, Aoyama K, Horie Y, Yoshimizu S, Horiuchi Y, Ishiyama A, Tsuchida T, Hirasawa T, Sakakibara Y, Yamada T, Yamaguchi S, Fujisaki J, Tada T. Application of artificial intelligence using convolutional neural networks in determining the invasion depth of esophageal squamous cell carcinoma. *Esophagus* 2020; **17**: 250-256 [PMID: 31980977 DOI: 10.1007/s10388-020-00716-x]
47 **Nakagawa K**, Ishihara R, Aoyama K, Ohmori M, Nakahira H, Matsuura N, Shichijo S, Nishida T, Yamada T, Yamaguchi S, Ogiyama H, Egawa S, Kishida O, Tada T. Classification for invasion depth of esophageal squamous cell carcinoma using a deep neural network compared with experienced endoscopists. *Gastrointest Endosc* 2019; **90**: 407-414 [PMID: 31077698 DOI: 10.1016/j.gie.2019.04.245]
48 **Zhao YY**, Xue DX, Wang YL, Zhang R, Sun B, Cai YP, Feng H, Cai Y, Xu JM. Computer-assisted diagnosis of early esophageal squamous cell carcinoma using narrow-band imaging magnifying endoscopy. *Endoscopy* 2019; **51**: 333-341 [PMID: 30469155 DOI: 10.1055/a-0756-8754]
49 **Everson M**, Herrera L, Li W, Luengo IM, Ahmad O, Banks M, Magee C, Alzoubaidi D, Hsu HM, Graham D, Vercauteren T, Lovat L, Ourselin S, Kashin S, Wang HP, Wang WL, Haidry RJ. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. *United European Gastroenterol J* 2019; **7**: 297-306 [PMID: 31080614 DOI: 10.1177/2050640618821800]
50 **Uema R**, Hayashi Y, Tashiro T, Saiki H, Kato M, Amano T, Tani M, Yoshihara T, Inoue T, Kimura K, Iwatani S, Sakatani A, Yoshii S, Tsujii Y, Shinzaki S, Iijima H, Takehara T. Use of a convolutional neural network for classifying microvessels of superficial esophageal squamous cell carcinomas. *J Gastroenterol Hepatol* 2021; **36**: 2239-2246 [PMID: 33694189 DOI: 10.1111/jgh.15479]
51 **Fukuda H**, Ishihara R, Kato Y, Matsunaga T, Nishida T, Yamada T, Ogiyama H, Horie M, Kinoshita K, Tada T. Comparison of performances of artificial intelligence *vs* expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video). *Gastrointest Endosc* 2020; **92**: 848-855 [PMID: 32505685 DOI: 10.1016/j.gie.2020.05.043]
52 **Shimamoto Y**, Ishihara R, Kato Y, Shoji A, Inoue T, Matsueda K, Miyake M, Waki K, Kono M, Fukuda H, Matsuura N, Nagaike K, Aoi K, Yamamoto K, Nakahara M, Nishihara A, Tada T. Real-time assessment of video images for esophageal squamous cell carcinoma invasion depth using artificial intelligence. *J Gastroenterol* 2020; **55**: 1037-1045 [PMID: 32778959 DOI: 10.1007/s00535-020-01716-5]
53 **Waki K**, Ishihara R, Kato Y, Shoji A, Inoue T, Matsueda K, Miyake M, Shimamoto Y, Fukuda H, Matsuura N, Ono Y, Yao K, Hashimoto S, Terai S, Ohmori M, Tanaka K, Kato M, Shono T, Miyamoto H, Tanaka Y, Tada T. Usefulness of an artificial intelligence system for the detection of

esophageal squamous cell carcinoma evaluated with videos simulating overlooking situation. *Dig Endosc* 2021 epub ahead of print [PMID: 33502046 DOI: 10.1111/den.13934]

54 **Struyvenberg MR**, de Groof AJ, Bergman JJ, van der Sommen F, de With PHN, Konda VJA, Curvers WL. Advanced Imaging and Sampling in Barrett's Esophagus: Artificial Intelligence to the Rescue? *Gastrointest Endosc Clin N Am* 2021; **31**: 91-103 [PMID: 33213802 DOI: 10.1016/j.giec.2020.08.006]

55 **Mutasa S**, Sun S, Ha R. Understanding artificial intelligence based radiology studies: What is overfitting? *Clin Imaging* 2020; **65**: 96-99 [PMID: 32387803 DOI: 10.1016/j.clinimag.2020.04.025]

56 **Lazăr DC**, Avram MF, Faur AC, Goldiș A, Romoșan I, Tăban S, Cornianu M. The Impact of Artificial Intelligence in the Endoscopic Assessment of Premalignant and Malignant Esophageal Lesions: Present and Future. *Medicina (Kaunas)* 2020; **56**: 364 [PMID: 32708343 DOI: 10.3390/medicina56070364]

57 **Zhang YH**, Guo LJ, Yuan XL, Hu B. Artificial intelligence-assisted esophageal cancer management: Now and future. *World J Gastroenterol* 2020; **26**: 5256-5271 [PMID: 32994686 DOI: 10.3748/wjg.v26.i35.5256]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com

ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 December 28; 2(6): 211-219

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 6 December 28, 2021

MINIREVIEWS

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 6 December 28, 2021

Contents

ABOUT COVER

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Lin-YuTong Wang*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jin-Lei Wang*.

