ISSN 2644-3236 (online)

Artificial Intelligence in *Gastroenterology*

Artif Intell Gastroenterol 2022 December 28; 3(5): 117-162

Published by Baishideng Publishing Group Inc

AIG

Artificial Intelligence in Gastroenterology

Contents

Bimonthly Volume 3 Number 5 December 28, 2022

REVIEW

Artificial Intelligence in Gastroenterology

Bimonthly Volume 3 Number 5 December 28, 2022

Contents

ABOUT COVER

Editorial Board Member of *Artificial Intelligence in Gastroenterology*, Alina Ioana Tantau, FACE, PhD, Academic Research, Chief Doctor, Professor, Senior Researcher, Department of Internal Medicine and Gastroenterology, "Iuliu Hatieganu" University of Medicine and Pharmacy, 4th Medical Clinic, Cluj-Napoca 400012, Romania. alitantau@gmail.com

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastroenterology* (*AIG*, *Artif Intell Gastroenterol*) is to provide scholars and readers from various fields of artificial intelligence in gastroenterology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIG mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastroenterology and covering a wide range of topics, including artificial intelligence in gastrointestinal cancer, liver cancer, pancreatic cancer, hepatitis B, hepatitis C, nonalcoholic fatty liver disease, inflammatory bowel disease, irritable bowel syndrome, and *Helicobacter pylori* infection.

INDEXING/ABSTRACTING

The *AIG* is now abstracted and indexed in Reference Citation Analysis, China Science and Technology Journal Database.

RESPONSIBLE EDITORS FOR THIS ISSUE

NAME OF JOURNAL
Artificial Intelligence in Gastroenterology

ISSN
ISSN 2644-3236 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Rajvinder Singh, Ferruccio Bonino

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2644-3236/editorialboard.htm

PUBLICATION DATE
December 28, 2022

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 Gastroenterology

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

Artif Intell Gastroenterol 2022 December 28; 3(5): 117-141

DOI: 10.35712/aig.v3.i5.117

ISSN 2644-3236 (online)

REVIEW

Artificial intelligence in gastroenterology: A narrative review

Jonathan S Galati, Robert J Duve, Matthew O'Mara, Seth A Gross

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): B
Grade C (Good): C
Grade D (Fair): D
Grade E (Poor): 0

P-Reviewer: Liu XQ, China; Qi XS, China

Received: October 9, 2022
Peer-review started: October 9, 2022
First decision: October 29, 2022
Revised: November 21, 2022
Accepted: December 21, 2022
Article in press: December 21, 2022
Published online: December 28, 2022

Jonathan S Galati, Department of Medicine, NYU Langone Health, New York, NY 10016, United States

Robert J Duve, Department of Internal Medicine, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, Buffalo, NY 14203, United States

Matthew O'Mara, Seth A Gross, Division of Gastroenterology, NYU Langone Health, New York, NY 10016, United States

Corresponding author: Jonathan S Galati, MD, Department of Medicine, NYU Langone Health, 550 First Avenue, New York, NY 10016, United States. jonathan.galati@nyulangone.org

Abstract

Artificial intelligence (AI) is a complex concept, broadly defined in medicine as the development of computer systems to perform tasks that require human intelligence. It has the capacity to revolutionize medicine by increasing efficiency, expediting data and image analysis and identifying patterns, trends and associations in large datasets. Within gastroenterology, recent research efforts have focused on using AI in esophagogastroduodenoscopy, wireless capsule endoscopy (WCE) and colonoscopy to assist in diagnosis, disease monitoring, lesion detection and therapeutic intervention. The main objective of this narrative review is to provide a comprehensive overview of the research being performed within gastroenterology on AI in esophagogastroduodenoscopy, WCE and colonoscopy.

Key Words: Artificial intelligence; Colonoscopy; Computer-aided detection; Deep learning; Endoscopy; Machine learning

Core Tip: Artificial intelligence (AI) is a complex concept that has the capacity to revolutionize medicine. Within gastroenterology, recent research efforts have focused on using AI in esophagogastroduodenoscopy, wireless capsule endoscopy (WCE) and colonoscopy to assist in diagnosis, disease monitoring, lesion detection and therapeutic intervention. This narrative review provides a comprehensive overview of the research being performed within gastroenterology on AI in esophagogastroduodenoscopy, WCE and colonoscopy.

Citation: Galati JS, Duve RJ, O'Mara M, Gross SA. Artificial intelligence in gastroenterology: A narrative review. *Artif Intell Gastroenterol* 2022; 3(5): 117-141
URL: https://www.wjgnet.com/2644-3236/full/v3/i5/117.htm
DOI: https://dx.doi.org/10.35712/aig.v3.i5.117

INTRODUCTION

Artificial intelligence (AI) is a complex concept, broadly defined in medicine as the development of computer systems to perform tasks that require human intelligence[1]. Since its inception in the 1950s, the field of AI has grown considerably (Figure 1)[2]. Often AI is accompanied by the terms machine learning (ML) and deep learning (DL), techniques used within the field of AI to develop systems that can learn and adapt without explicit instructions. Machine learning uses self-learning algorithms that derive knowledge from data to predict outcomes[1]. There are two main categories within ML: Supervised and unsupervised learning. In supervised learning, the AI is trained on a dataset in which human intervention has previously assigned a hierarchy of features which allows the algorithm to understand differences between data inputs and classify or predict outcomes[3]. In unsupervised learning, the system is provided a dataset that has not been categorized by human intervention. The algorithm then analyzes the data with the goal of identifying labels or patterns[3].

Deep learning is a subfield of ML that utilizes artificial neural networks (ANN) to analyze data. In DL, the system is able to analyze raw data and determine features that distinguish between data inputs. ANN systems are composed of interconnected nodes in a layered structure similar to how neurons are organized in the human brain. The weight of the connections between each node influences how the system can recognize, classify, and describe objects within data[3,4]. ANNs with multiple layers of nodes are classified as deep neural networks which form the backbone of deep learning.

Artificial intelligence has the capacity to revolutionize medicine. It can be used to increase efficiency by aiding in appointment scheduling, reviewing insurance eligibility, or tracking patient history. AI can also expedite data and image analysis and detect patterns, trends and associations[5]. Within gastroenterology, AI's prominence stems from its utility in image analysis[5,6]. Many gastrointestinal diseases rely on endoscopic evaluation for diagnosis, disease monitoring, lesion detection and therapeutic intervention. However, endoscopic evaluation is heavily operator dependent and thus subject to operator bias and human error. As such, recent efforts have focused on using AI in esophagogastroduodenoscopy, wireless capsule endoscopy (WCE) and colonoscopy to mitigate these issues, serving as an additional objective observer of the intestinal tract. The main objective of this narrative review is to provide a comprehensive overview of the research being performed within gastroenterology on artificial intelligence in esophagogastroduodenoscopy, WCE and colonoscopy. While other narrative reviews have been published regarding the use of artificial intelligence in esophagogastroduodenoscopy, WCE and colonoscopy, this narrative review goes a step further by providing a granular and more technical assessment of the literature. As such, this narrative review is intended for medical providers and researchers who are familiar with the use of artificial intelligence in esophagogastroduodenoscopy, WCE and colonoscopy and are interested in obtaining an in-depth review in a specific area.

LITERATURE REVIEW

Electronic databases Embase, Ovid Medicine, and PubMed were searched from inception to September 2022 using multiple search queries. Combinations of the terms "artificial intelligence", "AI", "computer aided", "computer aided detection", "CADe", "convolutional neural network", "deep learning", "DCNN", "machine learning", "colonoscopy", "endoscopy", "wireless capsule endoscopy", "capsule endoscopy", "WCE", "esophageal cancer", esophageal adenocarcinoma", "esophageal squamous cell carcinoma", "gastric cancer", "gastric neoplasia", "gastric lesions", "Barrett's esophagus", "celiac disease", "Helicobacter pylori", "Helicobacter pylori infection", "H pylori", "H pylori infection", "gastric ulcers", "duodenal ulcers", "inflammatory bowel disease", "IBD", "ulcerative colitis", "Crohn's disease", "parasitic infections", "hookworms", "bleeding", "gastrointestinal bleeding", "vascular lesions", "angioectasias", "polyp", "polyp detection", "tumor", "gastrointestinal tumor", "small bowel tumor", "bowel preparation", "Boston bowel preparation scale", "BBPS", "adenoma", "adenoma detection", "adenoma detection rate", "sessile serrated lesion", and "sessile serrated lesion rate" were used. We subsequently narrowed the results to clinical trials in human published within the last 10 years.

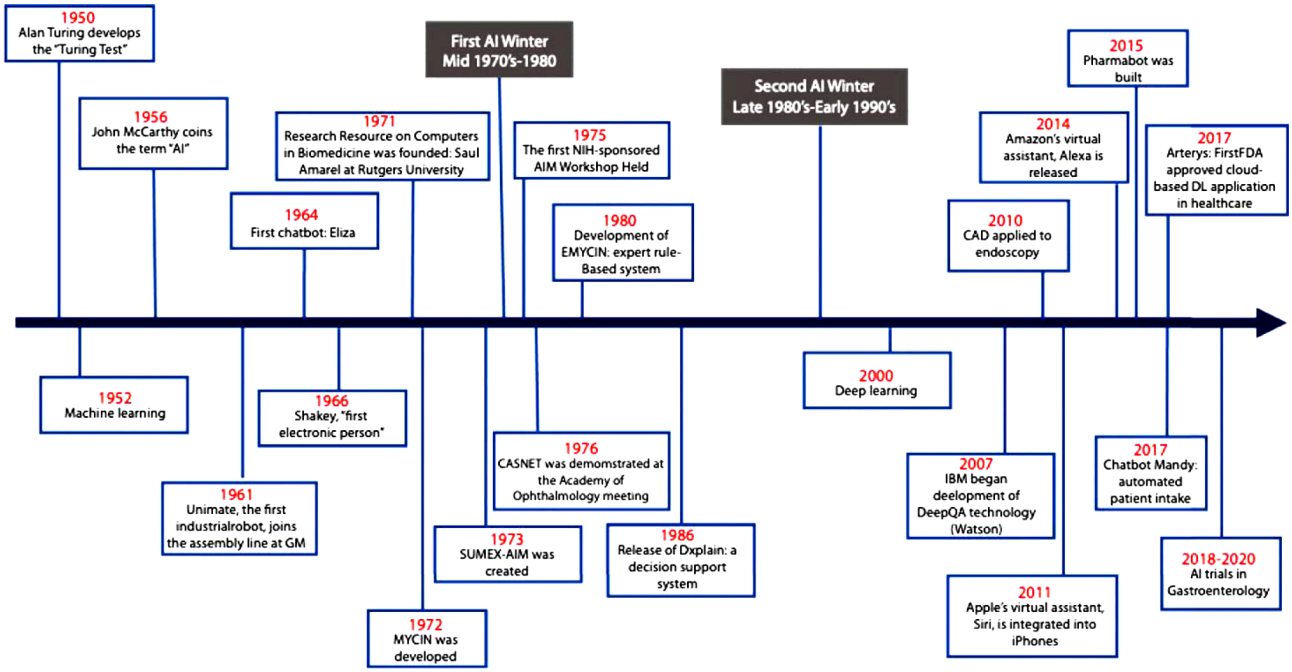

Figure 1 Timeline of the development and use of artificial intelligence in medicine. AI: Artificial intelligence; DL: Deep learning; FDA: U.S. Food and Drug Administration; CAD: Computer-aided diagnosis. Reprinted with permission from Elsevier Science & Technology Journals[2].

ESOPHAGOGASTRODUODENOSCOPY

Barrett's esophagus and esophageal adenocarcinoma

Barrett's esophagus (BE) is a premalignant condition associated with esophageal adenocarcinoma (EAC)[7-9]. It is caused by chronic inflammation and tissue injury of the lower esophagus as a result of gastric reflux[7-9]. Early detection and diagnosis can prevent the progression of BE to EAC[7-9]. Patients with BE should undergo routine surveillance endoscopies to monitor for progression. However, even with surveillance, dysplastic changes can be easily missed[7]. To improve the detection of dysplastic changes in BE, researchers have focused on developing AI systems to assist with the identification of dysplasia and early neoplasia during endoscopic evaluation.

Since 2016, a group of researchers from the Netherlands have developed numerous AI systems to identify neoplastic lesions in BE[10-16]. Their first publication detailed their experience using a support vector machine (SVM), a ML method, to identify early neoplastic lesions from white light endoscopy (WLE) images[10]. Their SVM achieved a sensitivity and specificity of 83% with respect to per-image detection and sensitivity of 86% and specificity of 87% with respect to per-patient detection[10]. In their next study, the group trialed several different feature extraction and ML methods using volumetric laser endomicroscopy (VLE) images[11]. They received the best results with the feature extraction module "layering and signal decay statistics", achieving high sensitivity (90%) and specificity (93%) with area under the curve (AUC) 0.95 for neoplastic lesion detection[11]. Following this, they conducted a second studying again using ML in VLE to identify neoplastic lesions in BE, however, they used a multiframe analysis approach, including frames neighboring the region of interest in the analysis[12]. With this approach, they found that multiframe analysis resulted in a significantly higher median AUC when compared to single frame analysis (0.91 *vs* 0.83; $P < 0.001$)[12]. Continuing to use ML methods, the group published their finding from the ARGOS project – a consortium of three international tertiary referral centers for Barrett's neoplasia[13]. In this study, de Groof *et al*[13] created a computer-aided detection (CADe) system that used SVM to classify images. The group tested the CADe with 60 images – 40 images from patients with a neoplastic lesion, 20 images from patients with non-dysplastic Barrett's esophagus. The CADe achieved an AUC of 0.92 and a sensitivity, specificity and accuracy of 95%, 85% and 92% respectively for detecting neoplastic lesions[13].

Following their successes creating ML systems for neoplastic lesion detection, the group of researchers from the Netherlands shifted their focus to DL methods. In their first foray into DL, they developed a hybrid CADe system using architecture from ResNet and U-Net models. The CADe was trained with 494364 labeled endoscopic images and subsequently refined with a data set comprised of 1247 WLE images. It was finally tested on a set of 297 images (129 images with early neoplasia, 168 with non-dysplastic BE) where the hybrid CADe system attained a sensitivity of 87.6%, specificity of 88.6% and accuracy of 88.2% for identifying early neoplasia[14]. The system was also tested in two external validation sets where it achieved similar results. A secondary outcome of the study was to see if within the images classified as having neoplasia if the CADe could delineate the neoplasia and recommend a

site for biopsy. The ground truth was determined by expert endoscopists. In two external data sets (external validation data set 4 and 5), the CADe identified the optimal biopsy site in 97.2% of cases and 91.9% of cases respectively[14]. Using a similar hybrid CADe, the group performed a pilot study testing the CADe during live endoscopic procedures[15]. Overall, the CADe achieved a sensitivity of 75.8%, specificity of 86.5% and accuracy of 84% in per-image analyses[15]. Their most recent study again used their hybrid ResNet and U-Net CADe to identify neoplastic lesions in narrow-band imaging (NBI)[16]. With respect to NBI images, the CADe was found to have sensitivity of 88% (95%CI 86%-94%), specificity of 78% (95%CI 72%-84%), and accuracy of 84% (95%CI 81%-88%) for identifying BE neoplasia [16]. In per frame and per video analyses, the CADe achieved sensitivities of 75% and 85%, specificities of 90% and 83% and accuracies of 85% and 83% respectively[16].

Outside of this group from the Netherlands, several other researchers have created DL systems for the detection of BE neoplasia[17-21]. Hong *et al*[17] created a CNN that could distinguish between intestinal metaplasia, gastric metaplasia and neoplasia from images obtained by endomicroscopy in patients with Barrett's esophagus with accuracy of 80.8%. Ebigbo *et al*[18] created a DL-CADe capable of detecting BE neoplasia with sensitivity 83.7%, specificity of 100.0% and accuracy of 89.9%. Two other groups achieved similar results to Ebigbo *et al*[18]: Hashimoto *et al*'s CNN detected early neoplasia with sensitivity of 96.4%, specificity of 94.2%, and accuracy of 95.4% and Hussein *et al*'s CNN detected early neoplasia with sensitivity 91%, specificity 79%, area under the receiver operating characteristic (AUROC) of 93%[19,20]. An overview of these studies is provided in Table 1.

In addition to neoplasia detection, some groups started to use AI to grade BE and predict submucosal invasion of lesions. Ali *et al*[22] recently published the results from a pilot study using a DL system to quantitatively assess BE area (BEA), circumference and maximal length (C&M). They tested their DL system on 3D printed phantom esophagus models with different BE patterns and 194 videos from 131 patients with BE. In the phantom esophagus models, the DL system achieved an accuracy of 98.4% for BEA and 97.2% for C&M[22]. In the patient videos, the DL system differed from expert endoscopists by 8% and 7% for C&M respectively[22]. Ebigbo *et al*[23], building upon their earlier success using a DL CADe to detect neoplasia, performed a pilot study using a 101-layer CNN to differentiate T1a (mucosal) and T1b (submucosal) BE related cancers. Using 230 WLE images obtained from three tertiary care centers in Germany, their CNN was capable of discerning T1a lesions from T1b lesions with sensitivity, specificity and accuracy of 77%, 64% and 71% respectively, comparable to the expert endoscopists enrolled in the study[23].

Despite BE's potential progression to EAC if left unmanaged, few studies have explicitly looked at using AI to detect EAC. Ghatwary *et al*[24] tested several DL models on 100 WLE images (50 featuring EAC, 50 featuring normal mucosa) to determine which was best at identifying EAC. They found that the Single-Shot Multibox Detector (SSD) method achieved the best results, attaining a sensitivity of 96% and specificity of 92%[24]. In 2021, Iwagami *et al*[25] focused on developing an AI system to identify esophagogastric junctional adenocarcinomas. They used SSD for their CNN, achieving a sensitivity, specificity and accuracy of 94%, 42% and 66% for detecting esophagogastric junctional adenocarcinomas. Their CNN performed similarly to endoscopists enrolled in the study (sensitivity 88%, specificity 43%, accuracy 66%)[25].

Esophageal squamous cell carcinoma

Esophageal squamous cell carcinoma (ESCC) is the most common histologic type of esophageal cancer in the world[26]. While certain imaging modalities such as Lugol's chromoendoscopy and confocal microendoscopy are effective at improving the accuracy, sensitivity and specificity of targeted biopsies, they are expensive and not universally available[27]. In recent years, efforts have focused on developing AI systems to support lower cost imaging modalities in order to improve their ability to detect ESCC.

Shin *et al*[27] and Quang *et al*[28] created ML algorithms which they tested on high-resolution microendoscope images, obtaining comparable sensitivities for the detection of ESCC (98% and 95% respectively). Following these studies, several groups created DL systems to detect ESCC[29-38]. In Cai *et al*'s study, their deep neural network-CADe was tested on 187 images obtained from WLE. The system obtained good sensitivity (97.8%), specificity (85.4%) and accuracy (91.4%) for identifying ESCC [29]. Similar findings occurred in three separate studies that used deep convolutional neural networks (DCNNs) to detect ESCC in WLE[30-32]. Using NBI, Guo *et al*[33] created a CADe that achieved high sensitivity (98.0%), specificity (95.0%) and an AUC of 0.99 for detecting ESCC in still images. Similar results were obtained in Li *et al*'s study[35]. For detecting ESCC in NBI video clips, Fukuda *et al*[34] obtained different results, finding similar sensitivity (91%) to Guo *et al*[33] however substantially lower specificity (51%). Three studies compared a DL-CADe with WLE to DL-CADe with NBI for the detection of ESCC[32,35,36]. The results from these three studies were quite discordant and as such a statement regarding whether a DL-CADe with WLE or DL-CADe with NBI is better for the detection of ESCC cannot be made at this time.

Interestingly, several studies used DL algorithms to assess ESCC invasion depth[39,40]. Everson *et al* [39] and Zhao *et al*[40] created CNNs to detect intrapapillary capillary loops, a feature of ESCC that correlates with invasion depth, in images obtained from magnification endoscopy with NBI. They achieved similar findings with Everson *et al*'s CNN achieving an accuracy of 93.7% and Zhao *et al*'s achieving an accuracy of 89.2%[39,40]. Using DL, two groups created DCNNs to directly detect ESCC

Table 1 Overview of findings from studies evaluating the detection accuracy of computer-aided detection for Barrett's esophagus-related neoplasia

Ref.	Country	Study design	AI Classifier	Lesions	Training dataset	Test dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUROC
Swager *et al* [11], 2017	Netherlands	Retrospective	ML[2] methods	NPL	-	60 VLE images	90	93	-	0.95
van der Sommen *et al* [10], 2016	Netherlands	Retrospective	SVM	NPL	-	100 WLE images	83	83	-	-
Hong *et al* [17], 2017	South Korea	Retrospective	CNN	NPL, IM, GM	236 endomicroscopy images	26 endomicroscopy images	-	-	80.77	-
de Groof *et al* [13], 2019	Netherlands, Germany, Belgium	Prospective	SVM	NPL	-	60 WLE images	95	85	91.7	0.92
Ebigbo *et al* [21], 2019	Germany, Brazil	Retrospective	CNN	EAC	Augsburg dataset: 148 WLE images and NBI; MICCAI dataset: 100 WLE images		97; 94[a]; 92	88; 80[a]; 100	-	-
Ghatwary *et al*[24], 2019	England, Egypt	Retrospective	Multiple CNNs	EAC	Images from 21 patients	Images from 9 patients	96	92	-	-
de Groof *et al* [14], 2020	Netherlands, France, Sweden, Germany, Belgium, Australia	Ambispective	CNN	NPL	Dataset 1: 494364 images; Dataset 2:1; 247 images; Dataset 3: 297 images	Dataset 3: 297 images; Dataset 4: 80 images; Dataset 5: 80 images	90[b]	87.5[b]	88.8[b]	-
de Groof *et al* [15], 2020	Netherlands, Belgium	Prospective	CNN	NPL	495611 images	20 patients; 144 WLE images	75.8	86.5	84	-
Ebigbo *et al* [18], 2020	Germany, Brazil	Prospective	CNN	EAC	129 images	62 images	83.7	100	89.9	-
Hashimoto *et al*[19], 2020	United States	Retrospective	CNN	NPL	1374 images	458 images	96.4	94.2	95.4	-
Struyvenberg *et al*[12], 2020	Netherlands	Prospective	ML[2] methods	NPL	-	3060 VLE frames	-	-	-	0.91
Iwagami *et al* [25], 2021	Japan	Retrospective	CNN	EJC	3443 images	232 images	94	42	66	-
Struyvenberg *et al*[16], 2021	Netherlands, Sweden, Belgium	Retrospective	CNN	NPL	495611 images	157 NBI zoom videos; 30021 frames	85[1]; 75	83[1]; 90	83[1]; 85	-
Hussein *et al* [20], 2022	England, Spain, Belgium, Austria	Prospective	CNN	DPL	148936 frames	264 *iscan-1* images	91	79	-	0.93

[a]Sensitivity and specificity reported by white light endoscopy images from the Augsburg dataset, narrow band images from the Augsburg dataset, and from the MICCAI dataset respectively.

[b]Results found from convolutional neural network analyzing dataset 4.

[1]Sensitivity, specificity and accuracy obtained from per-video analysis and from per-frame analysis respectively.

[2]Multiple machine learning (ML) methods tested. Results from best performing ML method reported.

AI: Artificial intelligence; AUROC: Area under the receiver operating characteristic; CNN: Convolutional neural network; DPL: Dysplasia; EAC: Esophageal adenocarcinoma; EJC: Esophagogastric junctional adenocarcinoma; GM: Gastric metaplasia; IM: Intestinal metaplasia; ML: Machine learning; NBI: Narrow band images; NPL: Neoplasia; SVM: Support vector machine; VLE: Volumetric laser endomicroscopy; WLE: White light endoscopy.

invasion depth[41-43]. One group from Osaka International Cancer Institute conducted two studies using SSD to create their DCNNs[41,42]. The DCNNs were made to classify images as EP-SM1 or EP-SM2-3 as this distinction in ESCC bares clinical significance. The studies (Nakagawa *et al*[41] and Shimamoto *et al*[42]) attained similar accuracies and specificities, however had substantially different sensitivities (90.1% *vs* 50% and 71%)[41,42]. The third study, Tokai *et al*[43], used SSD as well for their DCNN and also programed the DCNN to classify images as EP-SM1 or EP-SM2-3. Their observed sensitivity, specificity and accuracy were lower than those found by Nakagawa *et al*[41] (84.1%, 73.3% and 80.9% respectively).

Gastric cancer

Gastric cancer is the third leading cause of cancer-related mortality in the world[44,45]. Early detection of precancerous lesions or early gastric cancer with endoscopy can prevent progression to advanced disease[46]. However, a substantial number of upper gastrointestinal cancers are missed placing patients at risk for interval development[45]. To mitigate this risk, AI systems are being develop to assist with lesion detection.

In 2013, Miyaki *et al*[47] used a bag-of-features framework with densely sampled scale-invariant feature transform descriptors to classify still images obtained from magnifying endoscopy with flexible spectral imaging color enhancement as having or not having gastric cancer. Their system, a rudimentary version of ML, obtained good sensitivity (84.8%), specificity (87.0%) and accuracy (85.9%) for identifying gastric cancer[47]. Using SVM, Kanesaka *et al*[48] found higher sensitivity (96.7%), specificity (95%) and accuracy (96.3%).