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Fatih Altintoprak, Sahin Coban, Krish Ragunath

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
December 28, 2021

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

Artificial Intelligence in Gastrointestinal Endoscopy

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Gastrointest Endosc 2021 December 28; 2(6): 211-219

DOI: 10.37126/aige.v2.i6.211

ISSN 2689-7164 (online)

MINIREVIEWS

Artificial intelligence in polyp detection - where are we and where are we headed?

Kristen E Dougherty, Vatche J Melkonian, Grace A Montenegro

ORCID number: Kristen E Dougherty 0000-0002-2714-9062; Vatche J Melkonian 0000-0001-8111-5923; Grace A Montenegro 0000-0002-4443-4893.

Author contributions: Dougherty KE and Melkonian VJ and Montenegro GA contributed equally to this work; Dougherty KE and Melkonian VJ and Montenegro GA performed writing, review and editing of the manuscript.

Conflict-of-interest statement: None of the authors have disclosures.

Country/Territory of origin: United States

Specialty type: Gastroenterology and hepatology

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): D
Grade E (Poor): 0

Kristen E Dougherty, Vatche J Melkonian, Grace A Montenegro, Department of Surgery, St. Louis University Hospital, Saint Louis, MO 63110, United States

Corresponding author: Grace A Montenegro, FACS, MD, MS, Assistant Professor, Surgeon, Department of Surgery, St. Louis University Hospital, 1008 S. Spring Avenue, SLUCare Academic Pavilion, Saint Louis, MO 63110, United States. grace.montenegro@health.slu.edu

Abstract

The goal of artificial intelligence in colonoscopy is to improve adenoma detection rate and reduce interval colorectal cancer. Artificial intelligence in polyp detection during colonoscopy has evolved tremendously over the last decade mainly due to the implementation of neural networks. Computer aided detection (CADe) utilizing neural networks allows real time detection of polyps and adenomas. Current CADe systems are built in single centers by multidisciplinary teams and have only been utilized in limited clinical research studies. We review the most recent prospective randomized controlled trials here. These randomized control trials, both non-blinded and blinded, demonstrated increase in adenoma and polyp detection rates when endoscopists used CADe systems *vs* standard high definition colonoscopes. Increase of polyps and adenomas detected were mainly small and sessile in nature. CADe systems were found to be safe with little added time to the overall procedure. Results are promising as more CADe have shown to have ability to increase accuracy and improve quality of colonoscopy. Overall limitations included selection bias as all trials built and utilized different CADe developed at their own institutions, non-blinded arms, and question of external validity.

Key Words: Neural networks; Computer aided detection; Artificial intelligence in colonoscopy and polyp detection; Artificial intelligence in adenoma detection

Core Tip: Use of computer aided detection (CADe) in colonoscopy has been shown to increase polyp and adenoma detection rates compared to standard high-definition colonoscopy with little added procedure time. Additionally, CADe have been built to increase quality of screening colonoscopy. These advantages and features have been

Received: April 27, 2021
Peer-review started: April 28, 2021
First decision: May 19, 2021
Revised: July 2, 2021
Accepted: November 18, 2021
Article in press: November 18, 2021
Published online: December 28, 2021

P-Reviewer: Mochizuki K
S-Editor: Liu M
L-Editor: A
P-Editor: Liu M

demonstrated in blinded and non-blinded randomized controlled trials.

Citation: Dougherty KE, Melkonian VJ, Montenegro GA. Artificial intelligence in polyp detection - where are we and where are we headed? *Artif Intell Gastrointest Endosc* 2021; 2(6): 211-219
URL: https://www.wjgnet.com/2689-7164/full/v2/i6/211.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i6.211

INTRODUCTION

Colorectal cancer (CRC) is the second leading cause of cancer death in the United States. Adenomas are the most common type of precancerous polyp. Colonoscopy remains the gold standard for identifying these precancerous polyps and is the only nonsurgical intervention capable of removing them. The National polyp study showed that up to 90% of CRCs are preventable with polyp removal[1]. The adenoma detection rate (ADR) represents the percent of colonoscopies in which at least 1 adenoma is found. ADR is regarded as the main quality indicator of colonoscopy and ideally ADR should equal adenoma prevalence, estimated to be greater than 50%[2]. Unfortunately, ADRs vary widely, with some endoscopists having ADRs as low as 7%[3]. It has been shown that for each 1% increase in ADR, the interval CRC rate was decreased by 3%-6%[3,4]. The main cause of interval CRC incidence is overlooked lesions due to failure of recognition or incomplete mucosal exposure due to suboptimal technique during the withdrawal phase of colonoscopy[5]. Artificial intelligence in colonoscopy was expected to address these issues in hopes to reduce polyp detection miss rates and subsequently interval CRCs[6,7].

COMPUTER-AIDED DETECTION

The concept of computer-aided detection (CADe) in polyp detection was first described in the early 2000s where software was developed that utilized color and texture to identify polyps[8]. Polyp detection accuracy was as high as 95% however only applicable on static images due to high latency. Tajbakhsh *et al*[9] created CADe based on hybrid shape analysis. The system sensitivity reached close to 90% however proved un-competitive for real time video stream with high latency[9].

CADe of polyps has evolved exponentially since 2012 when deep learning models began utilizing convolutional neural networks (CNNs) to identify polyp-specific features independent of human input.

CNNs utilize statistical pattern recognition algorithms to identify an object, in this case, a polyp. In brief, the computer recognizes an array of numbers, or picture variables, based on pixel analysis of the captured images. The input layer is then filtered through several hidden layers each acting as distinctive feature identifiers or recognizable features. For example, if the desired outcome is for the CNN to recognize a discrete face, hidden layers would include the nose, mouth, eyebrows, *etc.* The fully connected layer comes at the end of this neural network and analyzes the output from previous layers to determine which features correlate best to a certain class, *i.e.*, the probability of the image being Jack's face *vs* Joe's face. The higher the probability of identifying the image in effect strengthens the network[10].

CNNs are created and utilized by multiple disciplines including computer science, bioinformatics, machine learning/intelligent systems, and increasingly in healthcare and medicine. CNNs afford the ability to detect images, in this case polyps, in real time analysis.

Training these networks involves providing a groundwork of data sets or images. Urban *et al*[11] utilized five different data sets: First, data including over one million images of non-medical objects. Second, a set of over 8600 colonoscopy images containing over 4000 images of unique polyps of varying size and morphologies, as well as over 4500 images without polyps. Third, a separate set of 1330 colonoscopy images, half showing unique polyps and half showing other non-polyp images collected from different patients. Fourth, videos of colonoscopies, and fifth, a larger data set augmenting the original set of colonoscopy images. This model identified

polyps with a 96.4% accuracy rate and demonstrated the ability to work in real time conditions with a processing rate of one frame per 10 milliseconds (ms). It identified all polyps discovered by expert viewers (ADRs > 50%) as well as any additional polyps that were missed. The authors believe utilizing real-time live analysis with this model during colonoscopies will prompt increasingly careful inspection and lead to discovery of additional polyps that may have been missed[11].

Five randomized control trials (RCTs) utilizing independent CADe systems are reviewed here demonstrating significant improvement in ADR compared to standard colonoscopy (Table 1).

RCT

In 2019, Wang *et al*[12] presented the first prospective single center RCT (Sichuan Provincial, China) investigating the influence of an automatic polyp detection method based on deep learning regarding the polyp detection rate and ADR. The study scheme was a non-blinded trial in which subjects who underwent diagnostic colonoscopy with or without assistance of a real-time automated polyp detection system. The primary outcome was ADR. The real-time automatic polyp detection system was based on SegNet architecture. The algorithm was authenticated and had a per image sensitivity of 94.4% and per image specificity of 95.9%. The system handled at least 25 frames per second with a dormancy of 76.8 ± 5.6 ms in simultaneous video analysis. The monitor was parallel with the original endoscopy monitor and it provided simultaneous visual notice and audible alarm when a polyp was detected. Subjects who had colonoscopy from September 2017 to February 2018 were suitable for enrollment and bowel preparation and high definition colonoscope's were standardized. Exclusion criteria included inflammatory bowel disease, CRC, previous unsuccessful colonoscopy, and high suspicion for polyposis syndromes.

Eight endoscopists participated in the study, half of which who were junior endoscopists. The experience was as follows – two seasoned endoscopists (> 20000 colonoscopies), two midlevel endoscopist (3000-10000), and four junior endoscopist (100-500).

Standard colonoscopy was completed in the control group. In the CADe the endoscopist was assisted by the real-time automatic detection system. The system captured the endoscopy video and displayed the polyp location with a blue box on a neighboring screen with a coinciding audible alert. The system was turned on during withdrawal only. The endoscopist was obligated to check every polyp location detected by the system devoid of assistance. A missed polyp was delineated as a polyp confirmed by the endoscopist but unobserved by the system. A false alarm was delineated as detected lesion which was interminably traced by the system deemed by the endoscopist not to be a polyp.

Five hundred thirty-six patients were randomized prospectively into the control group and 522 into the CADe group. There were no statistical differences in demographics, total time of colonoscopy, no polyp withdrawal time or withdrawal time excluding biopsy, bowel preparation and endoscopist experience. There was a statistical difference with withdrawal time of 6.39 min in the routine colonoscopy *vs* 6.89 minutes in the CADe group.