Following these successes, several groups began using CNNs for the identification of gastric cancer [44-46,49-59]. In 2018, Hirasawa *et al*[44] published one of the first papers to use a CNN (SSD) to detect gastric cancer. In a test set of 2296 images, the CNN had a sensitivity of 92.2% for identifying gastric cancer lesions[44]. In a larger study, Tang *et al*[49] created a DCNN to detect gastric cancer in a test set of 9417 images and 26 endoscopy videos. With respect to their test set, the DCNN performed well, achieving a sensitivity of 95.5% (95%CI 94.8%–96.1%), specificity of 81.7% (95%CI 80.7%–82.8%), accuracy of 87.8% (95%CI 87.1%–88.5%) and AUC 0.94[49]. The DCNN continued to perform well in external validation sets, achieving sensitivity of 85.9%-92.1%, specificity of 84.4%-90.3%, accuracy of 85.1%-91.2% and AUC 0.89-0.93[49]. Compared to expert endoscopists, the DCNN attained higher sensitivity, specificity and accuracy. In the video set, the DCNN achieved a sensitivity of 88.5% (95%CI 71.0%-96.0%)[49]. Several studies using DCNN to detect gastric cancer in endoscopy images obtained similar sensitivities, specificities and accuracies to Tang *et al*[49]. While one study reported a sensitivity of 58.4% for detecting gastric cancer, the sensitivity for the study's 67 endoscopists was 31.9%[55].

Recently, several groups from China and Japan have published studies using CNNs with magnified endoscopy with NBI (ME-NBI) in an effort to improve early gastric cancer detection[56-59]. Using a 22-layer CNN, Horiuchi *et al*[56] achieved a sensitivity, specificity and accuracy of 95.4%, 71.0% and 85.3% respectively for identifying early gastric cancer from a set of 258 ME-NBI images (151 gastric cancer, 107 gastritis). The same group published a similar study the following year however using ME-NBI videos instead of still images[57]. They obtained similar results: sensitivity of 87.4% (95%CI 78.8%-92.8%), specificity of 82.8% (95%CI 73.5%-89.3%) and accuracy of 85.1% (955 CI 79.0%-89.6%)[57]. Hu *et al*[58] and Ueyama *et al*[59] in their studies using CNN to identify gastric cancer in ME-NBI achieved similar sensitivities, specificities and accuracies as Horiuchi *et al*[56]. An overview of these studies is provided in Table 2.

Of increasing interest to researchers within this field is predicting invasion depth of gastric cancer using AI. Few studies have used CNNs to predict invasion depth[60-63]. Yoon *et al*[60] created a CNN to predict gastric cancer lesion depth from standard endoscopy images. The CNN achieved good sensitivity (79.2%) and specificity (77.8%) for differentiating T1a (mucosal) from T1b (submucosal) gastric cancers (AUC 0.851)[60]. Also using standard endoscopy images, Zhu *et al*[61] attained similar results. They trained their CNN to identify P0 (restricted to the mucosa or < 0.5 mm within the muscularis mucosae) *vs* P1 (≥ 0.5 mm deep into the muscularis mucosae) lesions. The CNN achieved a sensitivity of 76.6%, specificity of 95.6%, accuracy of 89.2% and AUROC 0.94 (95%CI 0.90-0.97). Cho *et al*[62] using DenseNet-161 as their CNN and Nagao *et al*[62] using ResNet50 as their CNN obtained comparable results to Zhu *et al*[61] for predicting gastric cancer invasion depth from endoscopy images.

Gastric ulcers

Within recent years, numerous studies have been published regarding the use of AI to assist with the detection and classification of gastric lesions. Few of these studies explicitly used AI systems to detect duodenal and gastric ulcers, however they report data pertaining to ulcer detection.

Using YOLOv5, a deep learning object detection model, Ku *et al*[64] created a CADe system capable of detecting multiple gastric lesions with good precision (98%) and sensitivity (89%). Also using YOLO for their DCNN, Yuan *et al*[53] achieved an overall system accuracy of 85.7% for gastric lesion identification. With respect to peptic ulcer detection, their system achieved an accuracy of 95.4% (93.5%-97.2%), sensitivity of 86.2% (77.5%–94.8%) and specificity of 96.8% (95.1%–98.4%)[53]. Guo *et al*[54] used ResNet50 to construct their CADe designed to detect gastric lesions. Their CADe achieved lower sensitivity 71.4% (95%CI 69.5–73.2%) and specificity 70.9% (95%CI 70.3–71.4%) than Yuan *et al*'s DCNN [53], however Guo *et al*[54] combined erosions and ulcers into one category for analysis. With their primary outcome being classifying gastric cancers and ulcers, Namikawa *et al*[52] developed a CNN capable of identifying gastric ulcers with high sensitivity (93.3%; 95%CI 87.3%−97.1%) and specificity (99.0%; 95%CI 94.6%-100%).

Helicobacter pylori infection

As a risk factor for future development of gastric cancer, early detection and eradication of *Helicobacter pylori* (*H. pylori*) in infected individuals is important. Endoscopic evaluation for *H. pylori* is highly

Table 2 Overview of findings from studies evaluating the detection accuracy of computer-aided detection for gastric cancer

Ref.	Country	Study design	AI classifier	Lesions	Training dataset	Test dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUROC
Miyaki *et al* [47], 2013	Japan	Prospective[a]	SVM	Gastric cancer	493 FICE-derived magnifying endoscopic images	92 FICE-derived magnifying endoscopic images	84.8	97	85.9	-
Kanesaka *et al*[48], 2018	Japan, Taiwan	Retrospective	SVM	EGC	126 M-NBI images	81 M-NBI images	96.7	95	96.3	-
Wu *et al* [50], 2019	China	Retrospective	CNN	EGC	9151 images	200 images	94	91	92.5	-
Cho *et al* [51], 2019	South Korea	Ambispective	CNN	Advanced gastric cancer, EGC, high grade dysplasia, low grade dysplasia, non-neoplasm	4205 WLE images	812 WLE images; 200 WLE images	-	-	86.6[b]; 76.4	0.877[b]
Tang *et al* [49], 2020	China	Retrospective	CNN	EGC	35823 WLE images	Internal: 9417 WLE images; External: 1514 WLE images[1]	95.5[1]; 85.9-92.1	81.7[1]; 84.4-90.3	87.8[1]; 85.1-91.2	0.94[1]; 0.887-0.925
Namikawa *et al*[52], 2020	Japan	Retrospective	CNN	Gastric cancer	18410 images	1459 images	99	93.3	99	-
Horiuchi *et al*[56], 2020	Japan	Retrospective	CNN	EGC	2570 M-NBI images	258 M-NBI images	95.4	71	85.3	0.852
Horiuchi *et al*[57], 2020	Japan	Retrospective	CNN	EGC	2570 M-NBI images	174 videos	87.4	82.8	85.1	0.8684
Guo *et al* [54], 2021	China	Retrospective	CNN	Gastric cancer, erosions/ulcers, polyps, varices	293162 WLE images	33959 WLE images	67.5[2]; 85.1	70.9[2]; 90.3	-	-
Ikenoyama *et al*[55], 2021	Japan	Retrospective	CNN	EGC	13584 WLE and NBI images	2940 WLE and NBI images	58.4	87.3	-	-
Hu *et al* [58], 2021	China	Retrospective	CNN	EGC	M-NBI images from 170 patients	Internal: M-NBI from 73 patients External: M-NBI images from 52 patients	79.2[3]; 78.2	74.5[3]; 74.1	77[3]; 76.3	0.808[3]; 0.813
Ueyama *et al*[59], 2021	Japan	Retrospective	CNN	EGC	5574 M-NBI images	2300 M-NBI	98	100	98.7	-
Yuan *et al* [53], 2022	China	Retrospective	CNN	EGC, advanced gastric cancer, submucosal tumor, polyp, peptic ulcer, erosion, and lesion-free gastric mucosa	29809 WLE images	1579 WLE images	59.2[4]; 100	99.3[4]; 98.1	93.5[4]; 98.4	-

[a]Presumed prospective based on manuscript.

[b]Accuracy of convolutional neural network (CNN) for detecting the five different lesions and detecting gastric cancer respectively. Area under the receiver operating characteristic (AUROC) pertains to detecting gastric cancer.

[1]The external dataset was comprised of images from 3 external sites. Sensitivity, specificity, accuracy and AUROC for the internal dataset and external dataset respectively.

[2]Sensitivity and specificity of CNN for detecting gastric cancers in a dataset comprised of images without annotations and for detecting gastric cancers in a dataset comprised of annotated images respectively.

[3]Sensitivity, specificity, accuracy and AUROC for the internal dataset and external dataset respectively.

[4]Sensitivity, specificity and accuracy for detecting early gastric cancer and for detecting advanced gastric cancer respectively.

"Internal" and "External" refer to internal and external datasets respectively.

CNN: Convolutional neural network; AUROC: Area under the receiver operating characteristic; EGC: Early gastric cancer; FICE: Flexible spectral imaging

color enhancement; M-NBI: Magnifying endoscopy with narrow band imaging; NBI: Narrow band imaging; SVM: Support vector machine; WLE: White light endoscopy.

operator dependent[65]. Pairing artificial intelligence with endoscopy for the detection of *H. pylori* could possibly reduce false results.

Shichijo *et al*[66] used GoogLeNet, a DCNN consisting of 22 layers, to evaluate 11481 images obtained from 397 patients (72 *H. pylori* positive, 325 negative) for the presence or absence of *H. pylori* infection. GoogLeNet attained a sensitivity of 81.9% (95%CI 71.1%-90.0%), specificity of 83.4% (95%CI 78.9%-87.3%) and accuracy of 83.1% (95%CI 79.1%-86.7%) with AUROC 0.89 for detecting *H. Pylori* infection [66]. When compared to endoscopists enrolled in the study, the sensitivity, specificity and accuracy attained by GoogLeNet was comparable to those attained by the endoscopists[66]. This same group published a second study in 2019 again using GoogLeNet for their DCNN[67]. However, a different optimization technique was used to prepare GoogLeNet. The DCNN was tasked with classifying images as *H. pylori* positive, negative or eradicated. In a set of 23699 images, the DCNN attained an accuracy of 80% for *H. pylori* negative, 84% for *H. pylori* eradicated, and 48% for *H. pylori* positive[67]. Also using GoogLeNet, Itoh *et al*[68] obtained similar results to Shichijo *et al*'s 2017 study with respect to sensitivity (86.7%) and specificity (86.7%)[66]. Using ResNet-50 as their architectural unit for their DCNN, Zheng *et al*[69] were successful in classifying images as *H. pylori* positive or negative, achieving a sensitivity, specificity, accuracy and AUC of 81.4% (95%CI 79.8%–82.9%), 90.1% (95%CI 88.4%–91.7%), 84.5% (95%CI 83.3%–85.7%) and 0.93 (95%CI 0.92–0.94) respectively.

Taking a different approach, Yasuda *et al*[70] used linked color imaging (LCI) with SVM to identify *H. pylori* infection. The LCI images were classified into high-hue and low-hue images based on redness and classified by SVM as *H. pylori* positive or negative. This method attained a sensitivity, specificity and accuracy of 90.4%, 85.7% and 87.6% respectively[70]. Combining LCI with a deep learning CADe system, Nakashima *et al*[71] achieved a sensitivity, specificity and accuracy of 92.5%, 80.0%, 84.2% for identifying *H. pylori* negative images, 62.5%, 92.5%, 82.5% for *H pylori* positive images, 65%, 86.2%, 79.2% for *H. pylori* post-eradication images respectively.

Celiac disease

While immunological tests can support the diagnosis of celiac disease, definitive diagnosis requires histological assessment of duodenal biopsies[72]. As such being able to identify changes in the duodenal mucosa consistent with celiac disease is important. However, these changes can be subtle and difficult to appreciate. Few studies have been published using a CADe system to detect or diagnose celiac disease.

In 2016, Gadermayr *et al*[73] created a system that combined expert knowledge acquisition with feature extraction to classify duodenal images obtained from 290 children as Marsh-0 (normal mucosa) or Marsh-3 (villous atrophy). Expert knowledge acquisition was achieved by having one of three study endoscopists assign a Marsh grade of 0 or 3 to an image. Feature extraction was accomplished using one of three methods: (1) multi-resolution local binary patterns; (2) multi-fractal spectrum; and (3) improved Fisher vectors. From expert knowledge acquisition and feature extraction, their classification algorithm identified images as Marsh-0 or Marsh-3. With optimal settings, the classification algorithm achieved an accuracy of 95.6%-99.6%[73]. In 2016, Wimmer *et al*[74] used CNN to detect celiac disease in a set of 1661 images (986 images of normal mucosa, 675 images of celiac disease) with varying convolutional blocks. Their CNN achieved the best overall classification rate (90.3%) with 4 convolutional blocks[74]. Taking their CNN a step further, they combined the CNN with 4 convolutional blocks with SVM which increased overall classification rate by 6.7%[74]. While interesting, Gadermayr *et al*'s method requires human intervention and the paper's methodology is quite complicated[73], largely in part to the extensive number of systems tested. Wimmer *et al*[74] provided a simpler method that attained a good overall classification rate.

WIRELESS CAPSULE ENDOSCOPY

Celiac disease

Few studies have assessed the utility of AI in the detection of celiac disease using WCE. In 2017, Zhou *et al*[75] trained GoogLeNet, a DCNN, to identify celiac disease using clips obtained during WCE. Their DCCN achieved a sensitivity and specificity of 100% for identifying patients with celiac disease from 10 WCE videos (5 from patients with celiac disease, 5 from healthy controls)[75]. Similarly, Wang *et al*[76] used DL to diagnose celiac disease from WCE videos, however their CNN utilized a block-wise channel squeeze and excitations attenuation module, a newer architectural unit thought to better mimic human visual perception[76]. Their system attained an accuracy of 95.9%, sensitivity of 97.2% and specificity of 95.6% for diagnosing celiac disease.

Inflammatory bowel disease

WCE is often used in patients with inflammatory bowel disease (IBD) to detect small bowel ulcers and erosions. While computed tomography enterography and MRI have been used to detect areas of disease activity and inflammation along the gastrointestinal tract in patients with IBD, these imaging modalities can miss early or small lesions. While WCE can directly visualize lesions, endoscopists reviewing the video may miss lesions or mistakenly identify imaging artifacts as lesions. AI systems could help reduce these errors. Several studies have been published using AI in WCE to detect intestinal changes consistent with Crohn's disease[77-83].

To discriminate ulcers from normal mucosa in Crohn's disease, Charisis *et al*[78] proposed combining bidimensional ensemble empirical mode decomposition and differential lacunarity to pre-process images followed by classification using several ML algorithms and a multilayer neural network. Using a dataset consisting of 87 ulcer and 87 normal mucosa images, their CADe achieved accuracy 89.0%-95.4%, sensitivity 88.2%-98.8%, and specificity 84.2%-96.6%[78]. Subsequently, Charisis and Hadjileontiadis published a paper in 2016 combining hybrid adaptive filtering and differential lacunarity (HAF-DLac) to process images followed by SVM to detect Crohn's disease related lesions in WCE[79]. In a set of 800 WCE images, the HAF-DLac system achieved a sensitivity, specificity and accuracy of 95.2%, 92.4% and 93.8% respectively for detecting lesions[79]. Using a similar approach to Charisis *et al*[78], Kumar *et al*[80] used MPEG-7 edge, color and texture features to pre-process images followed by image classification using SVM to detect and classify lesions in patients with Crohn's disease. Their system, tested against 533 images (212 normal mucosa, 321 images with lesions), obtained an accuracy of 93.0%-93.8% for detecting lesions and an accuracy of 78.5% for classifying them based on severity.

With respect to deep learning, few groups have used deep learning algorithms in WCE to identify Crohn's disease related lesions. Recently, Ferreira *et al*[82] used a DCNN to identify erosions and ulcers in patients with Crohn's disease. Their DCNN achieved a sensitivity of 98.0%, specificity of 99.0%, accuracy of 98.8% and AUROC of 1.00. Interestingly, Klang *et al*[83] developed a DCNN to detect intestinal strictures. Overall, their DCNN achieved an accuracy of 93.5% ± 6.7% and AUC of 0.989 for detecting strictures.

Hookworm infections

Three studies have used artificial intelligence to detect hookworms using WCE. The first to publish on this topic was Wu *et al*[84] in 2016. Using SVM, they were able to create a system that achieved a specificity of 99.0% and accuracy of 98.4% for detecting hookworms in WCE[84]. However, the system's sensitivity was 11.1%. He *et al*[85] created a DCNN using a novel deep hookworm detection framework that modeled the tubular appearance of hookworms. Their DCNN had an accuracy of 88.5% for identifying hookworm[85]. Gan *et al*[86] performed a similar study, finding an AUC of 0.97 (95%CI 0.967-0.978), sensitivity of 92.2%, specificity of 91.1% and accuracy of 91.2% The concordant findings of these three studies suggest a possible utility of using AI to diagnose hookworm infections.

Intestinal bleeding

One of the most common reasons to perform WCE is to evaluate for gastrointestinal bleeding after prior endoscopic attempts have failed to localize a source. Since the implementation of WCE in clinical practice, many methods, notably AI, have been employed to improve the detection of gastrointestinal sources of bleeding.

Several studies have looked at using supervised learning to identify bleeding in WCE. In 2014, Sainju *et al*[87] used an ML algorithm to interpret color quantization images and determine if bleeding was present. One of their models achieved a sensitivity, specificity and accuracy of 96%, 90% and 93%, respectively[87]. Using SVM, Usman *et al*[88] achieved similar results - sensitivity, specificity and accuracy of 94%, 91% and 92% respectively.

More recently, several groups have created DCNNs to identify bleeding and sources of bleeding in WCE. In 2021, Ghosh *et al*[89] used a system comprised of two CNN systems (CNN-1, CNN-2) to classify WCE images as bleeding or non-bleeding and subsequently to identify sources of bleeding within the bleeding images. For classifying images as bleeding or non-bleeding, CNN-1 had a sensitivity, specificity, accuracy and AUC of 97.5%, 99.9%, 99.4% and 0.99[89]. For identifying sources of bleeding within the bleeding images, CNN-2 had an accuracy of 94.4% and intersection over union (IoU) of 90.7%[89].

In 2020, Tsuboi *et al*[90] published the first study to use DCNN to detect small bowel angioectasias from WCE images. In their test set which included 488 images of small bowel angioectasias and 10000 images of normal small bowel mucosa, their DCNN achieved an AUC of 0.99 with sensitivity and specificity of 98.8% and 98.4%[90]. Similarly, in 2021 Ribeiro *at al*[91] developed a DCNN to identify vascular lesions, categorizing them by bleeding risk according to Saurin's classification: P0 – no hemorrhagic potential, P1 – uncertain/intermediate hemorrhagic potential and red spots, and P2 – high hemorrhagic potential (angioectasias, varices). In their validation set, the DCNN had a sensitivity, specificity, accuracy and AUROC of 91.7%, 95.3%, 94.1% and 0.97 respectively for identifying P1 lesions [91]. Regarding P2 lesions, the network had a sensitivity, specificity, accuracy and AUROC of 94.1%, 95.1%, 94.8% and 0.98 respectively[91]. This group published a similar study in 2022 however now using

their DCNN to detect and differentiate mucosal erosions and ulcers based on bleeding potential[92]. Saurin's classification was again used to classify lesions, additionally labeling P1 lesions as mucosal erosions or small ulcers and P2 lesions as large ulcers (> 2 cm)[92]. The DCNN achieved an overall sensitivity of 90.8% ± 4.7%, specificity of 97.1% ± 1.7%, and accuracy of 93.4% ± 3.3% in their test set of 1226 images[92]. For the detection of mucosal erosions (P1), their DCNN achieved a sensitivity of 87.2%, specificity of 95.0% and accuracy of 93.3% with AUROC of 0.98 (95%CI 0.97-0.99)[92]. With respect to small ulcers (P1), their DCNN achieved a sensitivity of 86.4%, specificity of 96.9% and accuracy of 94.5% with AUROC of 0.99 (95%CI 0.97-1.00)[92]. Finally, with respect to large ulcers (P2), their DCNN achieved a sensitivity of 95.3%, specificity of 99.2% and AUROC of 1.00 (95%CI 0.98-1.00)[92]. A third study published by this group aimed to develop a DCNN to identify colonic lesions and luminal blood/hematic vestiges had similar findings. In their training set of 1801 images, the DCNN achieved an overall sensitivity, specificity and accuracy of 96.3%, 98.2%, and 97.6% respectively[93]. For detecting mucosal lesions, the DCNN achieved a sensitivity of 92.0%, specificity of 98.5% and AUROC of 0.99 (95%CI 0.98-1.00)[93]. For luminal blood/hematic vestiges, the DCNN achieved a sensitivity of 99.5%, specificity of 99.8% and AUROC of 1.00 (95%CI 0.99-1.00)[93].

Polyp and tumor detection

Gastrointestinal tumors can be difficult to discern from normal mucosa and thus pose a higher degree of diagnostic difficulty compared to other lesions on traditional WCE[94]. As such, developing an AI system to aid with the detection of these easy to miss lesions could be beneficial.

Several groups have developed ML systems to aid with detection. Using SVM, Li *et al*[95] were able to develop a system capable of detecting small bowel tumors with sensitivity, specificity and accuracy of 88.6%, 96.2% and 92.4%. Similarly, Liu *et al*[96] and Faghih Dinevari *et al*[97] used SVM to identify tumors in WCE, however they used different image pre-processing algorithms. Liu *et al*[96] used discrete curvelet transform to pre-process images prior to being classified by SVM. Their ML system achieved a sensitivity of 97.8% ± 0.5, specificity of 96.7% ± 0.4 and accuracy of 97.3% ± 0.5 for identifying small bowel tumors[96]. Faghih Dinevari *et al*[97] relied on discrete wavelet transform and singular value decomposition for image pre-processing prior to classification by SVM. Their system achieved a sensitivity of 94.0%, specificity of 93.0% and accuracy of 93.5% for identifying small bowel tumors[97]. Sundaram and Santhiyakumari built upon these methodologies, using a region of interest-based color histogram to enhance WCE images prior to being classified by two SVM algorithms: SVM1 and SVM2 [98]. SVM1 classified the WCE image as normal or abnormal. If SVM1 classified the image as abnormal, it was further classified by SVM2 as benign, malignant or normal[98]. The system attained an overall sensitivity of 96.0%, specificity of 95.4% and accuracy of 95.7% for small bowel tumor detection and classification[98].

With respect to DL methods, Blanes-Vidal *et al*[99] created a DCNN to autonomously detect and localize colorectal polyps. Their study included 255 patients who underwent WCE and standard colonoscopy for positive fecal immunochemical tests. Of the 255 patients, 131 had at least 1 polyp. The DCNN obtained a sensitivity of 97.1%, specificity of 93.3% and accuracy of 96.4% for detecting polyps in WCE[99]. Saraiva *et al*[100] and Mascarenhas *et al*[101] similarly used DCNNs to detect colonic polyps in WCE and obtained similar results to Blanes-Vidal *et al*[99-101]. Using an ANN, Constantinescu *et al*[102] created a DL system able to detect small bowel polyps with sensitivity of 93.6% and specificity of 91.4%. For gastric polyps and tumors, Xia *et al*[103] created a novel CNN – a region-based convolutional neural network (RCNN) – to evaluate magnetically controlled capsule endoscopy (MCE) images. Tested on 201365 MCE images obtained from 100 patients, the RCNN detected gastric polyps with sensitivity of 96.5%, specificity of 94.8%, accuracy of 94.9% and AUC of 0.898 (95%CI 0.84-0.96)[103]. For submucosal tumors, the RCNN achieved a sensitivity of 87.2%, specificity of 95.3%, accuracy of 95.2% and AUC of 0.88 (95%CI 0.81-0.96)[103]. Taking a different approach, Yuan and Meng used a novel deep learning method – stacked sparse autoencoder image manifold constraint – to identify intestinal polyps on WCE, finding an accuracy of 98.00% for poly detection[104]. However, sensitivity, specificity and AUC analyses were not reported.

COLONOSCOPY

Bowel preparation assessment

Inadequate bowel preparation, present in 15% to 35% of colonoscopies, is associated with lower rates of cecal intubation, lower adenoma detection rate (ADR), and higher rates of procedure-related adverse events[105,106]. For patients with inadequate bowel preparation, the United States Multi-Society Task Force of Colorectal Cancer (MSTF) which represents the American College of Gastroenterology, the American Gastroenterological Association and the American Society for Gastrointestinal Endoscopy (ASGE), and the European Society of Gastrointestinal Endoscopy recommend repeating a colonoscopy within 1 year[105,107-109]. In addition, the MSTF and ASGE recommend that endoscopists document bowel preparation quality at time of colonoscopy[108,109].

Despite these recommendations and variety of bowel preparation rating scales available, documentation of bowel preparation quality remains variable with studies reporting appropriate documentation in 20% to 88% of colonoscopies[110-112]. Few studies have been published regarding the use of DCNN to assist in the objective assessment of bowel preparation. The first group to do so, Zhou *et al*[113] in 2019, found that their DCNN (ENDOANGEL) was more accurate (93.3%) at grading the bowel preparation quality of still images than novice (< 1 year of experience performing colonoscopies; 75.91%), senior (1-3 years of experience performing colonoscopies; 74.36%) and expert (> 3 years of experience performing colonoscopies; 55.11%) endoscopists. When tested on colonoscopy videos, ENDOANGEL remained accurate at grading bowel preparation quality (89.04%)[113].

Building upon their experience with ENDOANGEL, Zhou *et al*[114] created a new system using two DCNNs: DCNN1 filtered unqualified frames while DCNN2 classified images by Boston Bowel Preparation Scale (BBPS) scores. The BBPS is a validated rating scale for assessing bowel preparation quality[115]. Colonic segments are assigned scores on a scale from 0 to 3. Colonic segments unable to be evaluated due to the presence of solid, unremovable stool are assigned a score of 0 whereas colonic segments that are able to be easily evaluated and contain minimal to no stool are assigned a score of 3 [115]. Zhou *et al*'s DCNN2 classified images into two categories: well-prepared (BBPS score 2-3) and poorly prepared (BPPS score 0-1)[114]. There was no difference between the dual DCNN system and study endoscopists when calculating the unqualified image portion (28.35% *vs* 29.58%, $P = 0.285$) and e-BBPS scores (7.81% *vs* 8.74%, $P = 0.088$). In addition, a strong inverse relationship between e-BBPS and ADR ($\rho = -0.976$, $P = 0.022$) was found.

Two other groups developed similar dual DCNN systems as Zhou *et al*[114] to calculate BBPS and obtained concordant findings[116,117]. Lee *et al*[116] tested their system on colonoscopy videos and found the system had an accuracy of 85.3% and AUC of 0.918 for detecting adequate bowel preparation. Using still images, Low *et al*'s system was able to accurately determine bowel preparation adequacy (98%) and subclassify by BBPS (91%)[117].

Using a different approach, Wang *et al*[118] used U-Net to create a DCNN to perform automatic segmentation of fecal matter from still images. Compared to images segmented by endoscopists, U-Net achieved an accuracy of 94.7%.

Inflammatory bowel disease

Colonoscopy is essential for the assessment of IBD as it allows for real-time evaluation of colonic inflammation[119,120]. Despite there being endoscopic scoring systems available to quantify disease activity, assessment is operator-dependent resulting in high interobserver variability[119-121]. Recent efforts have focused on using artificial intelligence to objectively grade colonic inflammation[121,122].

Several studies have investigated using DCNNs to classify images obtained from patients with ulcerative colitis (UC) by endoscopic inflammation scoring systems. The most commonly used endoscopic scoring system in these studies is the Mayo Endoscopic Score (MES). Physicians assign scores on a scale from 0 to 3 based on the absence or presence of erythema, friability, erosions, ulceration and bleeding[123]. A score of 0 indicates normal or inactive mucosa whereas a score of 3 indicates severe disease activity[123]. In 2018, Ozawa *et al*[121] published the first study to use a DCNN to classify still images obtained from patients with UC into MES 0 *vs* MES 1-3 and MES 0-1 *vs* MES 2-3. Their DCNN had an AUROC of 0.86 (95%CI 0.84-0.87) and AUROC 0.98 (95%CI 0.97-0.98) when differentiating MES 0 *vs* MES 1-3 and MES 0-1 *vs* MES 2-3 respectively[121]. Stidham *et al*[122] performed a similar study and found an AUROC of 0.966 (95%CI 0.967-0.972) for differentiating still images into MES 0-1 *vs* MES 2-3. Using a combined deep learning and machine learning system, Huang *et al*[124] were able to achieve an AUC of 0.938 with accuracy of 94.5% for identifying MES 0-1 *vs* MES 2-3 from still images. While the binary classification used in the aforementioned studies can differentiate remission/mucosal healing (MES 0-1) and active inflammation (MES 2-3), knowing exact MESs also has clinical significance[125,126]. Bhambhvani and Zamora created a DCNN to assign individual MESs to still images. The model achieved an AUC of 0.89, 0.86 and 0.96 for classifying images into MES 1, MES 2 and MES 3 respectively and achieved an average specificity of 85.7%, average sensitivity of 72.4% and overall accuracy of 77.2%[127].

In order to simulate how MES is performed in practice, several groups developed systems using DL to predict MES from colonoscopy videos. Yao *et al*'s DCNN had good agreement with MES scoring performed by gastroenterologists in their internal video test set (k = 0. 84; 95%CI 0.75-0.92), however their DCCN did not perform as well in the external video test set (k = 0.59; 95%CI 0.46-0.71)[128]. Gottlieb *et al*[129] reported similar findings to Yao *et al*[128], finding that their DCNN had good agreement with MES scoring performed by gastroenterologists (quadratic weighted kappa of 0.844; 95%CI 0.787–0.901). Gutierrez Becker *et al*[130] created a DL system designed to perform multiple binary tasks: discriminating MES < 1 *vs* MES ≥ 1, MES < 2 *vs* MES ≥ 2, and MES < 3 *vs* MES ≥ 3. For these tasks, their DL system attained an AUROC of 0.84, 0.85, and 0.85 respectively.

A group from Japan published several studies using AI on endoscopic images to predict histologic activity in patients with UC[131-134]. Their first study in 2016 used machine learning to predict persistent histologic inflammation[131]. Their system attained a sensitivity of 74% (95%CI 65%-81%), specificity of 97% (95%CI 95%-99%) and accuracy of 91% (95%CI 83%-95%) for predicting persistent histologic inflammation in still images[131]. Their following studies used a deep neural network labeled

DNUC (deep neural network for evaluation of UC) to identify endoscopic remission and histologic remission[132,134]. In still images, DNUC had a sensitivity of 93.3% (95%CI 92.2%–94.3%), specificity of 87.8% (95%CI 87.0%–88.4%) and diagnostic accuracy of 90.1% (95%CI 89.2%–90.9%) for determining endoscopic remission[132]. With respect to histologic remission, DNUC had a sensitivity of 92.4% (95%CI 91.5%–93.2%), specificity of 93.5% (95%CI 92.6%–94.3%) and diagnostic accuracy of 90.1% (92.9%; 95%CI 92.1%–93.7%)[132]. In colonoscopy videos, DNUC showed a sensitivity of 81.5% (95%CI 78.5%–83.9%) and specificity of 94.7% (95%CI 92.5%–96.4%) for endoscopic remission[134]. For histologic remission, DNUC had a sensitivity of 97.9% (95%CI 97.0%–98.5%) and specificity of 94.6% (95%CI 91.1%–96.9%) in colonoscopy videos[134].

To date, only one study has been published using an AI system to distinguish normal from inflamed colonic mucosa in Crohn's disease[135]. The group paired a DCNN with a long short-term memory (LSTM), a type of neural network that uses previous findings to interpret its current input, and confocal laser endomicroscopy. Their DCNN-LSTM system attained an accuracy of 95.3% and AUC of 0.98 for differentiating normal from inflamed mucosa[135].

Polyp detection

Colorectal cancer is the third most common malignancy and second leading cause of cancer-related mortality in the world[136]. While colonoscopy is the gold standard for detection and treatment of premalignant and malignant lesions, a substantial number of adenomas are missed[137,138]. As such, efforts have focused on using AI to improve ADR and decrease adenoma miss rate (AMR).

At present, numerous pilot, validation and prospective studies[139-161], randomized controlled studies[162-174], and systematic reviews and meta-analyses[175-183] have been published regarding the use of AI for the detection of colonic polyps. Furthermore, there are commercially available AI systems for both polyp detection and interpretation. With respect to the systematic reviews and meta-analyses published on this topic, AI-assisted colonoscopy has consistently been shown to have higher ADR, polyp detection rate (PDR) and adenoma per colonoscopy (APC) compared to standard colonoscopy [175-183]. Recently, several large, randomized controlled trials have been published supporting these findings. Shaukat *et al*[162] published their findings from their multicenter, randomized controlled trial comparing CADe colonoscopy to standard colonoscopy. Their study included 1359 patients: 677 randomized to standard colonoscopy, 682 to CADe colonoscopy. They found an increase in ADR (47.8% *vs* 43.9%; $P = 0.065$) and APC (1.05 *vs* 0.83; $P = 0.002$) in the CADe colonoscopy group. However, they also found a decrease in the overall sessile serrated lesions per colonoscopy rate (0.20 *vs* 0.28; $P = 0.042$) and sessile serrated lesion detection rate (12.6% *vs* 16.0%; $P = 0.092$) in the CADe colonoscopy group [162]. Brown *et al*[163] in their CADeT-CS Trial which was a multicenter, single-blind randomized tandem colonoscopy study comparing CADe colonoscopy to high-definition white light colonoscopy found similar increases in ADR (50.44% *vs* 43.64%; $P = 0.3091$) and APC (1.19 *vs* 0.90; $P = 0.0323$) in their patients who underwent CADe colonoscopy first[163]. Additionally, polyp miss rate (PMR) (20.70% *vs* 33.71%; $P = 0.0007$), AMR (20.12% *vs* 31.25%; $P = 0.0247$), and sessile serrated lesion miss rate (7.14% *vs* 42.11%; $P = 0.0482$) were lower in the CADe colonoscopy first group. In a similarly designed study to Brown *et al*[163], Kamba *et al*'s multicenter, randomized tandem colonoscopy study comparing CADe colonoscopy to standard colonoscopy found lower AMR (13.8% *vs* 26.7%; $P < 0.0001$), PMR (14.2% *vs* 40.6%; $P < 0.0001$), and sessile serrated lesion miss rate (13.0% *vs* 38.5%' $P = 0.03$) and higher ADR (64.5% *vs* 53.6%; $P = 0.036$) and PDR (69.8% *vs* 60.9%; $P = 0.084$) in patients who underwent CADe colonoscopy first[164]. Similar to Shaukat *et al*[162], the sessile serrated lesion detection rate was lower in the CADe colonoscopy first group compared to standard colonoscopy first (7.6% *vs* 8.1%; $P = 0.866$) [164]. Similar increases in ADR, APC and PDR were appreciated in randomize controlled trials by Xu *et al*[172], Liu *et al*[173], Repici *et al*[170], Gong *et al*[166], Wang *et al*[167], and Su *et al*[169] as well[166-172].

The majority of AI-assisted colonoscopy studies focus on adenoma detection. While these studies report sessile serrated lesion rates, it is often a secondary outcome despite sessile serrated lesions being the precursors of 15%-30% of all colorectal cancers[184]. Few studies have created AI systems optimized for dedicating sessile serrated lesions. Recently, Yoon *et al*[184] used a generative adversarial network (GAN) to generate endoscopic images of sessile serrated lesions which were used to train their DCNN with the hope of improving sessile serrated lesion detection. In the validation set which was comprised of 1141 images of polyps and 1000 normal images, their best performing GAN-DCNN model, GAN-aug2, achieved a sensitivity of 95.44% (95%CI 93.71%-97.17%), specificity of 90.10% (95%CI 88.38%-91.77%), accuracy of 92.95% (95%CI 91.86%-94.04%) and AUROC of 0.96 (95%CI 0.9547-0.9709)[184]. In a type-separated polyp validation dataset, the GAN-aug2 achieved a sensitivity of 95.24%, 19.1% higher than the DCNN without augmentation[184]. Given the small number of sessile serrate lesions present in the initial set, Yoon *et al*[184] collected an additional 130 images depicting 133 sessile serrated lesions to create an additional validation set titled SSL temporal validation dataset[184]. The GAN-aug2 continued to outperform the DCNN without augmentation (sensitivity 93.98% *vs* 84.21%). Nemoto *et al*[185] created a DCNN to differentiate (1) tubular adenomas from serrated lesions; and (2) serrated lesions from hyperplastic polyps. In their 215-image training set, the DCNN was able to differentiate tubular adenomas from sessile serrated lesions with sensitivity of 72% (95%CI 62%-81%), specificity 89% (95%CI 82%-94%), accuracy 82% (95%CI 77%-87%) and AUC 0.86 (95%CI 0.80-0.91). For differentiating sessile serrated lesions from hyperplastic polyps, the DCNN achieved a sensitivity of 17% (95%CI 7%-32%),

specificity 85% (95%CI 76%-92%), accuracy 63% (95%CI 54%-72%) and AUC 0.55 (95%CI 0.44-0.66)[185]. An overview of studies investigating the detection accuracy of CADe is provided in Table 3. An overview of studies investigating ADR and PDR using CADe is provided in Table 4.

FUTURE DIRECTIONS

Artificial intelligence is in its early stages for medicine, especially in gastroenterology and endoscopy. AI will help is in the areas of "augmentation" and "automation". Augmentation like what is happening with polyp detection and interpretation. Automation by eliminating electronic paperwork, such as the use of natural language processing for procedure documentation. Artificial intelligence systems have repeatedly been shown to be effective at identifying gastrointestinal lesions with high sensitivity, specificity and accuracy. While lesion detection is important, this is only the beginning of AI's utility in esophagogastroduodenoscopy, WCE and colonoscopy.

After refining their AI systems for lesion detection, several groups discussed in this narrative review were able to add additional functions to their AI systems. In BE, ESCC and gastric cancer, several AI systems were capable of predicting tumor invasion depth. Within IBD, AI systems were able to generate endoscopic disease severity scores. One group was able to train their CADe to recommend neoplasia biopsy sites in BE[14]. Additional efforts should be dedicated to developing these functions, testing them in real-time and having the AI system provide management recommendations when clinically appropriate.

Additional areas in need of future research are using AI systems to make histologic predictions, to assist with positioning of the endoscopic ultrasound (EUS) transducer and interpretation of EUS images, to detect biliary diseases and make therapeutic recommendations in endoscopic retrograde cholangiopancreatography (ERCP), and, in combination with endoscopic mechanical attachments, to improve colorectal cancer screening and surveillance. While endoscopists may perform optical biopsies of gastrointestinal lesions to predict histology and make real-time management decisions, these predictions are highly operator-dependent and often require expensive equipment that is not readily available. Thus, developing an AI system capable of performing objective optical biopsies, especially in WLE, would preserve the quality of histologic predictions, be cost effective, and avoid the risks associated with endoscopic biopsy and resection.

Similarly, EUS is highly operator-dependent, requiring endoscopists to place the transducer in specific positions to obtain adequate views of the hepatopancreatobiliary system. Research should focus on using AI systems to assist with appropriate transducer positioning and perform real-time EUS image analysis[186-194].

Presently, several clinical studies are actively recruiting patients to evaluate the utility of AI systems in ERCP. Of particular interest is the diagnosis and management of biliary diseases. Some groups are planning to use AI to classify bile duct lesions and provide biopsy site recommendations[195]. One group is planning to use an AI system in patients requiring biliary stents to assist with biliary stent choice and stent placement[196]. It will be interesting to see how AI performs in these tasks as successes could pave the way for future studies investigating the utility of AI systems to make real-time management recommendations.

While this narrative review focused on the use of AI in colonoscopy, of growing interest is the use of endoscopic mechanical attachments in colonoscopy to assist with polyp detection in colorectal cancer screening and surveillance. Independently, AI systems and endoscopic mechanical attachments are known to increase ADR and PDR. Few studies have investigated how combining AI with endoscopic mechanical attachments impacts ADR and PDR. Future research should examine the impact that combining these modalities has on ADR and PDR.

LIMITATIONS

While substantial advances have been made in AI, it is important to note that AI is not without limitations. In many of the studies discussed in this narrative review, the authors trained their AI systems using internally obtained images labeled by a single endoscopist. Thus, the AI is subject to the same operator biases and human error as the labeling endoscopist[1,197]. In addition, by using internally obtained data, several of these training sets may have inherent institutional or geographic biases resulting in AI systems that are biased and nongeneralizable[197]. As AI continues to progress, large datasets comprised of high-quality images should be created and used for training AI systems to reduce these biases[1].

With the implementation of AI in clinical practice, medical error accountability must also be addressed. While many of the AI systems discussed in this narrative review boast high detection accuracies, none are perfect. It is undeniable that errors in detection and diagnosis will arise when using these technologies. Regulatory bodies are needed to continually supervise these AI systems and oversee problems as they arise[198].

Table 3 Overview of findings from studies evaluating the detection accuracy of computer-aided detection for colonic polyps

Ref.	Country	Study design	Lesions	Training dataset	Test dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUROC
Komeda *et al*[139], 2017	Japan	Retrospective	Adenomas	1200 images	10 images	80	60	70	-
Misawa *et al*[140], 2018	Japan	Retrospective	Polyps	411 video clips	135 video clips	90	63.3	76.5	0.87
Wang *et al* [149], 2018	China, United States	Retrospective	Polyps	4495 images	Dataset A: 27113 images; Dataset C: 138 video clips; Dataset D: 54 full-length videos	Dataset A: 94.38; Dataset C: 91.64	Dataset A: 95.92; Dataset D: 95.4	-	Dataset A: 0.984
Horiuchi *et al*[154], 2019	Japan	Prospective	Diminutive polyps	-	[a]	80	95.3	91.5	-
Hassan *et al*[141], 2020	Italy, United States	Retrospective	Polyps	-	338 video clips	99.7	-	-	-
Guo *et al* [142], 2021	Japan	Retrospective	Polyps	1991 images	100 video clips; 15 full videos	87[b]	98.3[b]	-	-
Neumann *et al*[143], 2021	Germany	Retrospective[1]	Polyps	> 500 videos	240 polyps within full-length videos	100	0	-	-
Li *et al* [144], 2021	Singapore	Retrospective	Polyps	6038 images	2571 images	74.1	85.1	-	-
Livovsky *et al*[151], 2021	Israel	Ambispective	Polyps	3611 h of videos	1393 h of videos	97.1	0	-	-
Pfeifer *et al* [158], 2021	Germany, Italy, Netherlands	Retrospective	Polyps	10467 images	45 videos	90	80	-	0.92
Ahmad *et al*[145], 2022[2]	England	Prospective	Polyps	Dataset A: 58849 frames; Dataset B: 10993 videos and still images	Dataset C: 110985 frames; Dataset D: 8950 frames; Dataset E: 542484 frames	Dataset C: 100, 84.1; Dataset D&E: 98.9, 85.2	Dataset C: 79.6; Dataset D&E: 79.3%		
Hori *et al* [146], 2022	Japan	Prospective	Polyps	1456 images	600 images	97	97.7	97.3	-
Pacal *et al* [152], 2022	Turkey	Retrospective	Polyps	Used images from 3 publicly available datasets (SUN, PICCOLO, Etis-Larib) to create training and test datasets		91.04	-	-	-
Yoon *et al* [184], 2022	South Korea	Retrospective	SSL	4397 images	Validation Set 2106; SSL Temporal Validation set 133	95.44; 93.89	90.1	92.95	0.96
Nemoto *et al*[185], 2022	Japan	Retrospective	TA, SSL	1849 images	400 images	72	89	82	0.86
Lux *et al* [148], 2022	Germany	Retrospective	Polyps	506338 images	41 full-length videos	-	-	95.3	-

[a]Tested CADe in a cohort of 95 patients.
[b]Per-frame analysis from full-length video dataset.
[1]Presumed retrospective based on manuscript.
[2]Sensitivity is reported as per-polyp and per-frame respectively. Specificity is reported as per-frame.
AUROC: Area under the receiver operating characteristic; SSL: Sessile serrate lesion; TA: Tubular adenoma. All studies used a convolutional neural network to classify images.

CONCLUSION

In this narrative review, we provide an objective overview of the AI-related research being performed within esophagogastroduodenoscopy, WCE and colonoscopy. We attempted to be comprehensive by using several electronic databases including Embase, Ovid Medicine, and PubMed. However, it is possible that some publications pertinent to our narrative review were missed.

Undoubtedly, AI within esophagogastroduodenoscopy, WCE and colonoscopy is rapidly evolving, moving from retrospectively tested supervised learning algorithms to large, multicenter clinical trials using completely autonomous systems within the span of 10 years. The systems developed by these researchers show promise for detecting lesions, diagnosing conditions, and monitoring diseases. In fact, two of the computer aided detection systems discussed in this narrative review designed to aid with colorectal polyp detection were approved by the United States Food and Drug Administration in 2021 [171,199]. Thus, the question is no longer if but when will AI become integrated with clinical practice. Medical providers at all levels of training should prepare to incorporate artificial intelligence systems into routine practice.

Table 4 Overview of findings from studies evaluating computer-aided detection for adenoma detection rate and polyp detection rate

Ref.	Country	Study design	Patients (*n*)		PDR (%)			ADR (%)		
			CADe	SC	CADe	SC	*P* value	CADe	SC	*P* value
Wang *et al* [168], 2019	China, United States	Randomized	522	536	45.02	29.1	< 0.001	29.12	20.34	< 0.001
Becq *et al*[155], 2020	United States, Turkey, Costa Rica	Prospective	50[b]		82	62	Not reported	-	-	-
Gong *et al*[166], 2020	China	Randomized	355	349	47	34	0.0016	16	8	0.001
Liu *et al*[171], 2020	China, United States	Randomized	393	397	47.07	33.25	< 0.001	29.01	20.91	0.009
Liu *et al*[173], 2020	China	Prospective	508	518	43.65	27.81	< 0.001	39.1	23.89	< 0.001
Repici *et al* [170], 2020	Italy, Kuwait, United States, Germany	Randomized	341	344	-	-	-	54.8	40.4	< 0.001
Su *et al*[169], 2020	China	Randomized	308	315	38.3	25.4	0.001	28.9	16.5	< 0.001
Wang *et al* [156], 2020	China, United States	Prospective, Tandem[1]	184	185	65.59	55.14	0.099	42.39	35.68	0.186
Wang *et al* [167], 2020	China, United States	Randomized	484	478	52	37	< 0.0001	34	28	0.03
Kamba *et al* [164], 2021	Japan	Randomized, Tandem[2]	172	174	69.8	60.9	0.084	64.5	53.6	0.036
Luo *et al*[174], 2021	China	Randomized, Tandem[1]	72	78	38.7	34	< 0.001	-	-	-
Pfeifer *et al* [158], 2021	Germany, Italy, Netherlands	Prospective, Tandem[1]	42[b]		50	38	0.023	36	26	0.044
Shaukat *et al* [157], 2021	United States, England	Prospective	83	283	-	-	-	54.2	40.6	0.028
Shen *et al*[150], 2021	China	Ambispective	64	64	78.1	56.3	0.008	53.1	29.7	0.007
Xu *et al*[172], 2021	China	Randomized	1177	1175	38.8	36.2	0.183	-	-	-
Glissen Brown *et al*[163], 2022	China, United States	Randomized, Tandem[2]	113	110	70.8	65.45	0.3923	50.44	43.64	0.3091
Ishiyama *et al* [159], 2022	Japan, Norway	Prospective	918	918	59	52.1	0.003	26.4	19.9	0.001
Lux *et al*[148], 2022	Germany	Retrospective	41	-	-	-	-	-	41.5	-

Quan *et al*[153], 2022	United States	Prospective	300	300	-	-	-	43.7[a]; 66.7	37.8[a]; 59.72	0.37[a]; 0.35
Repici *et al* [165], 2022	Italy, Switzerland, United States, Germany	Randomized	330	330	-	-	-	53.3	44.5	0.017
Shaukat *et al* [162], 2022	United States	Randomized	682	677	64.4	61.2	0.242	47.8	43.9	0.065
Zippelius *et al* [160], 2022	Germany, United States	Prospective	150[b]		-	-	-	50.7	52	0.5

[a]Quan *et al*[153] reported results by indication, screening and surveillance respectively.
[b]The same patients were used to compare CADe versus standard colonoscopy; Becq *et al*[155] recorded 50 colonoscopy videos that were analyzed by CADe and reviewed by endoscopists separately; Pfeifer *et al*[158] performed standard colonoscopy followed by CADe-assisted colonoscopy in all 42 patients; Zippelius *et al*[160] had their CADe analyze their patients while the endoscopists performed their colonoscopies.
[1]Performed analyses using data obtained from whole process.
[2]Performed analyses using data obtained from first pass.
ADR: Adenoma detection rate; CADe: Computer-aided detection; PDR: Polyp detection rate; SC: Standard colonoscopy. All studies used a convolutional neural network to classify images.

FOOTNOTES

Author contributions: Galati JS, Gross SA contributed to manuscript concept and design; Galati JS, Duve RJ, O'Mara M contributed to obtaining and interpreting literary sources, drafting of manuscript; Galati JS, Duve RJ, O'Mara M, Gross SA contributed to revision of manuscript; All authors read and approved the final version of the manuscript.

Conflict-of-interest statement: All the authors declare that they have no conflict of interest.

Country/Territory of origin: United States

ORCID number: Jonathan S Galati 0000-0001-8075-3376; Seth A Gross 0000-0001-8942-865X.