A 1.89-fold increase was found in the mean number of polyps discovered between the two groups [95% confidence interval (CI): 1.63 to 2.192, $P < 0.001$]. The PDR of the control and CADe group were 0.29 and 0.45, respectively [odds ratio (OR), 1.995; 95%CI: 1.532-2.544, $P < 0.001$]. They found a 1.72-fold increase in the mean number of adenomas discovered. The ADR of the control and CADe groups were 0.20 and 0.29, respectively (OR, 1.61; 95%CI: 1.213 to 2.135, $P < 0.001$). The number of detected polyps was significantly higher in the CADe group when looking specifically at non-pedunculated polyps, polyps 0 cm to 1 cm in size and polyps in all portions of the colon. There was also a considerably higher number of adenomas detected in the CADe group when looking at non-pedunculated polyps, polyps smaller than 0.5 cm and polyps in all portions of the colon except for the cecum and ascending colon. There was a total of 39 false positives in the CADe group. Of discovered polyps in the CADe cohort, none were missed by the automatic system.

ADR in the CADe group showed a trend of 6% increase in the subgroup of patients with excellent bowel preparation. In addition, their results, including the mean number of detected adenomas, mean number of detected polyps and PDR, were significantly increased. However, this was not statistically significant given the small sample size.

Table 1 Five randomized control trials utilizing independent computer aided detection systems

Ref.	Study design	Blinded?	Type of system	ADR control	ADR CADe	False alarms (per colonoscopy)	Missed polyps
Wang *et al*[12], 2019	RCT	No	CADe	0.2	0.29 ($P < 0.001$)	0.075	0
Repici *et al*[13], 2020	RCT	No	CADe	0.404	0.548 ($P < 0.001$)	-	-
Liu *et al*[14], 2020	RCT	No	CADe	0.23	039 ($P < 0.001$)	0.071	0
Wang *et al*[2], 2020	RCT	Yes	CADe	0.28	0.34 ($P = 0.03$)	0.1	0
Su *et al*[15], 2020	RCT	No	CADe + Quality	0.165	0.289 ($P < 0.001$)	0.2	0

RCT: Randomized control trial; CADe: Computer aided detection.

Limitations of this study include the inability to blind the endoscopists of each arm. In addition, the adenoma and polyp detection rates in this study are substantially lower than what is reported in Western countries, and thus there is a question of whether this study is applicable in centers with higher ADRs at baseline[12].

Repici *et al*[13] published work on a separate CNN by the GI genius, Medtronic system in 2020. The system was trained and validated with 99.7% per lesion sensitivity and 0.9% false-positive frames. Using a series of videos of 2684 histologically confirmed polyps from 840 patients. They performed a multicenter randomized trial to assess the safety and efficacy of this CADe in detection of colorectal neoplasia during real-time colonoscopy. Like Wang *et al*[12], the operator was not blinded to the study arm assigned to each patient. Colonoscopies were performed by 6 experienced endoscopist to from each center with over 2000 screening colonoscopies; inexperienced endoscopists were not included. High definition colonoscopes were utilized. The CADe system would signal the endoscopist with a bounding box only when a target polyp was recognized in the image. Primary outcome was ADR. Secondary outcomes were proximal ADR, total number of polyps detected, sessile serrated lesions detection rate, mean number of adenomas per colonoscopy (APC), cecal intubation rate and withdrawal time.

Patient's undergoing colonoscopy from September to November 2019 were included. Colonoscopy requirements included colorectal screening or post polypectomy surveillance as well as work-up following FIT positivity or patients with appropriate signs and symptoms warranting further work up. Patients were excluded in the case of personal history of CRC or Inflammatory bowel disease (IBD), previous colon resection, or antithrombotic therapy precluding polyp resection.

A total of 685 patients were randomized, 341 in CADe arm and 344 in the control arm. There was no significant difference in terms of bowel preparation or cecal intubation rate. ADR in the CADe group was 54.8% *vs* 40.4% in the control group. After adjusting for age, gender, and indication the ADR was significantly higher in the CADe group compared to the control.

The CADe group identified more non-polypoid (26.6% *vs* 18.4%) and polypoid (37.3% *vs* 26.5%) lesions compared to control. The proportion of patients with < 10 mm adenomas was higher in the CADe group, 44.3%, *vs* in the control group, 32.3%. The difference between the 2 arms was significant for both ≤ 5 mm and 6 mm to 9 mm adenomas. Regarding location, the proportion of patients with proximal adenomas was higher in the CADe group then in the control group. This was also true for distal adenomas. Forty-five patients were diagnosed with advanced neoplasia in the CADe group compared with 36 in control group, demonstrating a detection rate for advanced neoplasia of 13.3% and 10.5% respectively.

Of the 460 patients who underwent polyp resection, 120 did not have histologically proven adenomas, sessile serrated lesions, or CRC. The non-neoplastic resection rate for CADe and control were 26% and 28.8%, respectively.