S-Editor: Liu JH
L-Editor: A
P-Editor: Liu JH

REFERENCES

1 **Kröner PT**, Engels MM, Glicksberg BS, Johnson KW, Mzaik O, van Hooft JE, Wallace MB, El-Serag HB, Krittanawong C. Artificial intelligence in gastroenterology: A state-of-the-art review. *World J Gastroenterol* 2021; **27**: 6794-6824 [PMID: 34790008 DOI: 10.3748/wjg.v27.i40.6794]
2 **Kaul V**, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc* 2020; **92**: 807-812 [PMID: 32565184 DOI: 10.1016/j.gie.2020.06.040]
3 **Deo RC**. Machine Learning in Medicine. *Circulation* 2015; **132**: 1920-1930 [PMID: 26572668 DOI: 10.1161/CIRCULATIONAHA.115.001593]
4 **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]
5 **Christou CD**, Tsoulfas G. Challenges and opportunities in the application of artificial intelligence in gastroenterology and hepatology. *World J Gastroenterol* 2021; **27**: 6191-6223 [PMID: 34712027 DOI: 10.3748/wjg.v27.i37.6191]
6 **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]
7 **Hussein M**, González-Bueno Puyal J, Mountney P, Lovat LB, Haidry R. Role of artificial intelligence in the diagnosis of oesophageal neoplasia: 2020 an endoscopic odyssey. *World J Gastroenterol* 2020; **26**: 5784-5796 [PMID: 33132634 DOI: 10.3748/wjg.v26.i38.5784]
8 **Sharma P**. Barrett Esophagus: A Review. *JAMA* 2022; **328**: 663-671 [PMID: 35972481 DOI: 10.1001/jama.2022.13298]
9 **Bujanda DE**, Hachem C. Barrett's Esophagus. *Mo Med* 2018; **115**: 211-213 [PMID: 30228724]
10 **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]

11 **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]

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 **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]

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 **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]

16 **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]

17 **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]

18 **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]

19 **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]

20 **Hussein M**, González-Bueno Puyal J, Lines D, Sehgal V, Toth D, Ahmad OF, Kader R, Everson M, Lipman G, Fernandez-Sordo JO, Ragunath K, Esteban JM, Bisschops R, Banks M, Haefner M, Mountney P, Stoyanov D, Lovat LB, Haidry R. A new artificial intelligence system successfully detects and localises early neoplasia in Barrett's esophagus by using convolutional neural networks. *United European Gastroenterol J* 2022; **10**: 528-537 [PMID: 35521666 DOI: 10.1002/ueg2.12233]

21 **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]

22 **Ali S**, Bailey A, Ash S, Haghighat M; TGU Investigators, Leedham SJ, Lu X, East JE, Rittscher J, Braden B. A Pilot Study on Automatic Three-Dimensional Quantification of Barrett's Esophagus for Risk Stratification and Therapy Monitoring. *Gastroenterology* 2021; **161**: 865-878.e8 [PMID: 34116029 DOI: 10.1053/j.gastro.2021.05.059]

23 **Ebigbo A**, Mendel R, Rückert T, Schuster L, Probst A, Manzeneder J, Prinz F, Mende M, Steinbrück I, Faiss S, Rauber D, de Souza LA Jr, Papa JP, Deprez PH, Oyama T, Takahashi A, Seewald S, Sharma P, Byrne MF, Palm C, Messmann H. Endoscopic prediction of submucosal invasion in Barrett's cancer with the use of artificial intelligence: a pilot study. *Endoscopy* 2021; **53**: 878-883 [PMID: 33197942 DOI: 10.1055/a-1311-8570]

24 **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]

25 **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]

26 **Zhang Y**. Epidemiology of esophageal cancer. *World J Gastroenterol* 2013; **19**: 5598-5606 [PMID: 24039351 DOI: 10.3748/wjg.v19.i34.5598]

27 **Shin D**, Protano MA, Polydorides AD, Dawsey SM, Pierce MC, Kim MK, Schwarz RA, Quang T, Parikh N, Bhutani MS, Zhang F, Wang G, Xue L, Wang X, Xu H, Anandasabapathy S, Richards-Kortum RR. Quantitative analysis of high-resolution microendoscopic images for diagnosis of esophageal squamous cell carcinoma. *Clin Gastroenterol Hepatol* 2015; **13**: 272-279.e2 [PMID: 25066838 DOI: 10.1016/j.cgh.2014.07.030]

28 **Quang T**, Schwarz RA, Dawsey SM, Tan MC, Patel K, Yu X, Wang G, Zhang F, Xu H, Anandasabapathy S, Richards-Kortum R. A tablet-interfaced high-resolution microendoscope with automated image interpretation for real-time evaluation of esophageal squamous cell neoplasia. *Gastrointest Endosc* 2016; **84**: 834-841 [PMID: 27036635 DOI: 10.1016/j.gie.2016.03.1472]

29 **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]

30 **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]

31 **Liu W**, Yuan X, Guo L, Pan F, Wu C, Sun Z, Tian F, Yuan C, Zhang W, Bai S, Feng J, Hu Y, Hu B. Artificial Intelligence for Detecting and Delineating Margins of Early ESCC Under WLI Endoscopy. *Clin Transl Gastroenterol* 2022; **13**: e00433 [PMID: 35130184 DOI: 10.14309/ctg.0000000000000433]

32 **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]

33 **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]

34 **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]

35 **Li B**, Cai SL, Tan WM, Li JC, Yalikong A, Feng XS, Yu HH, Lu PX, Feng Z, Yao LQ, Zhou PH, Yan B, Zhong YS. Comparative study on artificial intelligence systems for detecting early esophageal squamous cell carcinoma between narrow-band and white-light imaging. *World J Gastroenterol* 2021; **27**: 281-293 [PMID: 33519142 DOI: 10.3748/wjg.v27.i3.281]

36 **Shiroma S**, Yoshio T, Kato Y, Horie Y, Namikawa K, Tokai Y, Yoshimizu S, Yoshizawa N, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Akazawa N, Akiyama J, Tada T, Fujisaki J. Ability of artificial intelligence to detect T1 esophageal squamous cell carcinoma from endoscopic videos and the effects of real-time assistance. *Sci Rep* 2021; **11**: 7759 [PMID: 33833355 DOI: 10.1038/s41598-021-87405-6]

37 **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]

38 **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]

39 **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]

40 **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]

41 **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]

42 **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]

43 **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]

44 **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]

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 **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]

47 **Miyaki R**, Yoshida S, Tanaka S, Kominami Y, Sanomura Y, Matsuo T, Oka S, Raytchev B, Tamaki T, Koide T, Kaneda K, Yoshihara M, Chayama K. Quantitative identification of mucosal gastric cancer under magnifying endoscopy with flexible spectral imaging color enhancement. *J Gastroenterol Hepatol* 2013; **28**: 841-847 [PMID: 23424994 DOI: 10.1111/jgh.12149]

48 **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]

49 **Tang D**, Wang L, Ling T, Lv Y, Ni M, Zhan Q, Fu Y, Zhuang D, Guo H, Dou X, Zhang W, Xu G, Zou X. Development and validation of a real-time artificial intelligence-assisted system for detecting early gastric cancer: A multicentre retrospective diagnostic study. *EBioMedicine* 2020; **62**: 103146 [PMID: 33254026 DOI: 10.1016/j.ebiom.2020.103146]

50 **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]

51 **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]

52 **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]

53 **Yuan XL**, Zhou Y, Liu W, Luo Q, Zeng XH, Yi Z, Hu B. Artificial intelligence for diagnosing gastric lesions under white-light endoscopy. *Surg Endosc* 2022; **36**: 9444-9453 [PMID: 35879572 DOI: 10.1007/s00464-022-09420-6]

54 **Guo L**, Gong H, Wang Q, Zhang Q, Tong H, Li J, Lei X, Xiao X, Li C, Jiang J, Hu B, Song J, Tang C, Huang Z. Detection of multiple lesions of gastrointestinal tract for endoscopy using artificial intelligence model: a pilot study. *Surg Endosc* 2021; **35**: 6532-6538 [PMID: 33185766 DOI: 10.1007/s00464-020-08150-x]

55 **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]

56 **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]

57 **Horiuchi Y**, Hirasawa T, Ishizuka N, Tokai Y, Namikawa K, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Performance of a computer-aided diagnosis system in diagnosing early gastric cancer using magnifying endoscopy videos with narrow-band imaging (with videos). *Gastrointest Endosc* 2020; **92**: 856-865.e1 [PMID: 32422155 DOI: 10.1016/j.gie.2020.04.079]

58 **Hu H**, Gong L, Dong D, Zhu L, Wang M, He J, Shu L, Cai Y, Cai S, Su W, Zhong Y, Li C, Zhu Y, Fang M, Zhong L, Yang X, Zhou P, Tian J. Identifying early gastric cancer under magnifying narrow-band images with deep learning: a multicenter study. *Gastrointest Endosc* 2021; **93**: 1333-1341.e3 [PMID: 33248070 DOI: 10.1016/j.gie.2020.11.014]

59 **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]

60 **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]

61 **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]

62 **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]

63 **Nagao S**, Tsuji Y, Sakaguchi Y, Takahashi Y, Minatsuki C, Niimi K, Yamashita H, Yamamichi N, Seto Y, Tada T, Koike K. Highly accurate artificial intelligence systems to predict the invasion depth of gastric cancer: efficacy of conventional white-light imaging, nonmagnifying narrow-band imaging, and indigo-carmine dye contrast imaging. *Gastrointest Endosc* 2020; **92**: 866-873.e1 [PMID: 32592776 DOI: 10.1016/j.gie.2020.06.047]

64 **Ku Y**, Ding H, Wang G. Efficient Synchronous Real-Time CADe for Multicategory Lesions in Gastroscopy by Using Multiclass Detection Model. *Biomed Res Int* 2022; **2022**: 8504149 [PMID: 36093395 DOI: 10.1155/2022/8504149]

65 **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]

66 **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]

67 **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]

68 **Itoh T**, Kawahira H, Nakashima H, Yata N. Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images. *Endosc Int Open* 2018; **6**: E139-E144 [PMID: 29399610 DOI: 10.1055/s-0043-120830]

69 **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]

70 **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]

71 **Nakashima H**, Kawahira H, Kawachi H, Sakaki N. Endoscopic three-categorical diagnosis of Helicobacter pylori infection using linked color imaging and deep learning: a single-center prospective study (with video). *Gastric Cancer* 2020; **23**: 1033-1040 [PMID: 32382973 DOI: 10.1007/s10120-020-01077-1]

72 **Stoleru CA**, Dulf EH, Ciobanu L. Automated detection of celiac disease using Machine Learning Algorithms. *Sci Rep* 2022; **12**: 4071 [PMID: 35260574 DOI: 10.1038/s41598-022-07199-z]

73 **Gadermayr M**, Kogler H, Karla M, Merhof D, Uhl A, Vécsei A. Computer-aided texture analysis combined with experts' knowledge: Improving endoscopic celiac disease diagnosis. *World J Gastroenterol* 2016; **22**: 7124-7134 [PMID: 27610022 DOI: 10.3748/wjg.v22.i31.7124]

74 **Wimmer G**, Hegenbart S, Vecsei A, Uhl A. Convolutional Neural Network Architectures for the Automated Diagnosis of Celiac Disease. Presented at: International Workshop on Computer-Assisted and Robotic Endoscopy. CARE 2016; 104–113 [DOI: 10.1007/978-3-319-54057-3_10]

75 **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]

76 **Wang X**, Qian H, Ciaccio EJ, Lewis SK, Bhagat G, Green PH, Xu S, Huang L, Gao R, Liu Y. Celiac disease diagnosis from videocapsule endoscopy images with residual learning and deep feature extraction. *Comput Methods Programs Biomed* 2020; **187**: 105236 [PMID: 31786452 DOI: 10.1016/j.cmpb.2019.105236]

77 **Yogapriya J**, Chandran V, Sumithra MG, Anitha P, Jenopaul P, Suresh Gnana Dhas C. Gastrointestinal Tract Disease Classification from Wireless Endoscopy Images Using Pretrained Deep Learning Model. *Comput Math Methods Med* 2021; **2021**: 5940433 [PMID: 34545292 DOI: 10.1155/2021/5940433]

78 **Charisis VS**, Hadjileontiadis LJ, Liatsos CN, Mavrogiannis CC, Sergiadis GD. Capsule endoscopy image analysis using texture information from various colour models. *Comput Methods Programs Biomed* 2012; **107**: 61-74 [PMID: 22056811 DOI: 10.1016/j.cmpb.2011.10.004]

79 **Charisis VS**, Hadjileontiadis LJ. Potential of hybrid adaptive filtering in inflammatory lesion detection from capsule endoscopy images. *World J Gastroenterol* 2016; **22**: 8641-8657 [PMID: 27818583 DOI: 10.3748/wjg.v22.i39.8641]

80 **Kumar R**, Zhao Q, Seshamani S, Mullin G, Hager G, Dassopoulos T. Assessment of Crohn's disease lesions in wireless capsule endoscopy images. *IEEE Trans Biomed Eng* 2012; **59**: 355-362 [PMID: 22020661 DOI: 10.1109/TBME.2011.2172438]

81 **Wei Z**, Wang W, Bradfield J, Li J, Cardinale C, Frackelton E, Kim C, Mentch F, Van Steen K, Visscher PM, Baldassano RN, Hakonarson H; International IBD Genetics Consortium. Large sample size, wide variant spectrum, and advanced machine-learning technique boost risk prediction for inflammatory bowel disease. *Am J Hum Genet* 2013; **92**: 1008-1012 [PMID: 23731541 DOI: 10.1016/j.ajhg.2013.05.002]

82 **Ferreira JPS**, de Mascarenhas Saraiva MJDQEC, Afonso JPL, Ribeiro TFC, Cardoso HMC, Ribeiro Andrade AP, de Mascarenhas Saraiva MNG, Parente MPL, Natal Jorge R, Lopes SIO, de Macedo GMG. Identification of Ulcers and Erosions by the Novel Pillcam™ Crohn's Capsule Using a Convolutional Neural Network: A Multicentre Pilot Study. *J Crohns Colitis* 2022; **16**: 169-172 [PMID: 34228113 DOI: 10.1093/ecco-jcc/jjab117]

83 **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]

84 **Wu X**, Chen H, Gan T, Chen J, Ngo CW, Peng Q. Automatic Hookworm Detection in Wireless Capsule Endoscopy Images. *IEEE Trans Med Imaging* 2016; **35**: 1741-1752 [PMID: 26886971 DOI: 10.1109/TMI.2016.2527736]

85 **He JY**, Wu X, Jiang YG, Peng Q, Jain R. Hookworm Detection in Wireless Capsule Endoscopy Images With Deep Learning. *IEEE Trans Image Process* 2018; **27**: 2379-2392 [PMID: 29470172 DOI: 10.1109/TIP.2018.2801119]

86 **Gan T**, Yang Y, Liu S, Zeng B, Yang J, Deng K, Wu J, Yang L. Automatic Detection of Small Intestinal Hookworms in Capsule Endoscopy Images Based on a Convolutional Neural Network. *Gastroenterol Res Pract* 2021; **2021**: 5682288 [PMID: 34868306 DOI: 10.1155/2021/5682288]

87 **Sainju S**, Bui FM, Wahid KA. Automated bleeding detection in capsule endoscopy videos using statistical features and region growing. *J Med Syst* 2014; **38**: 25 [PMID: 24696394 DOI: 10.1007/s10916-014-0025-1]

88 **Usman MA**, Satrya GB, Usman MR, Shin SY. Detection of small colon bleeding in wireless capsule endoscopy videos. *Comput Med Imaging Graph* 2016; **54**: 16-26 [PMID: 27793502 DOI: 10.1016/j.compmedimag.2016.09.005]

89 **Ghosh T**, Chakareski J. Deep Transfer Learning for Automated Intestinal Bleeding Detection in Capsule Endoscopy Imaging. *J Digit Imaging* 2021; **34**: 404-417 [PMID: 33728563 DOI: 10.1007/s10278-021-00428-3]

90 **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]

91 **Ribeiro T**, Saraiva MM, Ferreira JPS, Cardoso H, Afonso J, Andrade P, Parente M, Jorge RN, Macedo G. Artificial intelligence and capsule endoscopy: automatic detection of vascular lesions using a convolutional neural network. *Ann Gastroenterol* 2021; **34**: 820-828 [PMID: 34815648 DOI: 10.20524/aog.2021.0653]

92 **Afonso J**, Saraiva MM, Ferreira JPS, Cardoso H, Ribeiro T, Andrade P, Parente M, Jorge RN, Macedo G. Automated detection of ulcers and erosions in capsule endoscopy images using a convolutional neural network. *Med Biol Eng Comput* 2022; **60**: 719-725 [PMID: 35038118 DOI: 10.1007/s11517-021-02486-9]

93 **Mascarenhas M**, Ribeiro T, Afonso J, Ferreira JPS, Cardoso H, Andrade P, Parente MPL, Jorge RN, Mascarenhas Saraiva M, Macedo G. Deep learning and colon capsule endoscopy: automatic detection of blood and colonic mucosal lesions using a convolutional neural network. *Endosc Int Open* 2022; **10**: E171-E177 [PMID: 35186665 DOI: 10.1055/a-1675-1941]

94 **Lewis BS**, Eisen GM, Friedman S. A pooled analysis to evaluate results of capsule endoscopy trials. *Endoscopy* 2005; **37**: 960-965 [PMID: 16189768 DOI: 10.1055/s-2005-870353]

95 **Li B**, Meng MQ. Tumor recognition in wireless capsule endoscopy images using textural features and SVM-based feature selection. *IEEE Trans Inf Technol Biomed* 2012; **16**: 323-329 [PMID: 22287246 DOI: 10.1109/TITB.2012.2185807]

96 **Liu G**, Yan G, Kuang S, Wang Y. Detection of small bowel tumor based on multi-scale curvelet analysis and fractal technology in capsule endoscopy. *Comput Biol Med* 2016; **70**: 131-138 [PMID: 26829705 DOI: 10.1016/j.compbiomed.2016.01.021]

97 **Faghih Dinevari V**, Karimian Khosroshahi G, Zolfy Lighvan M. Singular Value Decomposition Based Features for Automatic Tumor Detection in Wireless Capsule Endoscopy Images. *Appl Bionics Biomech* 2016; **2016**: 3678913 [PMID: 27478364 DOI: 10.1155/2016/3678913]

98 **Shanmuga Sundaram P**, Santhiyakumari N. An Enhancement of Computer Aided Approach for Colon Cancer Detection in WCE Images Using ROI Based Color Histogram and SVM2. *J Med Syst* 2019; **43**: 29 [PMID: 30612188 DOI: 10.1007/s10916-018-1153-9]

99 **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]

100 **Saraiva MM**, Ferreira JPS, Cardoso H, Afonso J, Ribeiro T, Andrade P, Parente MPL, Jorge RN, Macedo G. Artificial intelligence and colon capsule endoscopy: development of an automated diagnostic system of protruding lesions in colon capsule endoscopy. *Tech Coloproctol* 2021; **25**: 1243-1248 [PMID: 34499277 DOI: 10.1007/s10151-021-02517-5]

101 **Mascarenhas M**, Afonso J, Ribeiro T, Cardoso H, Andrade P, Ferreira JPS, Saraiva MM, Macedo G. Performance of a Deep Learning System for Automatic Diagnosis of Protruding Lesions in Colon Capsule Endoscopy. *Diagnostics (Basel)* 2022; **12** [PMID: 35741255 DOI: 10.3390/diagnostics12061445]

102 **Constantinescu AF**, Ionescu M, Iovănescu VF, Ciurea ME, Ionescu AG, Streba CT, Bunescu MG, Rogoveanu I, Vere CC. A computer-aided diagnostic system for intestinal polyps identified by wireless capsule endoscopy. *Rom J Morphol Embryol* 2016; **57**: 979-984 [PMID: 28002513]

103 **Xia J**, Xia T, Pan J, Gao F, Wang S, Qian YY, Wang H, Zhao J, Jiang X, Zou WB, Wang YC, Zhou W, Li ZS, Liao Z. Use of artificial intelligence for detection of gastric lesions by magnetically controlled capsule endoscopy. *Gastrointest Endosc* 2021; **93**: 133-139.e4 [PMID: 32470426 DOI: 10.1016/j.gie.2020.05.027]

104 **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]

105 **Sharma P**, Burke CA, Johnson DA, Cash BD. The importance of colonoscopy bowel preparation for the detection of colorectal lesions and colorectal cancer prevention. *Endosc Int Open* 2020; **8**: E673-E683 [PMID: 32355887 DOI: 10.1055/a-1127-3144]

106 **Ahmad OF**. Deep learning for automated bowel preparation assessment during colonoscopy: time to embrace a new approach? *Lancet Digit Health* 2021; **3**: e685-e686 [PMID: 34538737 DOI: 10.1016/S2589-7500(21)00143-6]

107 **Hassan C**, East J, Radaelli F, Spada C, Benamouzig R, Bisschops R, Bretthauer M, Dekker E, Dinis-Ribeiro M, Ferlitsch M, Fuccio L, Awadie H, Gralnek I, Jover R, Kaminski MF, Pellisé M, Triantafyllou K, Vanella G, Mangas-Sanjuan C, Frazzoni L, Van Hooft JE, Dumonceau JM. Bowel preparation for colonoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2019. *Endoscopy* 2019; **51**: 775-794 [PMID: 31295746 DOI: 10.1055/a-0959-0505]

108 **Rex DK**, Boland CR, Dominitz JA, Giardiello FM, Johnson DA, Kaltenbach T, Levin TR, Lieberman D, Robertson DJ. Colorectal Cancer Screening: Recommendations for Physicians and Patients from the U.S. Multi-Society Task Force on Colorectal Cancer. *Am J Gastroenterol* 2017; **112**: 1016-1030 [PMID: 28555630 DOI: 10.1038/ajg.2017.174]

109 **ASGE Standards of Practice Committee**, Saltzman JR, Cash BD, Pasha SF, Early DS, Muthusamy VR, Khashab MA, Chathadi KV, Fanelli RD, Chandrasekhara V, Lightdale JR, Fonkalsrud L, Shergill AK, Hwang JH, Decker GA, Jue TL, Sharaf R, Fisher DA, Evans JA, Foley K, Shaukat A, Eloubeidi MA, Faulx AL, Wang A, Acosta RD. Bowel preparation before colonoscopy. *Gastrointest Endosc* 2015; **81**: 781-794 [PMID: 25595062 DOI: 10.1016/j.gie.2014.09.048]

110 **Hadlock SD**, Liu N, Bernstein M, Gould M, Rabeneck L, Ruco A, Sutradhar R, Tinmouth JM. The Quality of Colonoscopy Reporting in Usual Practice: Are Endoscopists Reporting Key Data Elements? *Can J Gastroenterol Hepatol* 2016; **2016**: 1929361 [PMID: 27579299 DOI: 10.1155/2016/1929361]

111 **Gavin DR**, Valori RM, Anderson JT, Donnelly MT, Williams JG, Swarbrick ET. The national colonoscopy audit: a nationwide assessment of the quality and safety of colonoscopy in the UK. *Gut* 2013; **62**: 242-249 [PMID: 22661458 DOI: 10.1136/gutjnl-2011-301848]

112 **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]

113 **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]

114 **Zhou W**, Yao L, Wu H, Zheng B, Hu S, Zhang L, Li X, He C, Wang Z, Li Y, Huang C, Guo M, Zhang X, Zhu Q, Wu L, Deng Y, Zhang J, Tan W, Li C, Zhang C, Gong R, Du H, Zhou J, Sharma P, Yu H. Multi-step validation of a deep learning-based system for the quantification of bowel preparation: a prospective, observational study. *Lancet Digit Health* 2021; **3**: e697-e706 [PMID: 34538736 DOI: 10.1016/S2589-7500(21)00109-6]

115 **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]

116 **Lee JY**, Calderwood AH, Karnes W, Requa J, Jacobson BC, Wallace MB. Artificial intelligence for the assessment of bowel preparation. *Gastrointest Endosc* 2022; **95**: 512-518.e1 [PMID: 34896100 DOI: 10.1016/j.gie.2021.11.041]

117 **Low DJ**, Hong Z, Jugnundan S, Mukherjee A, Grover SC. Automated Detection of Bowel Preparation Scoring and Adequacy With Deep Convolutional Neural Networks. *JCAG* 2022; **XX**: 1-5 [DOI: 10.1093/jcag/gwac013]

118 **Wang YP**, Jheng YC, Sung KY, Lin HE, Hsin IF, Chen PH, Chu YC, Lu D, Wang YJ, Hou MC, Lee FY, Lu CL. Use of U-Net Convolutional Neural Networks for Automated Segmentation of Fecal Material for Objective Evaluation of Bowel Preparation Quality in Colonoscopy. *Diagnostics (Basel)* 2022; **12** [PMID: 35328166 DOI: 10.3390/diagnostics12030613]

119 **Kohli A**, Holzwanger EA, Levy AN. Emerging use of artificial intelligence in inflammatory bowel disease. *World J Gastroenterol* 2020; **26**: 6923-6928 [PMID: 33311940 DOI: 10.3748/wjg.v26.i44.6923]

120 **Cohen-Mekelburg S**, Berry S, Stidham RW, Zhu J, Waljee AK. Clinical applications of artificial intelligence and

machine learning-based methods in inflammatory bowel disease. *J Gastroenterol Hepatol* 2021; **36**: 279-285 [PMID: 33624888 DOI: 10.1111/jgh.15405]

121 **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]

122 **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]

123 **Dave M**, Loftus EV Jr. Mucosal healing in inflammatory bowel disease-a true paradigm of success? *Gastroenterol Hepatol (N Y)* 2012; **8**: 29-38 [PMID: 22347830]

124 **Huang TY**, Zhan SQ, Chen PJ, Yang CW, Lu HH. Accurate diagnosis of endoscopic mucosal healing in ulcerative colitis using deep learning and machine learning. *J Chin Med Assoc* 2021; **84**: 678-681 [PMID: 34050105 DOI: 10.1097/JCMA.0000000000000559]

125 **Barreiro-de Acosta M**, Vallejo N, de la Iglesia D, Uribarri L, Bastón I, Ferreiro-Iglesias R, Lorenzo A, Domínguez-Muñoz JE. Evaluation of the Risk of Relapse in Ulcerative Colitis According to the Degree of Mucosal Healing (Mayo 0 vs 1): A Longitudinal Cohort Study. *J Crohns Colitis* 2016; **10**: 13-19 [PMID: 26351390 DOI: 10.1093/ecco-jcc/jjv158]

126 **Osterman MT**, Scott FI, Fogt FF, Gilroy ED, Parrott S, Galanko J, Cross R, Moss A, Herfarth HH, Higgins PDR. Endoscopic and Histological Assessment, Correlation, and Relapse in Clinically Quiescent Ulcerative Colitis (MARQUEE). *Inflamm Bowel Dis* 2021; **27**: 207-214 [PMID: 32170946 DOI: 10.1093/ibd/izaa048]

127 **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]

128 **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.e1 [PMID: 32810479 DOI: 10.1016/j.gie.2020.08.011]

129 **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]

130 **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]

131 **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]

132 **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]

133 **Takenaka K**, Ohtsuka K, Fujii T, Oshima S, Okamoto R, Watanabe M. Deep Neural Network Accurately Predicts Prognosis of Ulcerative Colitis Using Endoscopic Images. *Gastroenterology* 2021; **160**: 2175-2177.e3 [PMID: 33485853 DOI: 10.1053/j.gastro.2021.01.210]

134 **Takenaka K**, Fujii T, Kawamoto A, Suzuki K, Shimizu H, Maeyashiki C, Yamaji O, Motobayashi M, Igarashi A, Hanazawa R, Hibiya S, Nagahori M, Saito E, Okamoto R, Ohtsuka K, Watanabe M. Deep neural network for video colonoscopy of ulcerative colitis: a cross-sectional study. *Lancet Gastroenterol Hepatol* 2022; **7**: 230-237 [PMID: 34856196 DOI: 10.1016/S2468-1253(21)00372-1]

135 **Udristoiu AL**, Stefanescu D, Gruionu G, Gruionu LG, Iacob AV, Karstensen JG, Vilman P, Saftoiu A. Deep Learning Algorithm for the Confirmation of Mucosal Healing in Crohn's Disease, Based on Confocal Laser Endomicroscopy Images. *J Gastrointestin Liver Dis* 2021; **30**: 59-65 [PMID: 33723558 DOI: 10.15403/jgld-3212]

136 **Xi Y**, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl Oncol* 2021; **14**: 101174 [PMID: 34243011 DOI: 10.1016/j.tranon.2021.101174]

137 **Roselló S**, Simón S, Cervantes A. Programmed colorectal cancer screening decreases incidence and mortality. *Transl Gastroenterol Hepatol* 2019; **4**: 84 [PMID: 32039289 DOI: 10.21037/tgh.2019.12.13]

138 **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]

139 **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]

140 **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]

141 **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]

142 **Guo Z**, Nemoto D, Zhu X, Li Q, Aizawa M, Utano K, Isohata N, Endo S, Kawarai Lefor A, Togashi K. Polyp detection algorithm can detect small polyps: Ex vivo reading test compared with endoscopists. *Dig Endosc* 2021; **33**: 162-169 [PMID: 32173917 DOI: 10.1111/den.13670]

143 **Neumann H**, Kreft A, Sivanathan V, Rahman F, Galle PR. Evaluation of novel LCI CAD EYE system for real time detection of colon polyps. *PLoS One* 2021; **16**: e0255955 [PMID: 34437563 DOI: 10.1371/journal.pone.0255955]

144 **Li JW**, Chia T, Fock KM, Chong KW, Wong YJ, Ang TL. Artificial intelligence and polyp detection in colonoscopy: Use of a single neural network to achieve rapid polyp localization for clinical use. *J Gastroenterol Hepatol* 2021; **36**: 3298-3307 [PMID: 34327729 DOI: 10.1111/jgh.15642]

145 **Ahmad OF**, González-Bueno Puyal J, Brandao P, Kader R, Abbasi F, Hussein M, Haidry RJ, Toth D, Mountney P, Seward E, Vega R, Stoyanov D, Lovat LB. Performance of artificial intelligence for detection of subtle and advanced colorectal neoplasia. *Dig Endosc* 2022; **34**: 862-869 [PMID: 34748665 DOI: 10.1111/den.14187]

146 **Hori K**, Ikematsu H, Yamamoto Y, Matsuzaki H, Takeshita N, Shinmura K, Yoda Y, Kiuchi T, Takemoto S, Yokota H, Yano T. Detecting colon polyps in endoscopic images using artificial intelligence constructed with automated collection of annotated images from an endoscopy reporting system. *Dig Endosc* 2022; **34**: 1021-1029 [PMID: 34748658 DOI: 10.1111/den.14185]

147 **Brand M**, Troya J, Krenzer A, Saßmannshausen Z, Zoller WG, Meining A, Lux TJ, Hann A. Development and evaluation of a deep learning model to improve the usability of polyp detection systems during interventions. *United European Gastroenterol J* 2022; **10**: 477-484 [PMID: 35511456 DOI: 10.1002/ueg2.12235]

148 **Lux TJ**, Banck M, Saßmannshausen Z, Troya J, Krenzer A, Fitting D, Sudarevic B, Zoller WG, Puppe F, Meining A, Hann A. Pilot study of a new freely available computer-aided polyp detection system in clinical practice. *Int J Colorectal Dis* 2022; **37**: 1349-1354 [PMID: 35543874 DOI: 10.1007/s00384-022-04178-8]

149 **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]

150 **Shen P**, Li WZ, Li JX, Pei ZC, Luo YX, Mu JB, Li W, Wang XM. Real-time use of a computer-aided system for polyp detection during colonoscopy, an ambispective study. *J Dig Dis* 2021; **22**: 256-262 [PMID: 33742774 DOI: 10.1111/1751-2980.12985]

151 **Livovsky DM**, Veikherman D, Golany T, Aides A, Dashinsky V, Rabani N, Ben Shimol D, Blau Y, Katzir L, Shimshoni I, Liu Y, Segol O, Goldin E, Corrado G, Lachter J, Matias Y, Rivlin E, Freedman D. Detection of elusive polyps using a large-scale artificial intelligence system (with videos). *Gastrointest Endosc* 2021; **94**: 1099-1109.e10 [PMID: 34216598 DOI: 10.1016/j.gie.2021.06.021]

152 **Pacal I**, Karaman A, Karaboga D, Akay B, Basturk A, Nalbantoglu U, Coskun S. An efficient real-time colonic polyp detection with YOLO algorithms trained by using negative samples and large datasets. *Comput Biol Med* 2022; **141**: 105031 [PMID: 34802713 DOI: 10.1016/j.compbiomed.2021.105031]

153 **Quan SY**, Wei MT, Lee J, Mohi-Ud-Din R, Mostaghim R, Sachdev R, Siegel D, Friedlander Y, Friedland S. Clinical evaluation of a real-time artificial intelligence-based polyp detection system: a US multi-center pilot study. *Sci Rep* 2022; **12**: 6598 [PMID: 35449442 DOI: 10.1038/s41598-022-10597-y]

154 **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]

155 **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]

156 **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]

157 **Shaukat A**, Colucci D, Erisson L, Phillips S, Ng J, Iglesias JE, Saltzman JR, Somers S, Brugge W. Improvement in adenoma detection using a novel artificial intelligence-aided polyp detection device. *Endosc Int Open* 2021; **9**: E263-E270 [PMID: 33553591 DOI: 10.1055/a-1321-1317]

158 **Pfeifer L**, Neufert C, Leppkes M, Waldner MJ, Häfner M, Beyer A, Hoffman A, Siersema PD, Neurath MF, Rath T. Computer-aided detection of colorectal polyps using a newly generated deep convolutional neural network: from development to first clinical experience. *Eur J Gastroenterol Hepatol* 2021; **33**: e662-e669 [PMID: 34034272 DOI: 10.1097/MEG.0000000000002209]

159 **Ishiyama M**, Kudo SE, Misawa M, Mori Y, Maeda Y, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Miyachi H, Ishida F, Itoh H, Oda M, Mori K. Impact of the clinical use of artificial intelligence-assisted neoplasia detection for colonoscopy: a large-scale prospective, propensity score-matched study (with video). *Gastrointest Endosc* 2022; **95**: 155-163 [PMID: 34352255 DOI: 10.1016/j.gie.2021.07.022]

160 **Zippelius C**, Alqahtani SA, Schedel J, Brookman-Amissah D, Muehlenberg K, Federle C, Salzberger A, Schorr W, Pech O. Diagnostic accuracy of a novel artificial intelligence system for adenoma detection in daily practice: a prospective nonrandomized comparative study. *Endoscopy* 2022; **54**: 465-472 [PMID: 34293812 DOI: 10.1055/a-1556-5984]

161 **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]

162 **Shaukat A**, Lichtenstein DR, Somers SC, Chung DC, Perdue DG, Gopal M, Colucci DR, Phillips SA, Marka NA, Church TR, Brugge WR; SKOUT™ Registration Study Team. Computer-Aided Detection Improves Adenomas per Colonoscopy for Screening and Surveillance Colonoscopy: A Randomized Trial. *Gastroenterology* 2022; **163**: 732-741 [PMID: 35643173 DOI: 10.1053/j.gastro.2022.05.028]

163 **Glissen Brown JR**, Mansour NM, Wang P, Chuchuca MA, Minchenberg SB, Chandnani M, Liu L, Gross SA, Sengupta N, Berzin TM. Deep Learning Computer-aided Polyp Detection Reduces Adenoma Miss Rate: A United States Multi-center Randomized Tandem Colonoscopy Study (CADeT-CS Trial). *Clin Gastroenterol Hepatol* 2022; **20**: 1499-1507.e4

[PMID: 34530161 DOI: 10.1016/j.cgh.2021.09.009]

164 **Kamba S**, Tamai N, Saitoh I, Matsui H, Horiuchi H, Kobayashi M, Sakamoto T, Ego M, Fukuda A, Tonouchi A, Shimahara Y, Nishikawa M, Nishino H, Saito Y, Sumiyama K. Reducing adenoma miss rate of colonoscopy assisted by artificial intelligence: a multicenter randomized controlled trial. *J Gastroenterol* 2021; **56**: 746-757 [PMID: 34218329 DOI: 10.1007/s00535-021-01808-w]

165 **Repici A**, Spadaccini M, Antonelli G, Correale L, Maselli R, Galtieri PA, Pellegatta G, Capogreco A, Milluzzo SM, Lollo G, Di Paolo D, Badalamenti M, Ferrara E, Fugazza A, Carrara S, Anderloni A, Rondonotti E, Amato A, De Gottardi A, Spada C, Radaelli F, Savevski V, Wallace MB, Sharma P, Rösch T, Hassan C. Artificial intelligence and colonoscopy experience: lessons from two randomised trials. *Gut* 2022; **71**: 757-765 [PMID: 34187845 DOI: 10.1136/gutjnl-2021-324471]

166 **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]

167 **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]

168 **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]

169 **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]

170 **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]

171 **Liu P**, Wang P, Glissen Brown JR, Berzin TM, Zhou G, Liu W, Xiao X, Chen Z, Zhang Z, Zhou C, Lei L, Xiong F, Li L, Liu X. The single-monitor trial: an embedded CADe system increased adenoma detection during colonoscopy: a prospective randomized study. *Therap Adv Gastroenterol* 2020; **13**: 1756284820979165 [PMID: 33403003 DOI: 10.1177/1756284820979165]

172 **Xu L**, He X, Zhou J, Zhang J, Mao X, Ye G, Chen Q, Xu F, Sang J, Wang J, Ding Y, Li Y, Yu C. Artificial intelligence-assisted colonoscopy: A prospective, multicenter, randomized controlled trial of polyp detection. *Cancer Med* 2021; **10**: 7184-7193 [PMID: 34477306 DOI: 10.1002/cam4.4261]

173 **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]

174 **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* 2021; **25**: 2011-2018 [PMID: 32968933 DOI: 10.1007/s11605-020-04802-4]

175 **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]

176 **Aziz M**, Fatima R, Dong C, Lee-Smith W, Nawras A. The impact of deep convolutional neural network-based artificial intelligence on colonoscopy outcomes: A systematic review with meta-analysis. *J Gastroenterol Hepatol* 2020; **35**: 1676-1683 [PMID: 32267558 DOI: 10.1111/jgh.15070]

177 **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]

178 **Li J**, Lu J, Yan J, Tan Y, Liu D. Artificial intelligence can increase the detection rate of colorectal polyps and adenomas: a systematic review and meta-analysis. *Eur J Gastroenterol Hepatol* 2021; **33**: 1041-1048 [PMID: 32804846 DOI: 10.1097/MEG.0000000000001906]

179 **Zhang Y**, Zhang X, Wu Q, Gu C, Wang Z. Artificial Intelligence-Aided Colonoscopy for Polyp Detection: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. *J Laparoendosc Adv Surg Tech A* 2021; **31**: 1143-1149 [PMID: 33524298 DOI: 10.1089/lap.2020.0777]

180 **Deliwala SS**, Hamid K, Barbarawi M, Lakshman H, Zayed Y, Kandel P, Malladi S, Singh A, Bachuwa G, Gurvits GE, Chawla S. Artificial intelligence (AI) real-time detection vs. routine colonoscopy for colorectal neoplasia: a meta-analysis and trial sequential analysis. *Int J Colorectal Dis* 2021; **36**: 2291-2303 [PMID: 33934173 DOI: 10.1007/s00384-021-03929-3]

181 **Nazarian S**, Glover B, Ashrafian H, Darzi A, Teare J. Diagnostic Accuracy of Artificial Intelligence and Computer-Aided Diagnosis for the Detection and Characterization of Colorectal Polyps: Systematic Review and Meta-analysis. *J Med Internet Res* 2021; **23**: e27370 [PMID: 34259645 DOI: 10.2196/27370]

182 **Xu Y**, Ding W, Wang Y, Tan Y, Xi C, Ye N, Wu D, Xu X. Comparison of diagnostic performance between convolutional neural networks and human endoscopists for diagnosis of colorectal polyp: A systematic review and meta-analysis. *PLoS One* 2021; **16**: e0246892 [PMID: 33592048 DOI: 10.1371/journal.pone.0246892]

183 **Bang CS**, Lee JJ, Baik GH. Computer-Aided Diagnosis of Diminutive Colorectal Polyps in Endoscopic Images:

Systematic Review and Meta-analysis of Diagnostic Test Accuracy. *J Med Internet Res* 2021; **23**: e29682 [PMID: 34432643 DOI: 10.2196/29682]

184 **Yoon D**, Kong HJ, Kim BS, Cho WS, Lee JC, Cho M, Lim MH, Yang SY, Lim SH, Lee J, Song JH, Chung GE, Choi JM, Kang HY, Bae JH, Kim S. Colonoscopic image synthesis with generative adversarial network for enhanced detection of sessile serrated lesions using convolutional neural network. *Sci Rep* 2022; **12**: 261 [PMID: 34997124 DOI: 10.1038/s41598-021-04247-y]

185 **Nemoto D**, Guo Z, Peng B, Zhang R, Nakajima Y, Hayashi Y, Yamashina T, Aizawa M, Utano K, Lefor AK, Zhu X, Togashi K. Computer-aided diagnosis of serrated colorectal lesions using non-magnified white-light endoscopic images. *Int J Colorectal Dis* 2022; **37**: 1875-1884 [PMID: 35861862 DOI: 10.1007/s00384-022-04210-x]

186 **Yao L**, Zhang J, Liu J, Zhu L, Ding X, Chen D, Wu H, Lu Z, Zhou W, Zhang L, Xu B, Hu S, Zheng B, Yang Y, Yu H. A deep learning-based system for bile duct annotation and station recognition in linear endoscopic ultrasound. *EBioMedicine* 2021; **65**: 103238 [PMID: 33639404 DOI: 10.1016/j.ebiom.2021.103238]

187 **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]

188 **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]

189 **Goyal H**, Sherazi SAA, Gupta S, Perisetti A, Achebe I, Ali A, Tharian B, Thosani N, Sharma NR. Application of artificial intelligence in diagnosis of pancreatic malignancies by endoscopic ultrasound: a systemic review. *Therap Adv Gastroenterol* 2022; **15**: 17562848221093873 [PMID: 35509425 DOI: 10.1177/17562848221093873]

190 **Dumitrescu EA**, Ungureanu BS, Cazacu IM, Florescu LM, Streba L, Croitoru VM, Sur D, Croitoru A, Turcu-Stiolica A, Lungulescu CV. Diagnostic Value of Artificial Intelligence-Assisted Endoscopic Ultrasound for Pancreatic Cancer: A Systematic Review and Meta-Analysis. *Diagnostics (Basel)* 2022; **12** [PMID: 35204400 DOI: 10.3390/diagnostics12020309]

191 **Minoda Y**, Ihara E, Komori K, Ogino H, Otsuka Y, Chinen T, Tsuda Y, Ando K, Yamamoto H, Ogawa Y. Efficacy of endoscopic ultrasound with artificial intelligence for the diagnosis of gastrointestinal stromal tumors. *J Gastroenterol* 2020; **55**: 1119-1126 [PMID: 32918102 DOI: 10.1007/s00535-020-01725-4]

192 **Jang SI**, Kim YJ, Kim EJ, Kang H, Shon SJ, Seol YJ, Lee DK, Kim KG, Cho JH. Diagnostic performance of endoscopic ultrasound-artificial intelligence using deep learning analysis of gallbladder polypoid lesions. *J Gastroenterol Hepatol* 2021; **36**: 3548-3555 [PMID: 34431545 DOI: 10.1111/jgh.15673]

193 **Zhang B**, Zhu F, Li P, Zhu J. Artificial intelligence-assisted endoscopic ultrasound in the diagnosis of gastrointestinal stromal tumors: a meta-analysis. *Surg Endosc* 2022 [PMID: 36100781 DOI: 10.1007/s00464-022-09597-w]

194 **Ye XH**, Zhao LL, Wang L. Diagnostic accuracy of endoscopic ultrasound with artificial intelligence for gastrointestinal stromal tumors: A meta-analysis. *J Dig Dis* 2022; **23**: 253-261 [PMID: 35793389 DOI: 10.1111/1751-2980.13110]

195 **Marya NB**, Powers PD, Petersen BT, Law R, Storm A, Abusaleh RR, Rau P, Stead C, Levy MJ, Martin J, Vargas EJ, Abu Dayyeh BK, Chandrasekhara V. Identification of patients with malignant biliary strictures using a cholangioscopy-based deep learning artificial intelligence (with video). *Gastrointest Endosc* 2022 [PMID: 36007584 DOI: 10.1016/j.gie.2022.08.021]

196 **Huang L**, Lu X, Huang X, Zou X, Wu L, Zhou Z, Wu D, Tang D, Chen D, Wan X, Zhu Z, Deng T, Shen L, Liu J, Zhu Y, Gong D, Zhong Y, Liu F, Yu H. Intelligent difficulty scoring and assistance system for endoscopic extraction of common bile duct stones based on deep learning: multicenter study. *Endoscopy* 2021; **53**: 491-498 [PMID: 32838430 DOI: 10.1055/a-1244-5698]

197 **Uche-Anya E**, Anyane-Yeboa A, Berzin TM, Ghassemi M, May FP. Artificial intelligence in gastroenterology and hepatology: how to advance clinical practice while ensuring health equity. *Gut* 2022; **71**: 1909-1915 [PMID: 35688612 DOI: 10.1136/gutjnl-2021-326271]

198 **Stidham RW**. Artificial Intelligence for Understanding Imaging, Text, and Data in Gastroenterology. *Gastroenterol Hepatol (N Y)* 2020; **16**: 341-349 [PMID: 34035738]

199 **Spadaccini M**, Marco A, Franchellucci G, Sharma P, Hassan C, Repici A. Discovering the first US FDA-approved computer-aided polyp detection system. *Future Oncol* 2022; **18**: 1405-1412 [PMID: 35081745 DOI: 10.2217/fon-2021-1135]

Artificial Intelligence in Gastroenterology

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

Artif Intell Gastroenterol 2022 December 28; 3(5): 142-162

DOI: 10.35712/aig.v3.i5.142

ISSN 2644-3236 (online)

REVIEW

Artificial intelligence applications in predicting the behavior of gastrointestinal cancers in pathology

Aysen Yavuz, Anil Alpsoy, Elif Ocak Gedik, Mennan Yigitcan Celik, Cumhur Ibrahim Bassorgun, Betul Unal, Gulsum Ozlem Elpek

Specialty type: Pathology

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): B
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Tsoulfas G, Greece; Wan XH, China

Received: October 16, 2022
Peer-review started: October 16, 2022
First decision: November 15, 2022
Revised: November 25, 2022
Accepted: December 13, 2022
Article in press: December 13, 2022
Published online: December 28, 2022

Aysen Yavuz, Anil Alpsoy, Elif Ocak Gedik, Mennan Yigitcan Celik, Cumhur Ibrahim Bassorgun, Betul Unal, Gulsum Ozlem Elpek, Department of Pathology, Akdeniz University Medical School, Antalya 07070, Turkey

Corresponding author: Gulsum Ozlem Elpek, MD, Professor, Department of Pathology, Akdeniz University Medical School, Dumlupinar bulvarı, Antalya 07070, Turkey. elpek@akdeniz.edu.tr

Abstract

Recent research has provided a wealth of data supporting the application of artificial intelligence (AI)-based applications in routine pathology practice. Indeed, it is clear that these methods can significantly support an accurate and rapid diagnosis by eliminating errors, increasing reliability, and improving workflow. In addition, the effectiveness of AI in the pathological evaluation of prognostic parameters associated with behavior, course, and treatment in many types of tumors has also been noted. Regarding gastrointestinal system (GIS) cancers, the contribution of AI methods to pathological diagnosis has been investigated in many studies. On the other hand, studies focusing on AI applications in evaluating parameters to determine tumor behavior are relatively few. For this purpose, the potential of AI models has been studied over a broad spectrum, from tumor subtyping to the identification of new digital biomarkers. The capacity of AI to infer genetic alterations of cancer tissues from digital slides has been demonstrated. Although current data suggest the merit of AI-based approaches in assessing tumor behavior in GIS cancers, a wide range of challenges still need to be solved, from laboratory infrastructure to improving the robustness of algorithms, before incorporating AI applications into real-life GIS pathology practice. This review aims to present data from AI applications in evaluating pathological parameters related to the behavior of GIS cancer with an overview of the opportunities and challenges encountered in implementing AI in pathology.

Key Words: Digital pathology; Colorectal cancer; Gastric cancer; Machine learning; Deep learning; Prognosis

Core Tip: This review outlines the potential of artificial intelligence applications for evaluating pathological parameters related to the behavior of gastrointestinal cancers. The role of these methods in determining the behavior of esophageal cancers remains to be investigated. On the other hand, the results are promising, supporting that these models can assist in the determination of conventional pathological parameters and perform molecular subtyping in gastric and colorectal cancers. Furthermore, these applications encourage digital prognostic biomarker discovery by revealing predictions that are impossible when using traditional visual methods. However, further studies are needed to overcome the obstacles to implementing these applications into pathology practice.

Citation: Yavuz A, Alpsoy A, Gedik EO, Celik MY, Bassorgun CI, Unal B, Elpek GO. Artificial intelligence applications in predicting the behavior of gastrointestinal cancers in pathology. *Artif Intell Gastroenterol* 2022; 3(5): 142-162
URL: https://www.wjgnet.com/2644-3236/full/v3/i5/142.htm
DOI: https://dx.doi.org/10.35712/aig.v3.i5.142

INTRODUCTION

Gastrointestinal (GIS) cancers, including tumors of the esophagus, stomach, colon, and rectum, are an important health problem worldwide. Although the incidence of esophageal cancer (EC) is relatively low, gastric cancer (GC) and colorectal cancer (CRC) are among the most common types of cancer (fifth and third, respectively)[1]. They are also responsible for a substantial proportion of cancer mortality, with GC being the third and CRC the second most common cause of cancer-related death[2]. Although various predictive and prognostic parameters are currently available, the mortality rates for patients with GIS cancer are, unfortunately, still very high[2]. It has been shown that rectifying this situation may depend on paving the way for more personalized treatment strategies that lead to a better prognosis and/or fewer treatment side effects[3,4]. Therefore, the meticulous and complete evaluation of patients to determine the appropriate treatment is critical.

In this context, in addition to providing a definitive diagnosis, the role of an accurate evaluation of pathological parameters related to the behavior and proper treatment of GIS tumors cannot be ignored. However, pathology, a morphology-based specialty, is susceptible to subjectivity regarding intraobserver and interobserver variations, particularly in oncology. That is why, in recent years, the search for more objective criteria to eliminate bias, as well as to reduce the growing workload and to contribute time-saving, has allowed the improvement of image analysis-based digital pathology (DP), which has an important place in modern pathological applications[5,6].