Repici *et al*[13] demonstrated that addition of real-time CADe to colonoscopy resulted in 30% and 46% relative increase in ADR and APC. Safety of the system was demonstrated by the lack of increase of both useless resections and withdrawal time. Computer aided detection efficacy appeared to be independent of morphology and location of neoplasia and was mainly explained by the additional detection of polyps that were less than 5 mm, or between 6 mm to 9 mm in size. Limitations of this study

were like Wang *et al*[12] in that the endoscopists in each arm were not blinded. In addition, they did not include inexperienced endoscopists in their study. They demonstrated the safety and efficacy of integrating a CADe with colonoscopy with a substantial improvement of ADR and adenoma per colonoscopy without increasing the removal of non-neoplastic lesions. This is likely to improve the quality of colonoscopy without affecting efficiency[13].

In 2020, Liu *et al*[14] published their work using yet another CNN or CADe system. Polyp positive videos (151) and polyp negative videos (384) were used to design the system. This system utilized spatiotemporal data to recognize polyps, which is presumed to be more suitable for video data sets. This was a prospective, single center, randomized control study (China) to demonstrate the effective of CADe on the detection rate of polyps and adenomas during colonoscopy. Bowel preparations and high definition colonoscope's were standardized. Exclusion criteria included inflammatory bowel disease, history of CRC surgery, history of radiotherapy and/or chemotherapy and biopsy contraindications. CADe was only utilized during withdrawal phase. The system processed each frame and displayed the detected polyp. When the lesion would appear on the screen a voice alarm would prompt endoscopist to view the system. This study was done without the assistance of nurses, trainees, or staff. Polyps identified by the endoscopist but not identified by the CADe system were deemed "missed polyps" and were documented. False alarms were defined as lesions detected and continuously tracked but were not identified as polyps by the system.

A total of 1026 patients were eligible: 518 in the control and 508 in the computer-aided detected group. The two groups were similar in demographics and risk factors. Total withdrawal time in the standard group *vs* the CADe group was 6.74 min and 6.82 min, respectively ($P < 0.001$).

A total of 734 polyps were detected. Three hundred and ninety-two of these were adenomas representing 53.41%, and 31 were sessile serrated adenomas representing 4.22%. In total 248 polyps were detected in the control group and 486 polyps in the CADe group, a rate of 33.79% *vs* 66.21%, respectively. The corresponded to 1.53 times increase in the average number of polyps detected in both (95%CI: 1.652-2.297, $P < 0.001$). The polyp detection rates in control and CADe were 0.28 and 0.44, respectively which corresponded to 1.51 times increase in number of adenomas detected (95%CI: 1.423-2.016, $P < 0.001$). The detection rates of adenomas in control *vs* CADe group were 0.23 and 0.39, respectively (CI: 1.201- 2.220, $P < 0.001$). The number of polyps detected in the CADe group was significantly higher than control group when looking at sessile polyps, polyps 0 cm to 1 cm in size, and polyps in all portions of the colon. The number of adenomas detected in the CADe group also increased significantly when considering sessile polyps, polyps ≤ 0.5 cm, and polyps in all parts of the colon excluding the cecum and ascending colon.

Similar to Wang *et al*[12], detection rate in the CADe was higher when intestinal preparation was deemed adequate. Insufficient sample size of the subgroup analysis failed to show statistical significance. There were 36 false alarms in the CADe group corresponding to an average of 0.071 false alarms per colonoscopy. Of all polyps detected in the CADe group no polyps were missed. In mirroring the results of previous RCTs, this study again demonstrated significantly higher detection rates of adenoma, and average number of polyps and adenomas by colonoscopy in the CADe group when contrasted to control groups. However, the overall rise in adenoma detection was mainly due to the rise in detection of small adenomas, less than 1 cm in most instances.

The study review shows that integration of computer aided detection systems can effectively detect polyps that were otherwise missed by the endoscopist, however there is a blind area with polyps that remain undetected which remains an unanswered problem. Similar to Wang *et al*[12] study limitations include non-blinded endoscopists and low ADRs of endoscopists, as compared to Western countries. In conclusion the study showed this CADe increases the detection rate of colorectal polyps and adenomas, therefore depicting its feasibility for detection of polyps and adenomas on colonoscopy[14].

The first single center, randomized, double-blind trial to evaluate the use of automatic polyp detection using the CAD system during colonoscopy was published in early 2020 by Wang *et al*[12]. They enrolled consecutive patients between September 2018 and January 2019. After all exclusion criteria, there were 484 patients in the CADe group and 478 in the sham group. All qualified patients were randomized 1:1 to either white light colonoscopy with CADe assistance or to the control group consisting of white light colonoscopy with a mock system. Patients were not notified of their assignment and blinding of the operating endoscopists was achieved by the mock

system. Four senior endoscopists participated, each with at least 5 years' experience, completing a minimum of 1000 colonoscopy procedures per year.