In particular, significant advances in slide scanner technology, which can rapidly digitize all pathological slides at high resolution whole slide images (WSIs), has enabled not only the analysis of a wide range of morphological parameters but also the detection of biomarkers/genetic changes in many types of tumors[7-9]. The ability of computer-based analysis to detect prognostic and predictive markers from these images, depending on the fact that they are composed of number matrices containing a large amount of information that is not accessible to the human eye, has led to the adoption of artificial intelligence (AI) for DP[10,11]. Accordingly, the number of studies on AI applications associated with the diagnosis, follow-up, and treatment of many tumors has increased significantly over time. Regarding GIS, data from previous studies evaluating pathological prognostic parameters with various AI models suggest that using these methods may be beneficial. Unfortunately, these encouraging results have not overcome the wide range of challenges to be solved, from laboratory infrastructure to improving the robustness of algorithms, before incorporating AI applications into real-life pathology practice.

This review presented the applications of AI in the evaluation of pathological parameters related to the behavior of GIS cancer, along with a brief overview of the opportunities and challenges encountered in its implementation in pathology.

GENERAL VIEW OF AI IN PATHOLOGY LABORATORIES

In parallel with technological developments, the evolution of whole slide imaging (WSI) has provided remote diagnosis, consultation, and education[12-14]. In the recent past, it was suggested that the use of WSI is comparable to, or even better than, conventional microscopic examination for decision-making in pathology[15-17]. On the other hand, WSIs are also crucial in applying AI methods in pathological practice. They not only provide quick access to the archive without loss of image quality, but they can also render gigabit images, which are very difficult to process, suitable for processing by "tessellation"

[18]. This preprocessing is based on cutting a large image into nonoverlapping smaller patches called "tiles," making them amenable to computational analysis. It should be noted that although some pathological studies use selected images captured manually with a camera, WSI is currently recommended as a standard for AI applications, especially in tumors where heterogeneity is frequent, such as those of the GIS[19].

To achieve reliable results with WSIs, many steps, from preserving the structure of the tissue to the preparation of sections, must be carried out with care in the pathological laboratory. In particular, it is imperative to evaluate and check slides for artifacts (tears, floating contamination, thickness) that have the potential to adversely affect digitization and, thus, AI applications[20,21]. However, it should be noted that even with optimal protocols and slide scanner standardization, the importance of color normalization to ensure consistency in WSI databases should not be overlooked, as it can affect the robustness of deep learning (DL) models. Accordingly, histogram-matching color transfer and spectral matching methods can be applied[22-24]. However, as these methods depend on the expertise of pathologists and are impractical for manual adjustment, various algorithms have been proposed by researchers capable of performing this normalization. Although promising results have been obtained, there is a need for future studies on the performance of AI models using color normalization systems[25, 26].

The gradual evolution of traditional pathology into DP has led to the development of powerful and user-friendly WSI analysis software tools with the ability to manage substantial WSIs and metadata from different hardware manufacturers, as well as interactive drawing annotation capabilities to facilitate decision-making and reporting. Moreover, a significant proportion of them is freely available [27-29]. In addition, the high costs of hardware required for high-performance computation in software development have become more affordable, leading to the implementation of DP in major medical centers[16,30-32]. Increasing the number of centers capable of using DP will allow for the generation of large and high-quality WSI databases, enabling the acquisition of large datasets and the design of algorithms for AI. However, the requirement of a significant investment is still an obstacle to overcome for the widespread application of these technologies[33]. In addition, the problem of proprietary datasets persists, limiting the repeatability of the proposed methodologies and hindering advancement in this field.

As mentioned above, the ability of AI to extract meaningful information from images that the naked human eye cannot discriminate makes it an attractive tool in the field of image processing and analysis in pathology. Therefore, contemporary AI models have evolved from expert systems to different types, such as machine learning (ML) and DL (Table 1). In brief, ML is a subtype of AI that provides a computer system to automatically learn and develop from datasets on its own and solve problems without explicit programming[34-36]. DL is a subfield of ML that employs sophisticated algorithmic structures inspired by the neural network of the human brain (artificial neural network, ANN) in which statistical models are established from input training data[37-39]. Therefore, DL requires large, annotated datasets to develop its algorithms. At present, the annotation of datasets is a complex task in model development[9,40]. In practice, the time-consuming and challenging nature of annotation, especially in systems where heterogeneous lesions are common, such as GIS, may affect the accuracy of the model being trained[41]. Another limitation is that the dataset obtained by a study group does not show the same performance when compared to external validation sets from other institutions. Recently, studies have been conducted to overcome the hindering properties of annotation[42-44]. It has also been suggested that the adoption of DP for diagnosis could indirectly facilitate the generation of valuable datasets for future algorithm development by enabling pathologists to describe areas of interest during evaluation and reporting[45].

It has often been emphasized that the validation of AI-based technologies requires an evidence-based approach[42,46]. This should also be considered in a laboratory-based medical specialty such as pathology. On the other hand, analyzing the performance of AI techniques to that of pathologists is a significant challenge regarding interobserver and interobserver heterogeneity. Currently, the problems related to establishing "ground truth" in AI methods should not be overlooked[40,47]. It should be noted that this requires repeated testing of the effectiveness and consistency of AI applications in many different patient populations. The relative lack of a validation cohort in developing AI-powered DP applications is also related to the possible drawbacks of sharing histopathological slides. Despite interobserver heterogeneity and variability in pathological assessment also demonstrating the uncertainty of "ground truth" in this regard, multi center assessments involving multiple pathologists and datasets may be the best way to overcome this obstacle.

Before the integration of AI into the pathology workflow, the need to validate its benefits and address ethical recommendations increases the importance of AI-based tools being transparent and interpretable, resulting in an increasing demand for more explainable AI models. In this respect, there is a dilemma about the application of AI. Because most algorithms developed use DL, ensemble methods called "black box" models to tackle multidimensional problems are very complex. However, more straightforward methods that are not complex are not powerful enough to achieve the expected results [48]. For this reason, model interpretability, ethical concerns, and potential regulatory barriers should also be considered in newly developed AI tools to meet these expectations.

Table 1 General features of machine learning methods in the development of artificial intelligence models in gastrointestinal pathology

AI models	Strengths	Weaknesses
ML, Traditional, Supervised	Data output can be produced from the previously labeled training set	Labeling big data takes a considerable amount of time and can be challenging
	Allows users to reflect domain knowledge features	Feature extraction quality significantly affects the accuracy
ML, Traditional, Supervised	Users do not supervise the model or label any data	Input data is unknown and not labeled
	Patterns are detected automatically	Precise information related to data sorting is not provided
	Save time	Interpretation is challenging
SVM	Suitable for more efficient regression and classification analysis with high-dimensional data	Not suitable for large data sets. Requires more time for training; Low performance in overlapping classes
CNN	No labeling is required for important information and features	Lack of interpretability due to black boxes
	The performance capacity in image recognition is high	
FCN	Provides computational speed	A large amount of labeled data for training is required
	The background noise is automatically eliminated	The labeling cost is high
RNN	Able to decide which information to remember from past experiences	The model is hard to train
	A suitable deep learning model for sequential data	The computational cost is high
MIL	A detailed annotation is not required	A large amount of training data is required
	Suitable to be performed on large datasets	The computational cost is high
GAN	The potential to produce new realistic data that resembles the original data	The model is hard to train

AI: Artificial intelligence; ML: Machine learning; SVM: Support Vector Machine; CNN: Convolutional neural networks; FCN: Fully convolutional neural networks; RNN: Recurrent neural networks; MIL: Multi-instance learning; GAN: Generative adversarial networks.

AI IN THE PATHOLOGICAL DETERMINATION OF PRENEOPLASTIC LESIONS IN GIS

Barrett's esophagus

The majority of AI studies in EC consist of imaging studies. In pathology, there have been recent studies on the diagnosis of Barrett's esophagus (BE) and the evaluation of dysplasia in these lesions to predict the risk of EC[49,50]. A proposed attention-based deep NN framework for detecting BE and adenocarcinoma (ADC) was found to be reliable with a mean accuracy of 0.83[49]. Unlike existing methods based on the region of interest, this model is based on tissue-level annotations, suggesting that it may provide a new approach for applying DL in pathology. On the other hand, the fact that the study was performed in a single center and on a relatively small data set necessitates the development of the proposed model with further studies. Since trefoil factor 3 expression is the key finding of BE, a DL model (VGG16) using immunohistochemically stained sections showed significant adaptability, with an area under the curve (AUC) of 0.88[50]. Although the proposed approach reduced the pathologist workload by 57%, the underlying ML model still needs further optimization.

Colorectal polyp classification

In CRC, unlike GC, the classification of polyps is an important task to determine the risk of CRC and the future surveillance needs of patients[51]. In routine examinations, high-risk polyps are evaluated based on their histopathological features with considerable interobserver variability among pathologists[52,53]. However, a precise diagnosis of high-risk polyps is required for efficient and early detection of cancer. In addition, the recommendation for endoscopic screening of these lesions for an early diagnosis of CRC, especially in elderly individuals, increases the workload of daily pathology practice[54].

Therefore, AI applications have been developed to classify high-risk colorectal polyps and/or adenomas with high-grade dysplasia. In studies on the classification of these lesions and the identification of CRC, datasets of three to six specific categories and five models were used[55-62] (Table 2). Although most studies showed good performance with generally high AUCs and accuracies, because of the following restrictions, the evidence level of each model needed to be improved. The number of patches and WSIs that make up the datasets are different. Accordingly, in some studies, the number of datasets may affect the reliability of the results. In various studies, the annotation process is not delineated in detail. In addition, the fact that each model has a different focus and characteristics makes

Table 2 AI-based applications in pathology for the determination of tumor behavior in colorectal carcinomas

Ref.	Task	Data sets	Algorithm/Model	Performance	Comments
Xu *et al*[55]	NL/ADC/MC/SC/PC/CCTA	717 patches	AlexNet	Accuracy: 97%	The model provides the classifications of tumor subtypes
Korbar *et al*[56]	NL/HP/SSP/TSA/TA/TVA-VA	Training set: 458 WSIs; Test set: 239 WSIs	ResNET	F1 Score: 88.8%; Accuracy: 93%; Precision: 89.7%; Recall: 88.3%	The model may reduce the workload of pathologists in the assessment of colorectal polyps
Haj-Hassan *et al*[57]	NL/AD/ADC	30 patients, Multispectral image patches	CNN	Accuracy: 99.2%	CNN allows the classification of CRC tissue types using pre-segmented regions of interest
Ponzio *et al*[58]	NL/AD/ADC	27 WSIs	VGG16	Accuracy: 96%	TL considerably outperforms the CNN fully trained on CRC samples on the same test dataset
Sena *et al*[59]	NL/HP/AD/ADC	393 images	CNN	Accuracy: 80%	DL may provide a valuable tool to assist pathologists in the histological classification of CR tumors
Iizuka *et al*[60]	NL/AD/ADC	4036 WSIs + 500WSIs	CNN/RNN	AUCs: 0.96-0.99	Integrating DL models in pathology workflow would be of high benefit for easing the workload of pathologists
Wei *et al*[61]	NL//TA/TVA/VA/HP	1182 WSIs	ResNet	Accuracy: 93.5% (Internal test set); Accuracy: 87% (External test set)	This model may assist pathologists by improving the accuracy of CRC screening
Awan *et al*[62]	NL/Low GR/High GR	139 images	CNN	Accuracy: 97% (two-class), 91% (three-class)	The model provides the classifications of tumor subtypes based on the shape of glands
Sirinukunwattana *et al*[97]	Prediction of MSTs	510 WSIs (FOCUS), 431 WSIs (TCGA), 265 WSIs (GRAMPIAN cohort)	Inception V3	AUCs: 0.9 (FOCUS); 0.94 (TCGA), 0.85 (GRAMPIAN cohort)	RNA expression classifiers can predict from H-E stained images, opening the door to cheap and reliable biological stratification within routine workflows
Echle *et al*[98]	MSI *vs* MSS	6406 WSIs (Training); 771 WSIs (External validation)	ShuffleNet	AUC: 0.92 (Training); AUC: 0.96 (External validation)	The model provides a low-cost evaluation of MSI without molecular testing
Kather *et al*[80]	MSI *vs* MSS	60894 patches (TCGA-CRC-KR); 93408 patches (TCGA-CRC-DX)	ResNet18	AUC: 0.84 (TCGA-CRC-KR); AUC: 0.77 (TCGA-CRC-DX)	This method may lead to improvements in molecular subtype screening workload in pathology
Kather *et al*[77]	Prediction of molecular Als	426 patients (TCGA-CRC); 379 patients (DACHS)	ShuffleNet	AUROC: 0.76	The algorithm predicts a wide range of molecular alterations from routine, H-E stained slides
Kruger *et al*[99]	Prediction of MSTs	919 WSIs	ResNet 34	AUCs: Mean: 0.87; CMS1: 0.85; CMS2: 0.92, CMS3: 0.85; CMS4: 0.86	The MIL framework can identify morphological features indicative of different molecular subtypes
Popovici *et al*[100]	Prediction of MSTs	300 WSIs	VGG-F	Accuracy: 0.84;	The image-based

				Recall: 0.85; Precision: 0.84	classifier shows a significant prognostic value similar to the molecular counterparts
Cao *et al*[101]	MSI *vs* MSS	429 patients (TCGA-COAD); 785 patients (Asian-CRC)	EPLA	AUC: 0.88 (TCGA-COAD); AUC: 0.85 (Asian-CRC)	This pathomics-based model provides MSI estimation directly from images without molecular testing
Bilal *et al*[102]	Prediction of molecular AIs	502 slides (TCGA-CRC-DX); 47 slides (PAIP)	ResNet18, ResNet34, HoVerNet	AUROCS: HM (0.81 *vs* 0.71); MSI (0.86 *vs* 0.74); CIN (0.83 *vs* 0.73), BRAFmut (0.79 *vs* 0.66), TP53mut (0 *vs* 0.64), KRASmut (0.60), CIMP (0.79)	This algorithm is based on non-annotated images and uses only slide-level labels to predict the status of CRC pathways and mutations
Kwak *et al*[110]	LNM prediction	164 patients	CNN, U-Net	AUROC: 67%	PTS score is a potential prognostic parameter for LNM in CRC
Pai *et al*[111]	LNM prediction	230 patients (training), (136 testing)	CNN	AUROC: 79%	The model allows to identify and quantify a broad spectrum of histological features, including LNM in CRC
Kiehl *et al*[112]	LNM prediction	3013 patients	ResNET18	AUROC: 74.1%	DL-based analysis may help predict the LNM of patients with CRC using routine HE-stained slides
Weis *et al*[120]	Tumor Budding (Pan-CK)	381 patients	CNN	Spatial clusters of tumor buds correlates to N status (*P*: 0.003)	The model is a feasible and valid assessment tool for tumor budding on WSIs and can predict prognosis
Kather *et al*[121]	ADI, DEB, LYM, MUC, SM	86 slides (Training), 25 slides (Testing); 862 slide (TCGA-COAD)	VGG19	AUC: 98.7% HR: 2.29 (OS); 1.92 (RFS); Deep stroma score HR: 1.99 (*P*: 0.002), Shorter OS	This model can assess the human TME and predict prognosis directly from histopathological images
Shapcott *et al*[122]	TME (EC/IC/FC/MC)	853 patches, 142 images (TCGA-COAD)	CNN	Accuracy: 76% (detection), 65% (classification)	The model provides the assessment of TME in CRC slides
Sirinukunwattana *et al*[123]	a-4 tissues classes; b- prediction of DM	102 cases	Spatially Constrained CNN	a-AUROC: 90.4-99.9%; b-AUROC: 58.6-63.8%	The algorithm provides a digital marker for estimating the risk of DM
Swiderska-Chadaj *et al*[124]	TME Detection of ICs	28 WSIs	FCN/LSM/U-Net	F1-score of 0.80; Sensitivity: 74%; Precision: 86%	DL approaches are reliable for automatically detecting lymphocytes in IHC-stained CRC tissue sections
Geessink *et al*[115]	TSR	129 slides	CNN	HR: 2.48 (DSS); 2.05 (DFS)	CNN defined TSR as an independent prognosticator
Zhao *et al*[125]	TSR	499 patients (Discovery cohort); 315 patients (Validation cohort:)	CNN	TSR, independent prognostic parameter. HRs: 2.48 (Discovery cohort); 2.08 (Validation cohort)	CNN allows objective evaluation of TSR
Zhao *et al*[126]	Mucus tumor ratio low *vs* mucus tumor ratio high	814 patients	CNN	HRs: 1.88 (Discovery cohort); 2.09 (Validation cohort)	The DL quantified mucus tumor ratio is an independent prognostic factor in CRC
Bychkov *et al*[132]	Prognosis LR *vs* HR	420 TMA	VGG-16	HR: 2.3	The model extracts

					more prognostic information from the tissue morphology than the experienced human observer
Skrede *et al*[133]	Prognosis (CSS)	1122 patients (Validation cohort)	DoMorev1	HRs: 1.89 (uncertain *vs* good); 3.84 (poor *vs* good)	The digital marker has the potential to identify patients at LR and HR and provides the selection of treatment
Jiang *et al*[134]	a-HRR *vs* LRR b-Poor *vs* good prognosis	101 patients (Traning); 67 patients (Validation); 47 (TCGA-COAD)	InceptionResNetV2	a-HRs: 8.98 (training); 10.69 (other 2 test groups); b-HRs: 10.687 (training); 5.03 (other 2 test groups)	The selected model offers an independent prognostic predictor which allows stratification of stage III CRC into risk groups

NL: Normal; ADC: Adenocarcinoma; MC: Mucinous carcinoma; SC: Serrated carcinoma; PC: Papillary carcinoma; CCTA: Cribriform comedo-type adenocarcinoma; CRC: Colorectal cancer; HP: Hyperplastic polyp; SSP: Sessile serrated polyp; TSA: Traditional serrated adenoma; TA: Tubular adenoma; TVA: Tubulovillous adenoma; VA: Villous adenoma; AD: Adenoma; CNN: Convolutional neural networks; WSIs: Whole slide images; RNN: Recurrent neural networks; AUC: Area under the curve; GR: Grade; MSI: Microsatellite instable; MSS: Microsatellite stable; TCGA-CRC: Tumor Cancer Genome Atlas-Colorectal cancer; KR: Frozen tissues; DX: Formalin fixed paraffin embedded tissues; COAD: Colon adenocarcinoma; CRC: Colorectal carcinoma; EPLA: Ensemble patch likelihood aggregation; MST: Molecular subtype; CMS1: Tumor with MSI; CMS2: Tumors exhibiting epithelial gene expression, activated WNT and MYC signaling; CMS3: Tumors with metabolic disregulations; CMS4: Tumors that possess TGF-β; MIL: Multi instance learning; Als: Alterations; AUROC: Area under the receiver operating characteristics; PAIP: Pathology artificial intelligence platform; HM: Hypermutation; CIN: Chromosomally unstable; CIMP: CpG island methylator phenotype; CK: Cytokeratin; ML: Machine learning; ADI: Adipocyte; DEB: Debris; LYM: Lymphocytes; MUC: Mucus; SM: Smooth muscle; HR: Hazard ratio; OS: Overall survival; RFS: Recurrence free-survival; TME: Tumor microenvironment; ICs: Immune cells; FCN: Fully convolutional network; LSM: Liquid state machine; IHC: Immunohistochemistry; EC: Epithelial cell; FC: Fibroblast; MC: Miscellaneous; TSR: Tumor stroma ratio; LR: Low risk; HR: High risk; TMA: Tissue microarray; CSS: Cancer specific survival; HRR: High recurrence risk; LRR: Low recurrence risk; DM: Distant metastasis; LNM: Lymph node metastasis; PTS: The predictive value of the peritumoral stroma score.

their comparison across studies impossible. One of the most striking examples of these studies is Korbar *et al*[56], where a DL model (ResNet-152) trained with over 400 WSIs showed a high overall accuracy in subtyping polyps. In another study, Wei *et al*[61], who ensembled five layers of ResNet, could classify these lesions with WSIs from a single institution, even in external datasets with a performance comparable to that of histopathological evaluation. This data indicates that further manual annotations by various qualified GI pathologists may be required to decrease classification problems in future AI systems for colorectal polyp detection.

AI IN THE PATHOLOGICAL DETERMINATION OF TUMOR BEHAVIOR IN GIS

In this section relevant data on GC and CRC will be discussed. Unfortunately, no AI studies have identified the parameters that are important in determining tumor behavior and survival in EC. Similarly, studies of EC concerning molecular characterization have not been found. Therefore, in EC, a tumor with extremely high mortality, it is clear that additional pathology studies are necessary to reveal the effectiveness of AI applications in predicting tumor behavior.

TUMOR SUBTYPING

Gastric cancer

Although nearly all GC are ADC, the clinicopathological features and behaviors show considerable variation depending on the histopathological diversity of tumor cells[63,64]. In recent years, it has been reported that the survival of patients with GC at the same stage differs significantly among the different subtypes. Therefore, accurate histopathological classification is critical in determining their prognosis, monitoring, and treatment.

GC is often classified based on the ADC differentiation grade, including well-differentiated ADC and poorly differentiated ADC. The grading depends on the presence or absence of glandular structure formation. ADCs are divided into intestinal and diffuse subtypes based on the Lauren classification[65]. While the diffuse form comprises a poorly differentiated type and signet ring cell carcinoma (SRCC), the intestinal type exhibits glands with papillae, tubules, or solid regions. Diffuse-type carcinomas are commonly confused with other nonneoplastic diseases. Because they usually consist of solitary dispersed cells in a desmoplastic stroma and inflammation.

In most of the reported studies, the adenocarcinoma differentiation grade is judged through manual identification by pathologists. Although there have been many studies on AI applications in the pathological diagnosis of GC in the recent past, there are few studies regarding tumor subclassification (Table 3). Yasuda *et al*[66] investigated the features and classification of GC tissues by using supervised ML algorithms. The results showed that this method reliably identifies morphological changes in tumors with different grades. Interestingly, PD-L1 expression levels have been found to serve as a morphological classification in hematoxylin and eosin (HE)-stained slides and correlate with histological grades. Therefore, quantitative analyses of tissue morphology may reveal molecular alterations in malignancies, and molecular analyses may aid in the pathological evaluation of cancer tissues. In another study, four different DL models were used to classify GC into diffuse ADC *vs* other ADC subtypes[67]. From biopsy WSIs, the trained model performed well at identifying both poorly differentiated ADC and SRCC cells. The authors pointed out that while higher magnification can reduce the false positive rate in classification, applying an RNN model with a more comprehensive dataset yields good results even at low magnifications. Hybrid models such as StoHisNet have also distinguished tubular, mucinous, and papillary subtypes of GC. This model showed a higher performance for multiclassification of pathological images of GC than other CNN-based models[68]. Although the model performed well in the four classifications of gastric pathological images, the study group does not include SRCC and other types. Also, the inability of the supervised network in the study to use unlabeled data and the lack of information on which combination maximizes the performance of the model performance warrant further studies. More recently, Su *et al*[69] demonstrated that DL models constructed using a pre-trained ResNet-18 model based on ImageNet27 achieved tumor differentiation recognition or poorly differentiated ADC and well-differentiated ADC classes, respectively. Although these results suggest that AI may be useful in GC classification, the scarcity of data and the differences in classification parameters used in these studies make it difficult to come to any solid conclusions.

Recently, GC has also been classified by the Tumor Cancer Genome Atlas (TCGA) into four molecular subtypes that are also included in the latest World Health Organization classification: Epstein–Barr-virus (EBV)-positive (9%), microsatellite unstable (MSI) (22%), genomically stable (GS) (19%) and chromosomally unstable (CIN) (50%)[70,71]. The clinical significance of this classification comes from the fact that various factors, such as the prognosis and treatment response, differ among these subtypes [72,73]. In particular, among all subclasses of GC, tumors with MSI and positive EBV are associated with a better response to immunotherapy[72]. Consequently, recognizing these subtypes is crucial for categorizing patients who benefit from these treatments. Nevertheless, such classification requires the application of costly techniques, such as immunohistochemistry, and molecular testing, such as polymerase chain reaction, into pathological practice.

On the other hand, these two types have known characteristic histopathological findings. While EBV-positive GCs show prominent infiltration of lymphocytes into the neoplastic epithelium and the stroma, MSI subtype shows significant lymphocytic infiltration, intestinal-type histology, and expanding growth characteristics[63,74,75]. Therefore, these morphological features could be used to make predictions about the molecular subtype. In recent years, it has been suggested that molecular findings can be detected with AI *via* WSIs from HE-stained sections produced for pathological assessment[76-78]. Various models have been applied for molecular subtyping of GIS cancers. However, most of these studies have been conducted on CRCs (see below), whereas relatively few studies are available for GC (Table 3). For the detection of GC subtypes, Muti *et al*[79] demonstrated that DL could detect MSI and EBV positivity independently from each other in GC directly from HE-stained tissues in multi center pooled cohorts. They observed a high classification performance for the detection of MSI and EBV status. The relatively limited number of cases with positive findings and the fact that the ground truth methods for MSI were developed in CRC are presented as potential limitations of this study. On the other hand, their findings align with previous observations[69,80,81]. In addition, large-scale and multicenter validation broadens their work, which has considerable potential for integration into clinical procedures, suggesting that the application of DL could be a substitute for molecular techniques in the classification of GC. Furthermore, because these two subtypes share common morphological features and they are immunotherapy-sensitive tumors, Hinata *et al*[82] combined MSI and EBV in DL models and found they had a higher detection accuracy. This finding has been interpreted based on the possibility that these subtypes have similar distinctive pathological features, such as abundant stromal lymphocytic infiltration and intraepithelial lymphocytosis. On the other hand, the use of tissue microarray and manual labeling of tumor regions for TCGA presented as sources of bias compared to whole tissue slides, given the heterogeneity of tumor tissue. It was also emphasized that manual annotation by a pathologist might be a challenge to overcome by some weakly supervised methods (for example, attention-based deep multi instance learning) in the field of DL for the broad application of the proposed model.