Endoscopists were told to perform all colonoscopy procedures with the aid of a CADe system, and they were unaware of the use of the mock system. Authors utilized the same CADe system as previously discussed in their non-blinded trial. A mock system was designed to appropriately mask the endoscopists. This system simulated alert boxes on polyp-like non polyp structures without tracking actual polyps during colonoscopy. The sham model was built using a portion of the images used to develop the CADe system as previously described, producing a much lower sensitivity and specificity. The grouping and outputs of the mock and computer aided detection systems were only observable to a senior endoscopist using a separate monitor acting as a second observer.

In both study groups, if the operating endoscopist did not recognize an abnormality within an alert box, the observer was tasked with informing the location of any alert box for the operating endoscopist using a laser pointer on the principal monitor. An alarm sounded to the observer *via* earpiece if an alert box was visible. The observer was not blinded to the intervention and aware which system was being used. All observable alert boxes were documented by the observer, however only the alert boxes that were on non-polyp structures were recorded as consistent false detections in the CADe, *i.e.*, false alarms. Consistent false detections of the CADe were recorded and believed not to be a polyp by the operating endoscopist. In the CADe group the observer also recorded any missed detections, defined as a polyp discovered by the working endoscopist and proved by histology but not alerted by the CADe.

After the clinical analysis, they analyzed the videos of polyps that were detected by CADe, but initially missed by the operating endoscopist. The video clips were then independently reviewed by an addition 3 skilled, experienced endoscopists who did not partake in the clinical trial. Endoscopists labeled each video when they identified a polyp, and they were limited to single viewing of the videos. Analysis of sensitivity and specificity on these easily overlooked polyps was performed.

The primary outcome was proportion of individuals who underwent a complete colonoscopy and had 1 or more adenomas detected. Secondary outcomes were the proportion of individuals undergoing a complete colonoscopy who had 1 or more polyps detected, the number of polyps per colonoscopy, and the number of APC, which was calculated by dividing the total number of polyps that are adenomas detected by the total number of colonoscopies done.

Four hundred seventy-eight patients were allocated to sham group and 484 to the CADe. No difference was in terms of demographics and adenoma detection probability features. There were no recorded untoward events with these procedures. There was a statistically significant increase in withdrawal time with the CADe group, 7.46 min *vs* 6.99 min ($P < 0.0001$).

More biopsies were performed for polyps in the CADe group than in the mock group. When omitting the time taken to do biopsies, the mean withdrawal time was not statistically significant between the groups. Overall, 809 polyps were detected, of which 38% were found in the mock group and 62% were found in the CADe group. Of these polyps, 57% were adenomas and 4% were sessile serrated adenomas. When considering shape (sessile) and size (0-5 mm), the CADe group had a significantly higher number of detected polyps and adenomas.

Notably, there was a 1.61-fold increase in polyps detected per colonoscopy between the 2 groups (95%CI: 1.39-1.85; $P < 0.0001$). The PDR was significantly higher in the CADe group then with the mock system, 52% *vs* 37%. A 1.53-fold increase in APC between the groups (95%CI: 1.27-1.85; $P < 0.0001$). The ADR was significantly higher with the CADe system compared to the mock system with 28% in the sham control group and 34% in the CADe group having an adenoma detected. Based on the observers' judgement, here were 48 false detections in the CADe group, averaging 0.1 per colonoscopy. Of all the detected polyps in the experimental group, none were missed by the CADe system.

An average of 0.17 adenomas and 0.33 polyps per patient were overlooked initially by the endoscopist in the CADe group. These polyps were small (mean adenoma size 3.89 mm), isochromatic, flat in shape, had unclear boundaries, were partly behind colon folds, and were on the edge of the visual field. The sensitivity and specificity of three skilled endoscopist during review of the endoscopy videos was 17% and 64%, respectively.

Once again, Wang *et al*[12] demonstrated a CADe system can effectually increase the number of polyps and adenomas detected with colonoscopy, and after controlling for operational bias. The CADe system had higher sensitivity and specificity for detection of easy to overlook polyps compared to evaluation based solely on the utilities of the

human endoscopist.

The main contribution of CADe in this system demonstrated a rise detection of diminutive and non-pedunculated, non-advanced adenomas and hyperplastic polyps. The CADe system is safe and effectual approach to increase ADR during colonoscopy [2].

In 2020, Su *et al*[15] published a prospective randomized control study comparing CADe with control, however, their CADe was built with quality control, specifically supervising withdrawal stability, from five different neural networks (AQCS). Similar to the other trials mentioned, AQCS was turned on during withdrawal. In addition to visual cue when polyps detected, the added features in this system were notifying the endoscopist to slow down withdrawal speed and re-examine colonic segments when unstable or blurry frames were detected continuously and prompting endoscopist to clean mucosa when inadequate score (Boston Prep Score < 2) was given by system. Study was single center in Qilu Hospital in China, from October 2018 to May 2019. Study included patients over 18yo who were able to give consent for screening colonoscopy. Exclusion criteria included history of IBD, CRC or colorectal surgery, patients with previously failed colonoscopy, or highly suspicious for polyposis syndromes, or patients whose colonoscopy could not be completed due to stenosis or large occupying lesions. Six endoscopists participated, each with 4-6 years of experience and colonoscopy volume of 5000 to 8000. Endoscopists were not blinded to randomization status, however, patients, data collection and study analyses were blinded.

A total of 659 patients were randomized, after exclusions, 308 in AQCS group and 315 in control. There were no differences in demographics between the two groups. They performed retrospective review of 459 normal colonoscopies without positive findings performed by participating colonoscopists and there was no significant difference in the mean withdrawal times between the two groups.

A total number of 169 adenomas were detected, 56 control and 113 AQCS group. ADR in control 16.51% *vs* 28.9% AQCS group (OR, 2.055; 95%CI: 1.397-3.024; $P < 0.001$). Additionally, there were more adenomas detected in AQCS when non pedunculated, diminutive adenomas (≤ 5 mm), larger adenomas (> 5 mm), and adenomas in all segments with exception of cecum and rectum were considered.

Polyp detection rates were also higher in the AQCS group, 38.1% *vs* 25.4% in control (OR, 1.824; 95%CI: 1.296-2.569; $P = 0.001$).

Withdrawal time excluding biopsy time was significantly longer in the AQCS group than in the control group (7.03 ± 1.01 min *vs* 5.68 ± 1.26 min, $P < 0.001$). For the AQCS group there was a significant improvement in withdrawal time in this study compared to their retrospective withdrawal times before this study. Adequate bowel preparation rate was 87.3% in the AQCS group *vs* 80.6% in control group, (OR, 1.656; 95%CI: 1.070-2.564; $P = 0.023$).

There were 62 false prompts (false positives), averaging 0.201 false prompts per colonoscopy and no missed prompts (false negatives) in the AQCS group.

Su *et al*[15] demonstrated significant improvement in ADR when utilizing a quality-controlled CADe system that supervises in withdrawal stability and prompts endoscopists to clean colonic mucosa when inadequate prep scores are recognized. The AQCS demonstrated significant increase in ADR and an increase detection of larger adenomas, compared to previously mentioned CADe systems.

Limitations of this study are similar to prior RCTs where endoscopists are not blinded to randomization. It is also a single center study with only experienced endoscopists participating. Authors mention that the system utilized 4 intra-procedural quality metrics, and these combined, improved ADR. They did not perform preliminary testing to evaluate whether 2 or 3 metrics would increase ADR to standard colonoscopy.

CONCLUSION

The goal of computer aided detection of polyps and adenomas is to close the gap between ADR and adenoma prevalence and in turn reduce interval CRC rates. CADe systems could act as second observers and reduce miss-rates of polyps. Implementation of CNNs for image recognition has overhauled the playing field regarding artificial intelligence utilization in colonoscopy, as these networks are built to allow image recognition in real time. As mentioned above, multiple CADe systems are being built and programmed by multidisciplinary teams from bioinformatics, computer science, machine learning/intelligent systems and in medicine. A single system has

not been shown to be superior to others.

As demonstrated by randomized trials, the ability to integrate CADe with colonoscopy in real time has demonstrated overall ADR and polyp detection rates were significantly higher for CADe groups compared with control. These were most significant for small, diminutive polyps and adenomas ≤ 5 mm and those which were sessile in character. These CADe systems have been shown to be safe and efficient. The CADe systems mentioned here have scarce miss rates, if any, when it comes to polyp detection. Small adenomas have less probability for malignant transformation compared to larger, however the increase in total ADR may contribute to decreased risk of interval CRC. Conversely, increased detection and resection of diminutive lesions may represent additional unnecessary polypectomies and add to workload, cost, and pathology resource utilization. Wang *et al*[12] remarks that gaining the ability to identify small adenomas will provide the advantage of resecting small pre-malignant lesions along with distinguishing patients who are at higher risk for future adenomas and interval cancer. Sue *et al*[15]'s AQCS network demonstrated increase detection of larger adenomas. Their system was unique to others in that it was built with a quality control feature that essentially improved the quality of colonoscopy by improving withdrawal time and identifying inadequate exposure of mucosa. Detection of polyps and adenomas by CADe relies on exposure of the entirety of the colonic mucosal field by the endoscopist. Polyps that remain outside of the visual field still pose a major deficiency that have the potential to be addressed by this system.

Endoscopists are still responsible for proper execution of the colonoscopy procedure, including cecal intubation rate. Inexperienced endoscopists are likely to have suboptimal results in the technical exposure of colorectal mucosa and perhaps adding quality control to CADe is the answer as demonstrated by Su *et al*[15]. Additionally, the AQCS was not utilized by inexperienced endoscopists in that RCT.