Recently, a DL model called EBVNet that assists pathologists in predicting EBV from HE-stained slides has been introduced in GC[83]. The results suggested that human-machine fusion dramatically enhances the diagnostic ability of both EBVNet and the pathologist. However, this study has some limitations regarding its retrospective evaluation of training and validation. Additionally, the logistic regression model applied in the assessment is still an indirect way to interpret the model. More importantly, as in many DL models, the EBVNet decision-making procedure by the neural network is

Table 3 Artificial intelligence-based applications in pathology for the determination of tumor behavior in gastric cancer

Ref.	Task	Data sets	Algorithm/Model	Performance	Comments
Yasuda *et al*[66]	NC, GR1, GR2, GR3; PDL-1, ATF7IP/MCAF1	66 WSIs	SV, ML, wndchrm	AUCs: 0.98-0.99	The model allows grading emphasizing a correlation between molecular expression and tissue structures
Kanavati *et al*[67]	NC, ADC-D, ADC-O	1-stage training: 1950 WSIs, 2-stage training: 874 WSIs	CNN and RNN	AUCs: 0.95-0.99	The tool can aid pathologists by potentially accelerating their diagnostic workflow
Fu *et al*[68]	NC, TC, MC, PC	Training 2938 WSIs, Testing 980 WSIs	StoHisNet	The accuracy: 94.69%, F1 score: 94.96%, Recall: 94.95%, Precision: 94.97%	The model has high performance in the multi-classification on gastric images and shows strong generalization ability on other pathological datasets
Su *et al*[69]	NC, WD, PD, MSS *vs* MSI	GR: Training 348 WSIs, Testing 88 WSIs MSS: Training 212 WSIs, Testing: 52 WSIs, MSI: Training 136 WSIs, Testing: 36 WSIs	ResNet-18	PD *vs* WD, F1 score: 0.8615, PD *vs* WD *vs* NC, F1 score: 0.8977; MSI *vs* MSS accuracy: 0.7727	The proposed system integrated the tumor GR and MSI status recognition problems into the same workflow and was suitable for exploring the relationships between pathological features and molecular status
Muti *et al*[79]	MSI *vs* MSS; EBV (+) *vs* EBV (-)	2823 patients with known MSI status; 2685 patients with known EBV status	CNN, Shufflenet	MSI *vs* MSS, AUROCs: 0.723-0.863; EBV (+) *vs* EBV (-), AUROCs: 0.672-0.859	DL-based classifiers have the potential to provide faster decisions for pathologists and to offer therapeutic options tailored to the molecular profile of the individual patient
Kather *et al*[80]	MSI *vs* MSS	Training 81 patients +216 patients (TCGA-STAD)	ResNet-18	AUC: 0.84	This system provides significant improvements in molecular alterations screening workflow
Kather *et al*[81]	EBV (+) vs. EBV (-)	Training 317 patients (TCGA-STAD)	CNN, VGG19	AUC: 0.80	This workflow enables a fast and low-cost method to identify EBV and enables pathologists to check the plausibility of computer-based image classification (the black box of DL)
Hinata *et al*[82]	EBV+MSI/dMMR *vs* EBV- non MSI/dMMR	UTokyo training cohort: 326 patients; TCGA training cohort: 48 patients	CNNs,VGG16, VGG19, ResNet50, EfficientNetB0	AUCs: 0.901–0.992 (Utokyo cohort); AUCs: 0.809–0.931 (TCGA cohort)	The model detects immuno-therapy-sensitive GC subtypes from histological images at a lower cost and in a shorter time than the conventional methods
Zheng *et al*[83]	EBV (+) *vs* EBV (-)	EBV (+) 203 WSIs; EBV (-) 803 WSIs	EBVNet	AUROC: 0.969, Internal validation; AUROC: 0.941, External dataset AUROC: 0.895, TCGA dataset	The human-machine fusion significantly improves the diagnostic performance of both the EBVNet and the pathologist, provides an approach for the identification of EBV(+) GC, and may help effectively select patients for immuno-therapy
Flinner *et al*[87]	EBV, MSI, GS, CIN	Training 84 WSIs (TCGA-STAD); Testing: 133 WSIs (TCGA-STAD)	CNN, DenseNet161	AUC: 0.76 for four classes	The simplified molecular TCGA and GC subclasses could be predicted by DL directly based on H-E staining
Jang *et al*[88]	CDH1, ERBB2, KRAS, PIK3CA, TP53 mutations	425 FF slides (TCGA-STAD); 320 FT slides (TCGA-STAD)	CNN, Inception-v3	AUCs (FF-FT): CDH1 (0.667-0.778), ERBB2(0.63-0.833), KRAS (0.657-0.838); PIK3CA (0.688-0.761), TP53 (0.572-0.775)	When trained with appropriate tissue data, DL could predict genetic mutations in H-E-stained tissue slides

Huang *et al*[109]	Metastatic LNs	983 WSIs	ESCNN	AUC: 0.9936	ESCNN improves the accuracy of pathologists in identifying metastatic LNs, micrometastases, and isolated tumor cells, allowing for shortening the review time
Hu *et al*[107]	Metastatic LNs	222 patients	RCNN, Xception and DenseNet-121	Accuracy 97.13%; PPV: 93.53, NPV: 97.99%	The system can be implemented into clinical workflow to assist pathologists in preliminary screening for LN metastases in GC patients
Matsushima *et al* [108]	Metastatic LNs	827 lymph nodes	CNN	AUROC: 0.9994	This DL-based diagnosis-aid system can assist pathologists in detecting LN metastasis in GC and reduce their workload
Wang *et al*[106]	Metastatic LNs, T/LNM	9366 slides (7736 with metastasis)	Resnet-50	LNM (+) *vs* (-): Sensitivity 98.5%, Specificity 96.1%; T/LNM: HR: 2.05 (univariate analysis); 1.39 (multivariate analysis)	This system can assist pathologists in detecting LN metastasis in GC and reduce their workload. Besides, T/LNM is prognostic of OS in GC patients
Hong *et al*[116]	dTSR (HE and CK7)	Training 13 WSIs; Testing 358 WSIs	cGAN	Kappa value: 0.623 (dTSR and vTSR); AUROC: 0.907; OS (*P*: 0.0024)	By diagnosing TSR in GC, this model predicts OS in the advanced stage of GC
Meier *et al*[127]	TME + Ki-67	248 patients	CNN	HRs: Ki67&CD20: 1.364, CD20&CD68: 1.338; Ki67&CD68: 1.473	In combination with a panel of IHC markers, this model predicts the prognosis of patients with GC
Huang *et al*[128]	OS	Training: 2261 pictures; Internal validation: 960 pictures	GastroMIL	HR: 2.414 (univariate analysis), 1.843 (multivariate analysis)	The risk score computed by MIL-GC was proved to be the independent prognostic value of GC
Jiang *et al*[129]	5-YS, 5-YDFS	786 patients	ML, SVM	AUCs: 5-YS: 0.834; 5-YDFS: 0.828	The classifier can accurately distinguishes GC patients with different OS and DFS and identifies a subgroup of patients with stage II and III disease who could benefit from adjuvant chemotherapy
Jiang *et al*[130]	Low SVM *vs* High SVM, 5-YS, 5-YDFS	Training: 223 patients; Internal validation: 218 patientsExternal validation: 227 patients	ML, SVM	AUCs: 5-YS: 0.818; 5-YDFS: 0.827	SVM signature distinguish GC patients with different OS and DFS and identifies a subgroup of patients with stage II and III disease who could benefit from adjuvant chemotherapy
Wang *et al*[131]	TME	172 patients	$CG_{Signature}$ powered by AI	AUROCs: 0.960 ± 0.01 (binary classification), 0.771 ± 0.024 to 0.904 ± 0.012 (ternary classification)	Digital grade cancer staging produced by CGSignature predicts the prognosis of GC and significantly outperforms the AJCC 8th edition Tumor Node Metastasis staging system

NC: Non cancer; GR: Grade; ATF7IP/MCAF1: Activating Transcription Factor 7 Interacting Protein; WSIs: Whole slide images; SV: Supervised; ML: Machine learning; wndchrm: weighted neighbor distances using a compound hierarchy of algorithms representing morphology; AUC: Area under the curve; ADC-D: Diffuse adenocarcinoma; ADC-O: Adenocarcinoma other; DL: Deep learning; CNN: Convolutional neural networks; RNN: Recurrent neural network; TC: Tubular carcinoma; MC: Mucinous carcinoma; PC: Papillary carcinoma; WD: Well differentiated; PD: Poorly differentiated; MSI: Microsatellite instable; MSS: Microsatellite stable EBV: Epstein-Barr virus; TCGA-STAD: Tumor Cancer Genome Atlas, Stomach adenocarcinoma dMMR: Deficient mismatch repair; GC: Gastric cancer; AUROC: Area under the receiver operating characteristics; GS: Genomically stable; CIN: Chromosomally unstable; LN: Lymph node; ESCNN: Enhanced streaming CNN; RCNN: Region based CNN; PPV: Positive predictive value; NPV: Negative predictive value; T/LNM: Tumor area-to-metastatic LN-area ratio; dTSR: Digital tumor-stroma ratio; HE: Hematoxylin and eosin; CK7: Cytokeratin 7; cGAN: Conditional generative adversarial network; vTSR: Visual tumor-stroma ratio; OS: Overall survival; TME: Tumor microenvironment; HR: Hazard ratio; 5-YS: Five year survival; 5-YDFS: Five year disease free survival; SVM: Support vector machine; AI: Artificial intelligence.

nontransparent (black boxes). Since various methods have been proposed to solve black boxes in DL in the recent past, additional studies applying these methods will contribute to the determination of the molecular subtypes of AI models of GC[84-86]. In a more recent study, Flinner *et al*[87], in their study emphasizing the error-proneness of the morphological and staining methods used to determine GC subtypes for subclassification, found that DL could be more effective in this regard. On the other hand, they also pointed out that image tiles labeled with false ground truth associated with GC heterogeneity may reduce the accuracy of DL but this can be overcome by first experimentally defining the test data.

Recently, the feasibility of a DL approach has also been evaluated in the classification of GC for mutations in the CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes[88]. High AUCs observed in both frozen and formalin-fixed tissues highlight that DL-based classifiers could predict the mutational status of these tumors. Although these results are promising for the application of AI to subtyping GC, additional studies are necessary, with further refinement of these methods.

Colorectal cancer

Similar to GC, molecular subtyping of CRC is essential for targeted treatment against critical oncogenic signaling pathways. CRCs are divided by molecular consensus into four types (CMS): 1. CMS1: Tumors with MSI that have a good prognosis in non metastatic stages; CMS2: Tumors with intermediate prognosis exhibiting epithelial gene expression, activated WNT and MYC signaling; CMS3: Tumors with intermediate prognosis demonstrating metabolic dysregulations; CMS4: Tumors with a poor prognosis that possess transforming growth factor beta (TGF-β) activation[89-91]. The identification of CRC with MSI is paramount because this group is susceptible to immunomodulating therapies[92,93]. Although some findings, such as tissue architecture, growth pattern, cellular morphology, and distributions of tumor stroma ratio (TSR) and tumor microenvironment (TME) provide some clues about the subclassification of these tumors, molecular stratification of patients necessitates RNA analyses that are expensive and difficult to standardize[94-96]. Accordingly, some studies have investigated the contribution of AI to tumor subclassification from HE-stained tissue sections by DL models (Table 2). Sirinukunwattana *et al*[97] demonstrated that a CNN-based model could detect CMS subtypes. At the same time, they criticized the potential over fitting of the computational model to the training cohort as a limitation of the study. In a more recent study, Echle *et al*[98] developed a DL model in a large series of 8836 cases of CRC to predict MSI tumors. In the international validation of the study group, the algorithm achieved a high performance [area under the receiver operating curve (AUROC) of 0.96][80]. Other investigators have also reported similar results, pointing out the potential use of DL models for detecting molecular subtypes of CRC[77,99-101]. In a retrospective study, a DL pipeline method was developed based on experimental setups similar to previous studies[102]. Three models were used to predict mutation density (low *vs* high), MSI, CIN, and GpG island methylator phenotype. The mutated and wild-type BRAF, TP53, and KRAS types were also investigated. This method showed higher AUROCs for the prediction of hypermutation, MSI, CIN, BRAF, and TP53 compared to previously reported data, suggesting that AI methods may provide the stratification of patients with CRC for targeted therapies. However, further large-scale validations with multicenter datasets are required before their implementation in pathological practice.

LYMPH NODE METASTASIS

Gastric cancer

Another important parameter that predicts GC behavior and treatment is lymph node metastasis (LNM) [103]. However, identifying LNM is still a challenging and tedious task in pathological practice, making the implementation of AI an attractive tool to reduce the workload[104,105]. Although numerous studies have demonstrated that DL-based algorithms can detect metastatic lymph nodes in GC with a similar level of accuracy to human specialists, these algorithms have not yet been implemented into pathology practice[106-108] (Table 3). The failure to integrate these algorithms is related to the characteristics of WSIs, the excessive effort required to apply the annotation, and the limited associated data. Recently, Huang *et al*[109] developed a weakly supervised end-to-end technique termed enhanced streaming CNN (ESCNN). Their results revealed that the routine pathological evaluation benefitted from the AI-assisted LN assessment workflow regarding review time, sensitivity, and consistency. On the other hand, AI-attributable false alarms that misled the pathologists on negative results led to a decrease in specificity from 94% to 84%, which needs more large-scale or multicenter studies to check the effectiveness of the workflow.

Colorectal cancer

Recent evidence indicates that features extracted by DL models from routine histologic slides can predict LNM in CRC[110-112] (Table 2). For example, Kwak *et al*[110] detected LNM by generating a score based on the ratio of peritumoral stroma to tumor tissue on a test set. In another study, the presence of LNM was detected with a model which segmented WSIs into areas such as tumor budding or poorly differentiated clusters[111]. More recently, Kiehl *et al*[112] performed an approach that uses

DL-based image analysis (slide-based artificial intelligence predictor) in association with patient data to estimate LNM in CRC patients. Their results indicated that LNM could be predicted in patients with CRC through AI applications from histological slides to a similar level to using a classifier containing clinical data.

THE TUMOR STROMA RATIO, TUMOR MICROENVIRONMENT AND TUMOR BUDDING

Gastric cancer

In recent years, it has been shown that the TSR in many organ tumors is an important clue to the course of the disease. In particular, stromal dominance has been observed to be an independent prognostic factor in many tumors, including GIS[113,114]. However, TSRs are not included in pathology report protocols because of the lack of a standard procedure among different methodologies and a low reproducibility related to the high interobserver variation[115]. Recently, a DL pipeline has been introduced to facilitate the automated assessment of TSR in GC[116]. Although this model has been shown to be effective in detecting survival according to the low and high TSR rates in advanced GC, it was emphasized that some limitations, such as the nonautomatic selection of hot spots and the use of a single test, should be eliminated. Therefore, there is a need for many studies on the use of AI applications in TSR determination of GC.

In a recent study, a DL model determined the tumor-to-metastatic lymph node-area ratio in metastatic lymph nodes in patients with GC[106]. Statistical analysis also revealed that this ratio is an independent prognostic factor warranting further investigation.

Colorectal cancer

In CRC, recent studies have demonstrated that lymphocytes and fibroblasts profoundly shape the TME and significantly impact tumor behavior[117-119]. In addition, it has been shown that CRC may have a poor prognosis due to tumor budding (1-5 cells in the invasive area)[120]. In the literature, seven studies of AI methods have been identified to determine these parameters in a more objective and time-saving manner (Table 2). However, many of them used different methods. Three models focused on the classification of the cell types, such as epithelial, inflammatory, fibroblast, lymphocytes, and others (mucus, smooth muscle, normal mucosa, stroma, and cancer epithelium)[121-123]. In an elegant study, a DL algorithm was proposed for estimating the risk of distant metastasis by analyzing the TME[123]. Cell detection and cell classification were evaluated in two CNNs used to build a cell network. In each tumor, a tissue phenotype signature was obtained by proportioning the area of tissue phenotypes to the total tissue area. Statistical analysis revealed that the connection frequency (CF) of the smooth muscle ratio, the CF of the inflammation ratio, and the appearance (AP) based on inflammation could independently estimate the development of distant metastasis. Distant metastasis-free survival analysis indicated that CF smooth muscle and AP inflammation ratios were potential prognosticators. Although the hazard ratios for CF of the smooth muscle ratio and AP inflammation were 2.11 and 0.39, respectively, the AUC values for distant metastasis prediction were 0.59 for the CF of the smooth muscle ratio and 0.64 for AP based on inflammation. As emphasized by the authors, specific immunohistochemical staining can improve the prediction of distant metastases by increasing the informative value of histological slides. Another limitation of this study is the small number of metastatic cases. Another recent study was performed to detect CD3- and CD8-positive immune cells on WSIs of slides stained by immunohistochemistry in a multicenter cohort by four different methods[124]. U-Net obtained the highest performance and highest agreement with manual evaluation (0.72), which was higher than that of pathologists (K = 0.64), supporting that DL models are helpful for automatically detecting lymphocytes in immunohistochemically stained tissue sections.

In CRC, the automatic tumor budding evaluation on immunohistochemical pankeratin-stained slides revealed that the absolute number of buds per image was significantly correlated with manually segmented ground truth (R: 0.86)[120]. Interestingly, the number of spatial clusters of buds in hot spots was significantly correlated with the prognosis. In three studies, the impact of detecting the TSR or deep stroma score in CRC by DL algorithms was found to be an independent parameter to predict tumor behavior[115,121,125] (Table 2).

Recently, Zhao *et al*[126] demonstrated that the ratio of the mucinous component in the tumor area (MTR) quantified by AI is an independent prognostic factor in CRC. On the other hand, the most invasive part of primary tumors was selected for evaluation. As noted by the authors, measuring the exact proportion and prognostic value of mucus in the entire tumor is still worthy of further investigation.

SURVIVAL OUTCOMES

Gastric cancer

Another continuing research topic is evaluating survival outcomes in GC with AI models[127-129] (Table 3). Recently, support vector machine (SVM), one of the popular algorithms in ML, has been applied to predict the survival of GC. Jiang *et al*[129] demonstrated that SVM could be useful in predicting the outcome and identifying patients with GC who might benefit from adjuvant therapy. In this study, the classifier incorporated patient gender, carcinoembryonic antigen levels, LNM, and the protein expression level of eight features, composed of CD3 invasive margin (IM), CD3 center of the tumor (CT), CD8IM, CD45ROCT, CD57IM, CD66bIM, CD68CT, and CD34. There were significant variations between the high- and low-GC-SVM classifiers. Recently, Huang *et al*[128] designed MIL-GC (a DL-based model) to predict overall survival (OS) in patients with GC. They observed C-indices of 0.728 and 0.671 in the training and internal validation sets, respectively. The external validation likewise exhibited strong prognostic prediction performance (C-index = 0.657), confirming the resilience of the two models. Furthermore, univariate and multivariate Cox analyses demonstrated that the risk score derived by MIL-GC has independent prognostic significance, indicating the potential of AI approaches to predict GC behavior. Additionally, tumor progression includes complex interactions between malignant cells and their surrounding microenvironment (TME)[130]. TME targeting and reprogramming is, in fact, can be a potential strategy to achieve antitumor effects in many cancers. Several AI studies involving the TME have recently demonstrated that these methods can determine the prognosis of GIS cancers. Regarding GC, Wang *et al*[131], suggested a graph NN-based solution, CellGraph Signature powered AI, for the digital staging of TME and the exact prediction of patient survival by combining and converting multiplexed immunohistochemistry (mIHC) images as Cell-Graphs. The survival prediction achieved outstanding model performance for both binary and ternary classifications. Furthermore, survival analysis revealed that this method outperforms the AJCC 8th edition Tumor Node Metastasis staging system in discriminating both binary and ternary classes with statistical significance (P value < 0.0001), implying the effectiveness and advantages of such an AI-powered digital staging system in DP and precision oncology.

These data demonstrate that AI-based models allow prognosis prediction in GC. However, developing efficient models requires training on large sets reflecting scanning and staining protocols variability.

Colorectal cancer

Regarding prognostic evaluations from HE-stained slides by AI in CRC, some DL models have been developed for prognostication (Table 2). Bychkov *et al*[132] combined a CNN and a recurrent NN model to estimate the disease-specific five-year survival from tumor tissue microarray samples without tissue classification. The model classified patients into a low- or high-risk group (AUC of 0.69). This result was more significant than the AUC of the visual evaluation of the pathologist (AUC of 0.58) or the histological grade determined at the time of the original diagnosis (AUC of 0.57). However, an external dataset was not included. In another study by Skrede *et al*[133], diverse data from four different cohorts were used to develop an automatic prognostic marker to predict the outcome. The model included a CNN used to separate tumor tissue and two other CNN ensembles that identified individuals as having a favorable or poor survival. Patients were assigned as uncertain when the two CNN ensembles predicted different outcomes. In an external test group, the classifier was a strong predictor of survival. In addition, the output of the two CNN ensembles produced a strong predictive score related to patient outcome (AUC of 0.71). A generalization of this approach has been recommended, as an external test cohort from more than one medical center demonstrated similar hazard ratios.

Jiang *et al*[134], to achieve a shorter computational time, developed a hybrid model by synergizing ML algorithms with DL (InceptionResNetV2 and gradient boosting decision machine classifier) to predict the survival of patients with stage III CRC. While the internal test sets constituted a Chinese cohort, external testing was performed on the TCGA cohort. They revealed that the model stratifies patients with stage III colon cancer into high- and low-risk recurrence and poor and favorable prognostic groups directly from tissue sections. These data suggest that the analysis of H-E-stained tissue samples by AI methods could serve as a digital prognostic biomarker in CRC. However, additional studies are warranted to support the evaluation of the performance of these methods in larger patient series.

OVERALL LIMITATIONS OF AI-BASED APPLICATIONS IN REAL-LIFE PRACTICE

In the literature, there are some frequently discussed topics considering the general challenges of AI such as identification of the clinical need, ethical considerations, funding, optimization of data-sets, annotation of the dataset, regulation, validation, and implementation[46].

Recognizing the actual clinical need and defining a potential solution is the first stage in developing the AI application. However, there can be an imbalance between the benefits in daily pathological practice and the total cost of its implementation. As a result, the market for a particular AI tool may be too tiny and it may not be profitable.

Although patients can provide permission for data to be used for studies, constructing AI models may have issues if commercial use is not approved[135]. In order to develop a framework for global data sharing, patient consent should include the possibility of its commercial use for product development[40].

Training on huge datasets is necessary for developing AI systems with high performance in digital pathology. Changes related to differences in fixation, tissue thickness, and variations in staining and scanning protocols encountered in preanalytical and analytical phases may influence data accuracy[136,137]. For example, it is difficult to convert a glass slide to WSI, and changing the hue of the slide could affect AI accuracy. Many AI algorithms have emerged for this purpose recently, including staining and color features[138,139]. In addition, a number of algorithms are presented to optimize WSI quality. These algorithms identify areas of the highest quality and exclude areas that are out of focus or affected by artifacts[140,141].

Concerning the implementation of AI, to enable users to shift the daily routine practice in the pathology laboratory, from glass slides to WSIs, the first step is to install an institutional IT infrastructure. In addition to these changes in infrastructure, pathology residency training might need to be adjusted in accordance with the availability of this new tool. Preventing residents from relying completely on AI while also allowing them to benefit from it as a helping instrument would require fine balancing and planning prior to its installation[142].

Similar to other clinical tests, quality assurance is crucial, hence it is urgently necessary to develop a plan for external quality assurance for applications. Furthermore, laboratory workers should also be familiar with the quality management system.

Although some algorithms and automated AI models are thought to perform better than pathologists, pathologists will always be required to audit technology and control mechanisms in AI implementation [143].

CONCLUSION

In this review, we outlined the potential of AI applications for evaluating pathological parameters related to the behavior of GIS cancers. Current data suggest the merit of AI-based approaches in assessing tumor grading, subtyping, detection of metastasis, and prognosis in GC and CRC. In addition, these methods encourage biomarker discovery by revealing predictions that are impossible when using traditional visual methods. Regarding EC, there is still much room for improvement in developing AI models to predict the behavior of these tumors in pathology. On the other hand, the enormous potential of AI in improving workflows, eliminating simple errors, and increasing objectivity during pathological evaluations to determine the behavior of GIS cancers should motivate researchers to overcome the many remaining hurdles. In algorithm development, variations in imaging data, interobserver variability during interpretations, model transparency, and interpretability are significant challenges to be solved. A large number of studies with external validation and quality controls implemented on large datasets are essential in meeting the standards of these methods. Thereby, AI applications that are practical, interpretable, manageable, and cost-effective can play a crucial role in the development of pathological evaluations to be performed in the prognosis and treatment of GIS tumors.