Artificial intelligence in colonoscopy has certainly made strides over the last decade, specifically in real time detection. Currently, CADe systems based on CNN for the use of polyp and adenoma detection during colonoscopy are being built in single centers. This poses a risk of selection bias leading to difficulty implementing any one CADe system on a wide scale. Appropriately curated large scale data sets are needed to limit data set bias. Collection of image and video inputs should be broad and include unsampled or under-represented lesions. The added complexity of developing CADe to assist in withdrawal stability and identification of inadequate exposure elevates the technology of AI as these enhance the ability and accuracy of the endoscopist who remains the critical portion of the colonoscopy, for now[15-17]. In addition to polyp detection, models built to aid in diagnosis and classification of inflammatory bowel disease have been described[18,19]. Current systems should have controlled and practical set up, as not add to the workflow of standard colonoscopy. Ideally these systems should predict pathology and size and improve accuracy, minimizing unnecessary pathologic assessment and avoidable resection of non-neoplastic lesions. While it is expected that technologic cost will increase initially, when used effectively and efficiently, CADe systems should ultimately reduce cost. The review of RCTs demonstrates undeniable improvement of ADR when utilizing CADe compared to standard colonoscopy. Collectively they demonstrate CADe are safe and practical when used in real-time and more complex CADe systems have the potential to improve accuracy of the endoscopist improving quality of colonoscopy.

REFERENCES

1 **Winawer SJ**, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, Waye JD, Schapiro M, Bond JH, Panish JF. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med* 1993; **329**: 1977-1981 [PMID: 8247072 DOI: 10.1056/NEJM199312303292701]

2 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

3 **Corley DA**, Jensen CD, Marks AR, Zhao WK, Lee JK, Doubeni CA, Zauber AG, de Boer J, Fireman BH, Schottinger JE, Quinn VP, Ghai NR, Levin TR, Quesenberry CP. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 1298-1306 [PMID: 24693890 DOI: 10.1056/NEJMoa1309086]

4 **Kaminski MF**, Wieszczy P, Rupinski M, Wojciechowska U, Didkowska J, Kraszewska E, Kobiela J, Franczyk R, Rupinska M, Kocot B, Chaber-Ciopinska A, Pachlewski J, Polkowski M, Regula J. Increased Rate of Adenoma Detection Associates With Reduced Risk of Colorectal Cancer and Death.

Gastroenterology 2017; **153**: 98-105 [PMID: 28428142 DOI: 10.1053/j.gastro.2017.04.006]

5 **Zhao S**, Wang S, Pan P, Xia T, Chang X, Yang X, Guo L, Meng Q, Yang F, Qian W, Xu Z, Wang Y, Wang Z, Gu L, Wang R, Jia F, Yao J, Li Z, Bai Y. Magnitude, Risk Factors, and Factors Associated With Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology* 2019; **156**: 1661-1674.e11 [PMID: 30738046 DOI: 10.1053/j.gastro.2019.01.260]

6 **Hassan C**, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, Antonelli G, Yu H, Areia M, Dinis-Ribeiro M, Bhandari P, Sharma P, Rex DK, Rösch T, Wallace M, Repici A. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021; **93**: 77-85.e6 [PMID: 32598963 DOI: 10.1016/j.gie.2020.06.059]

7 **Bisschops R**, East JE, Hassan C, Hazewinkel Y, Kamiński MF, Neumann H, Pellisé M, Antonelli G, Bustamante Balen M, Coron E, Cortas G, Iacucci M, Yuichi M, Longcroft-Wheaton G, Mouzyka S, Pilonis N, Puig I, van Hooft JE, Dekker E. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. *Endoscopy* 2019; **51**: 1155-1179 [PMID: 31711241 DOI: 10.1055/a-1031-7657]

8 **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**: 151-166 [PMID: 12507791 DOI: 10.1016/s0169-2607(02)00007-x]

9 **Tajbakhsh N**, Gurudu SR, Liang J. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information. *IEEE Trans Med Imaging* 2016; **35**: 630-644 [PMID: 26462083 DOI: 10.1109/TMI.2015.2487997]

10 **Lee H**, Grosse R, Ranganath R, Ng A. Convolutional Deep Belief networks for scalable Unsupervised learning of hierarchical representations. ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning. Association for Computing Machinery, 2009: 609-616 [DOI: 10.1145/1553374]

11 **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078.e8 [PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037]

12 **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

13 **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology* 2020; **159**: 512-520.e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

14 **Liu WN**, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, Huang J. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; **26**: 13-19 [PMID: 31898644 DOI: 10.4103/sjg.SJG_377_19]

15 **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc* 2020; **91**: 415-424.e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

16 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: 30268542 DOI: 10.1016/j.gie.2018.09.024]

17 **Karnes WE**, Ninh A, Dao T, Requa J, Samarasena JB. Real-time identification of anatomic landmarks during colonoscopy using deep learning. *Gastrointest Endosc* 2018; **87**: AB252 [DOI: 10.1016/j.gie.2018.04.447]

18 **Mossotto E**, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep* 2017; **7**: 2427 [PMID: 28546534 DOI: 10.1038/s41598-017-02606-2]

19 **Karnes WE**, Ninh A, Dao T, Requa J, Samarasena JB. Unambiguous real-time scoring of bowel preparation using artificial intelligence. *Gastrointest Endosc* 2018; **87**: AB258 [DOI: 10.1016/j.gie.2018.04.461]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com