FOOTNOTES

Author contributions: Yavuz A, Alpsoy A, Gedik EO, Celik MY performed the data acquisition; Bassorgun CI, Unal B and Elpek GO designed the outline and coordinated the writing of the paper; all authors equally contributed to the writing of the paper and preparation of the tables.

Conflict-of-interest statement: All the authors declare that they have no conflict of interest.

Country/Territory of origin: Turkey

ORCID number: Aysen Yavuz 0000-0001-9991-5515; Anil Alpsoy 0000-0003-4978-7652; Elif Ocak Gedik 0000-0003-2618-498X; Mennan Yigitcan Celik 0000-0001-8769-5156; Cumhur Ibrahim Bassorgun 0000-0003-2440-511X; Betul Unal 0000-

0001-7680-1808; Gulsum Ozlem Elpek 0000-0002-1237-5454.

S-Editor: Liu JH
L-Editor: A
P-Editor: Liu JH

REFERENCES

1 **Arnold M**, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, Bray F. Global Burden of 5 Major Types of Gastrointestinal Cancer. *Gastroenterology* 2020; **159**: 335-349.e15 [PMID: 32247694 DOI: 10.1053/j.gastro.2020.02.068]
2 **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]
3 **Johdi NA**, Sukor NF. Colorectal Cancer Immunotherapy: Options and Strategies. *Front Immunol* 2020; **11**: 1624 [PMID: 33042104 DOI: 10.3389/fimmu.2020.01624]
4 **Li K**, Zhang A, Li X, Zhang H, Zhao L. Advances in clinical immunotherapy for gastric cancer. *Biochim Biophys Acta Rev Cancer* 2021; **1876**: 188615 [PMID: 34403771 DOI: 10.1016/j.bbcan.2021.188615]
5 **Niazi MKK**, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol* 2019; **20**: e253-e261 [PMID: 31044723 DOI: 10.1016/S1470-2045(19)30154-8]
6 **Abels E**, Pantanowitz L, Aeffner F, Zarella MD, van der Laak J, Bui MM, Vemuri VN, Parwani AV, Gibbs J, Agosto-Arroyo E, Beck AH, Kozlowski C. Computational pathology definitions, best practices, and recommendations for regulatory guidance: a white paper from the Digital Pathology Association. *J Pathol* 2019; **249**: 286-294 [PMID: 31355445 DOI: 10.1002/path.5331]
7 **Dangott B**, Parwani A. Whole slide imaging for teleconsultation and clinical use. *J Pathol Inform* 2010; **1** [PMID: 20805956 DOI: 10.4103/2153-3539.65342]
8 **Evans AJ**, Depeiza N, Allen SG, Fraser K, Shirley S, Chetty R. Use of whole slide imaging (WSI) for distance teaching. *J Clin Pathol* 2021; **74**: 425-428 [PMID: 32646928 DOI: 10.1136/jclinpath-2020-206763]
9 **Saillard C**, Schmauch B, Laifa O, Moarii M, Toldo S, Zaslavskiy M, Pronier E, Laurent A, Amaddeo G, Regnault H, Sommacale D, Ziol M, Pawlotsky JM, Mulé S, Luciani A, Wainrib G, Clozel T, Courtiol P, Calderaro J. Predicting Survival After Hepatocellular Carcinoma Resection Using Deep Learning on Histological Slides. *Hepatology* 2020; **72**: 2000-2013 [PMID: 32108950 DOI: 10.1002/hep.31207]
10 **Wong ANN**, He Z, Leung KL, To CCK, Wong CY, Wong SCC, Yoo JS, Chan CKR, Chan AZ, Lacambra MD, Yeung MHY. Current Developments of Artificial Intelligence in Digital Pathology and Its Future Clinical Applications in Gastrointestinal Cancers. *Cancers (Basel)* 2022; **14** [PMID: 35954443 DOI: 10.3390/cancers14153780]
11 **Courtiol P**, Maussion C, Moarii M, Pronier E, Pilcer S, Sefta M, Manceron P, Toldo S, Zaslavskiy M, Le Stang N, Girard N, Elemento O, Nicholson AG, Blay JY, Galateau-Sallé F, Wainrib G, Clozel T. Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med* 2019; **25**: 1519-1525 [PMID: 31591589 DOI: 10.1038/s41591-019-0583-3]
12 **Dun XP**, Parkinson DB. Visualizing peripheral nerve regeneration by whole mount staining. *PLoS One* 2015; **10**: e0119168 [PMID: 25738874 DOI: 10.1371/journal.pone.0119168]
13 **Kim D**, Pantanowitz L, Schüffler P, Yarlagadda DVK, Ardon O, Reuter VE, Hameed M, Klimstra DS, Hanna MG. (Re) Defining the High-Power Field for Digital Pathology. *J Pathol Inform* 2020; **11**: 33 [PMID: 33343994 DOI: 10.4103/jpi.jpi_48_20]
14 **Heffner S**, Colgan O, Doolan C. Digital Pathology. Available from: https://www.leicabiosystems.com/en-br/knowledgepathway/digitalpathology/
15 **Borowsky AD**, Glassy EF, Wallace WD, Kallichanda NS, Behling CA, Miller DV, Oswal HN, Feddersen RM, Bakhtar OR, Mendoza AE, Molden DP, Saffer HL, Wixom CR, Albro JE, Cessna MH, Hall BJ, Lloyd IE, Bishop JW, Darrow MA, Gui D, Jen KY, Walby JAS, Bauer SM, Cortez DA, Gandhi P, Rodgers MM, Rodriguez RA, Martin DR, McConnell TG, Reynolds SJ, Spigel JH, Stepenaskie SA, Viktorova E, Magari R, Wharton KA, Qiu J, Bauer TW. Digital Whole Slide Imaging Compared With Light Microscopy for Primary Diagnosis in Surgical Pathology. *Arch Pathol Lab Med* 2020; **144**: 1245-1253 [PMID: 32057275 DOI: 10.5858/arpa.2019-0569-OA]
16 **Hanna MG**, Reuter VE, Ardon O, Kim D, Sirintrapun SJ, Schüffler PJ, Busam KJ, Sauter JL, Brogi E, Tan LK, Xu B, Bale T, Agaram NP, Tang LH, Ellenson LH, Philip J, Corsale L, Stamelos E, Friedlander MA, Ntiamoah P, Labasin M, England C, Klimstra DS, Hameed M. Validation of a digital pathology system including remote review during the COVID-19 pandemic. *Mod Pathol* 2020; **33**: 2115-2127 [PMID: 32572154 DOI: 10.1038/s41379-020-0601-5]
17 **Aloqaily A**, Polonia A, Campelos S, Alrefae N, Vale J, Caramelo A, Eloy C. Digital Versus Optical Diagnosis of Follicular Patterned Thyroid Lesions. *Head Neck Pathol* 2021; **15**: 537-543 [PMID: 33128731 DOI: 10.1007/s12105-020-01243-y]
18 **Yeh FC**, Ye Q, Hitchens TK, Wu YL, Parwani AV, Ho C. Mapping stain distribution in pathology slides using whole slide imaging. *J Pathol Inform* 2014; **5**: 1 [PMID: 24672736 DOI: 10.4103/2153-3539.126140]
19 **Calderaro J**, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. *Gut* 2021; **70**: 1183-1193 [PMID: 33214163 DOI: 10.1136/gutjnl-2020-322880]
20 **Taqi SA**, Sami SA, Sami LB, Zaki SA. A review of artifacts in histopathology. *J Oral Maxillofac Pathol* 2018; **22**: 279 [PMID: 30158787 DOI: 10.4103/jomfp.JOMFP_125_15]
21 **Salvi M**, Acharya UR, Molinari F, Meiburger KM. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput Biol Med* 2021; **128**: 104129

[PMID: 33254082 DOI: 10.1016/j.compbiomed.2020.104129]
22 **Tabesh A**, Teverovskiy M, Pang HY, Kumar VP, Verbel D, Kotsianti A, Saidi O. Multifeature prostate cancer diagnosis and Gleason grading of histological images. *IEEE Trans Med Imaging* 2007; **26**: 1366-1378 [PMID: 17948727 DOI: 10.1109/TMI.2007.898536]
23 **Macenko M**, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, Schmitt C, Thomas NE. A method for normalizing histology slides for quantitative analysis. Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009 June 28–July 1; Boston, USA, 2009: 1107–1110 [DOI: 10.1109/ISBI.2009.5193250]
24 **Yagi Y**. Color standardization and optimization in whole slide imaging. *Diagn Pathol* 2011; **6** Suppl 1: S15 [PMID: 21489185 DOI: 10.1186/1746-1596-6-S1-S15]
25 **Lee JS**, Ma YX. Stain Style Transfer for Histological Images Using S3CGAN. *Sensors (Basel)* 2022; **22** [PMID: 35161789 DOI: 10.3390/s22031044]
26 **Kausar T**, Kausar A, Ashraf MA, Siddique MF, Wang M, Sajid M, Siddique MZ, Haq AU, Riaz I. SA-GAN: Stain Acclimation Generative Adversarial Network for Histopathology Image Analysis. Appl Sci 2022; 12: 288 [DOI: 10.3390/app12010288]
27 **Schneider CA**, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012; **9**: 671-675 [PMID: 22930834 DOI: 10.1038/nmeth.2089]
28 **Bankhead P**, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, James JA, Salto-Tellez M, Hamilton PW. QuPath: Open source software for digital pathology image analysis. *Sci Rep* 2017; **7**: 16878 [PMID: 29203879 DOI: 10.1038/s41598-017-17204-5]
29 **Aubreville M**, Bertram C, Klopfleisch R, Maier A. SlideRunner. In Bildverarbeitung für die Medizin 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 309–314 [DOI: 10.007/978-3-662-56-537-7_81]
30 **Tizhoosh HR**, Pantanowitz L. Artificial Intelligence and Digital Pathology: Challenges and Opportunities. *J Pathol Inform* 2018; **9**: 38 [PMID: 30607305 DOI: 10.4103/jpi.jpi_53_18]
31 **Williams B**, Hanby A, Millican-Slater R, Verghese E, Nijhawan A, Wilson I, Besusparis J, Clark D, Snead D, Rakha E, Treanor D. Digital pathology for primary diagnosis of screen-detected breast lesions - experimental data, validation and experience from four centres. *Histopathology* 2020; **76**: 968-975 [PMID: 31994224 DOI: 10.1111/his.14079]
32 **Eloy C**, Vale J, Curado M, Polónia A, Campelos S, Caramelo A, Sousa R, Sobrinho-Simões M. Digital Pathology Workflow Implementation at IPATIMUP. *Diagnostics (Basel)* 2021; **11** [PMID: 34829458 DOI: 10.3390/diagnostics11112111]
33 **The Royal College of Pathologists**. Digital Pathology. (accessed on 20 March 2022). Available from: https://www.rcpath.org/profession/digital-pathology.html
34 **Gurcan MN**, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng* 2009; **2**: 147-171 [PMID: 20671804 DOI: 10.1109/RBME.2009.2034865]
35 **Rashidi HH**, Tran NK, Betts EV, Howell LP, Green R. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Acad Pathol* 2019; **6**: 2374289519873088 [PMID: 31523704 DOI: 10.1177/2374289519873088]
36 **Saxena S**, Gyanchandani M. Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review. *J Med Imaging Radiat Sci* 2020; **51**: 182-193 [PMID: 31884065 DOI: 10.1016/j.jmir.2019.11.001]
37 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]
38 **Wang X**, Chen H, Gan C, Lin H, Dou Q, Tsougenis E, Huang Q, Cai M, Heng PA. Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. *IEEE Trans Cybern* 2020; **50**: 3950-3962 [PMID: 31484154 DOI: 10.1109/TCYB.2019.2935141]
39 **Silva-Rodríguez J**, Colomer A, Naranjo V. WeGleNet: A weakly-supervised convolutional neural network for the semantic segmentation of Gleason grades in prostate histology images. *Comput Med Imaging Graph* 2021; **88**: 101846 [PMID: 33485056 DOI: 10.1016/j.compmedimag.2020.101846]
40 **Yoshida H**, Kiyuna T. Requirements for implementation of artificial intelligence in the practice of gastrointestinal pathology. *World J Gastroenterol* 2021; **27**: 2818-2833 [PMID: 34135556 DOI: 10.3748/wjg.v27.i21.2818]
41 **Cosatto E**, Laquerre PF, Malon C, Graf HP, Saito A, Kiyuna T, Marugame A, Kamijo K. Automated gastric cancer diagnosis on H and E-stained sections; training a classifier on a large scale with multiple instance machine learning. Proceedings of SPIE - Progress in Biomedical Optics and Imaging. MI: 2013 [DOI: 10.1117/12.2007047]
42 **Mattocks CJ**, Morris MA, Matthijs G, Swinnen E, Corveleyn A, Dequeker E, Müller CR, Pratt V, Wallace A; EuroGentest Validation Group. A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet* 2010; **18**: 1276-1288 [PMID: 20664632 DOI: 10.1038/ejhg.2010.101]
43 **Joshi**, R. ; Kruger, A.J.; Sha, L.; Kannan, M.; Khan, A.A.; Stumpe, M.C. Learning relevant H&E slide morphologies for prediction of colorectal cancer tumor mutation burden using weakly supervised deep learning. *J Clin Oncol* 2020; **38**: e15244 [DOI: 10.1200/JCO.2020.38.15_suppl.e15244]
44 **Lu MY**, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng* 2021; **5**: 555-570 [PMID: 33649564 DOI: 10.1038/s41551-020-00682-w]
45 **van der Laak J**, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med* 2021; **27**: 775-784 [PMID: 33990804 DOI: 10.1038/s41591-021-01343-4]
46 **Colling R**, Pitman H, Oien K, Rajpoot N, Macklin P; CM-Path AI in Histopathology Working Group, Snead D, Sackville T, Verrill C. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. *J Pathol* 2019; **249**: 143-150 [PMID: 31144302 DOI: 10.1002/path.5310]
47 **Elmore JG**, Longton GM, Carney PA, Geller BM, Onega T, Tosteson AN, Nelson HD, Pepe MS, Allison KH, Schnitt SJ, O'Malley FP, Weaver DL. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 2015; **313**: 1122-1132 [PMID: 25781441 DOI: 10.1001/jama.2015.1405]

48 **Amann J**, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020; **20**: 310 [PMID: 33256715 DOI: 10.1186/s12911-020-01332-6]

49 **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]

50 **Gehrung M**, Crispin-Ortuzar M, Berman AG, O'Donovan M, Fitzgerald RC, Markowetz F. Triage-driven diagnosis of Barrett's esophagus for early detection of esophageal adenocarcinoma using deep learning. *Nat Med* 2021; **27**: 833-841 [PMID: 33859411 DOI: 10.1038/s41591-021-01287-9]

51 **Vu HT**, Lopez R, Bennett A, Burke CA. Individuals with sessile serrated polyps express an aggressive colorectal phenotype. *Dis Colon Rectum* 2011; **54**: 1216-1223 [PMID: 21904135 DOI: 10.1097/DCR.0b013e318228f8a9]

52 **Foss FA**, Milkins S, McGregor AH. Inter-observer variability in the histological assessment of colorectal polyps detected through the NHS Bowel Cancer Screening Programme. *Histopathology* 2012; **61**: 47-52 [PMID: 22486166 DOI: 10.1111/j.1365-2559.2011.04154.x]

53 **Osmond A**, Li-Chang H, Kirsch R, Divaris D, Falck V, Liu DF, Marginean C, Newell K, Parfitt J, Rudrick B, Sapp H, Smith S, Walsh J, Wasty F, Driman DK. Interobserver variability in assessing dysplasia and architecture in colorectal adenomas: a multicentre Canadian study. *J Clin Pathol* 2014; **67**: 781-786 [PMID: 25004943 DOI: 10.1136/jclinpath-2014-202177]

54 **US Preventive Services Task Force**, Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, Davis EM, Donahue KE, Doubeni CA, Krist AH, Kubik M, Li L, Ogedegbe G, Owens DK, Pbert L, Silverstein M, Stevermer J, Tseng CW, Wong JB. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA* 2021; **325**: 1965-1977 [PMID: 34003218 DOI: 10.1001/jama.2021.6238]

55 **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]

56 **Korbar B**, Olofson AM, Miraflor AP, Nicka CM, Suriawinata MA, Torresani L, Suriawinata AA, Hassanpour S. Deep Learning for Classification of Colorectal Polyps on Whole-slide Images. *J Pathol Inform* 2017; **8**: 30 [PMID: 28828201 DOI: 10.4103/jpi.jpi_34_17]

57 **Haj-Hassan H**, Chaddad A, Harkouss Y, Desrosiers C, Toews M, Tanougast C. Classifications of Multispectral Colorectal Cancer Tissues Using Convolution Neural Network. *J Pathol Inform* 2017; **8**: 1 [PMID: 28400990 DOI: 10.4103/jpi.jpi_47_16]

58 **Ponzio F**, Macii E, Ficarra E, Di Cataldo S. Colorectal Cancer Classification using Deep Convolutional Networks-An Experimental Study. Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2. Bioimaging, 2018: 58-66. Available from: https://www.sciencepress.org/papers/2018/66431.pdf

59 **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]

60 **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]

61 **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]

62 **Awan R**, Sirinukunwattana K, Epstein D, Jefferyes S, Qidwai U, Aftab Z, Mujeeb I, Snead D, Rajpoot N. Glandular Morphometrics for Objective Grading of Colorectal Adenocarcinoma Histology Images. *Sci Rep* 2017; **7**: 16852 [PMID: 29203775 DOI: 10.1038/s41598-017-16516-w]

63 **Arai T**, Sakurai U, Sawabe M, Honma N, Aida J, Ushio Y, Kanazawa N, Kuroiwa K, Takubo K. Frequent microsatellite instability in papillary and solid-type, poorly differentiated adenocarcinomas of the stomach. *Gastric Cancer* 2013; **16**: 505-512 [PMID: 23274922 DOI: 10.1007/s10120-012-0226-6]

64 **Sugimura H**. Editorial: an obsession with subtyping gastric cancer. *Gastric Cancer* 2013; **16**: 451-453 [PMID: 23483302 DOI: 10.1007/s10120-013-0243-0]

65 **Lauren P**. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An Attempt At A Histo-Clinical Classification. *Acta Pathol Microbiol Scand* 1965; **64**: 31-49 [PMID: 14320675 DOI: 10.1111/apm.1965.64.1.31]

66 **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]

67 **Kanavati F**, Tsuneki M. A deep learning model for gastric diffuse-type adenocarcinoma classification in whole slide images. *Sci Rep* 2021; **11**: 20486 [PMID: 34650155 DOI: 10.1038/s41598-021-99940-3]

68 **Fu B**, Zhang M, He J, Cao Y, Guo Y, Wang R. StoHisNet: A hybrid multi-classification model with CNN and Transformer for gastric pathology images. *Comput Methods Programs Biomed* 2022; **221**: 106924 [PMID: 35671603 DOI: 10.1016/j.cmpb.2022.106924]

69 **Su F**, Li J, Zhao X, Wang B, Hu Y, Sun Y, Ji J. Interpretable tumor differentiation grade and microsatellite instability recognition in gastric cancer using deep learning. *Lab Invest* 2022; **102**: 641-649 [PMID: 35177797 DOI: 10.1038/s41374-022-00742-6]

70 **Fukayama M**, Goldblum JR, Miettinen M, et al Digestive System Tumours, WHO Classification of Tumours (5th edn). International Agency for Research on Cancer: Lyon, 2019. Available from: http://iarc.who.int

71 **Cancer Genome Atlas Research Network**. Comprehensive molecular characterization of gastric adenocarcinoma.

Nature 2014; **513**: 202-209 [PMID: 25079317 DOI: 10.1038/nature13480]

72 **Kim HS**, Shin SJ, Beom SH, Jung M, Choi YY, Son T, Kim HI, Cheong JH, Hyung WJ, Noh SH, Chung H, Park JC, Shin SK, Lee SK, Lee YC, Koom WS, Lim JS, Chung HC, Rha SY, Kim H. Comprehensive expression profiles of gastric cancer molecular subtypes by immunohistochemistry: implications for individualized therapy. *Oncotarget* 2016; **7**: 44608-44620 [PMID: 27331626 DOI: 10.18632/oncotarget.10115]

73 **Sohn BH**, Hwang JE, Jang HJ, Lee HS, Oh SC, Shim JJ, Lee KW, Kim EH, Yim SY, Lee SH, Cheong JH, Jeong W, Cho JY, Kim J, Chae J, Lee J, Kang WK, Kim S, Noh SH, Ajani JA, Lee JS. Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. *Clin Cancer Res* 2017; **23**: 4441-4449 [PMID: 28747339 DOI: 10.1158/1078-0432.CCR-16-2211]

74 **Grogg KL**, Lohse CM, Pankratz VS, Halling KC, Smyrk TC. Lymphocyte-rich gastric cancer: associations with Epstein-Barr virus, microsatellite instability, histology, and survival. *Mod Pathol* 2003; **16**: 641-651 [PMID: 12861059 DOI: 10.1097/01.MP.0000076980.73826.C0]

75 **Fukayama M**, Abe H, Kunita A, Shinozaki-Ushiku A, Matsusaka K, Ushiku T, Kaneda A. Thirty years of Epstein-Barr virus-associated gastric carcinoma. *Virchows Arch* 2020; **476**: 353-365 [PMID: 31836926 DOI: 10.1007/s00428-019-02724-4]

76 **Fu Y**, Jung AW, Torne RV, Gonzalez S, Vöhringer H, Shmatko A, Yates LR, Jimenez-Linan M, Moore L, Gerstung M. Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *Nat Cancer* 2020; **1**: 800-810 [PMID: 35122049 DOI: 10.1038/s43018-020-0085-8]

77 **Kather JN**, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, Krause J, Niehues JM, Sommer KAJ, Bankhead P, Kooreman LFS, Schulte JJ, Cipriani NA, Buelow RD, Boor P, Ortiz-Brüchle NN, Hanby AM, Speirs V, Kochanny S, Patnaik A, Srisuwananukorn A, Brenner H, Hoffmeister M, van den Brandt PA, Jäger D, Trautwein C, Pearson AT, Luedde T. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* 2020; **1**: 789-799 [PMID: 33763651 DOI: 10.1038/s43018-020-0087-6]

78 **Schmauch B**, Romagnoni A, Pronier E, Saillard C, Maillé P, Calderaro J, Kamoun A, Sefta M, Toldo S, Zaslavskiy M, Clozel T, Moarii M, Courtiol P, Wainrib G. A deep learning model to predict RNA-Seq expression of tumours from whole slide images. *Nat Commun* 2020; **11**: 3877 [PMID: 32747659 DOI: 10.1038/s41467-020-17678-4]

79 **Muti HS**, Heij LR, Keller G, Kohlruss M, Langer R, Dislich B, Cheong JH, Kim YW, Kim H, Kook MC, Cunningham D, Allum WH, Langley RE, Nankivell MG, Quirke P, Hayden JD, West NP, Irvine AJ, Yoshikawa T, Oshima T, Huss R, Grosser B, Roviello F, d'Ignazio A, Quaas A, Alakus H, Tan X, Pearson AT, Luedde T, Ebert MP, Jäger D, Trautwein C, Gaisa NT, Grabsch HI, Kather JN. Development and validation of deep learning classifiers to detect Epstein-Barr virus and microsatellite instability status in gastric cancer: a retrospective multicentre cohort study. *Lancet Digit Health* 2021; **3**: e654-e664 [PMID: 34417147 DOI: 10.1016/S2589-7500(21)00133-3]

80 **Kather JN**, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tacke F, Neumann UP, Grabsch HI, Yoshikawa T, Brenner H, Chang-Claude J, Hoffmeister M, Trautwein C, Luedde T. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019; **25**: 1054-1056 [PMID: 31160815 DOI: 10.1038/s41591-019-0462-y]

81 **KatherJN**, Schulte J, Grabsch HI, Loeffler C, Muti H, Dolezal J, Srisuwananukorn A, Agrawal N, Kochanny S, von Stillfried S, Boor P, Yoshikawa T, JaegerD, Trautwein C, Bankhead P, Cipriani NA, Luedde T, Pearson AT. Deep learning detects virus presence in cancer histology. bioRxiv [DOI: 10.1101/690206]

82 **Hinata M**, Ushiku T. Detecting immunotherapy-sensitive subtype in gastric cancer using histologic image-based deep learning. *Sci Rep* 2021; **11**: 22636 [PMID: 34811485 DOI: 10.1038/s41598-021-02168-4]

83 **Zheng X**, Wang R, Zhang X, Sun Y, Zhang H, Zhao Z, Zheng Y, Luo J, Zhang J, Wu H, Huang D, Zhu W, Chen J, Cao Q, Zeng H, Luo R, Li P, Lan L, Yun J, Xie D, Zheng WS, Cai M. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. *Nat Commun* 2022; **13**: 2790 [PMID: 35589792 DOI: 10.1038/s41467-022-30459-5]

84 **Montavon G**, Samek W, Müller KR. Methods for interpreting and understanding deep neuralnetworks. Digit Signal Process 2018; 73: 1-15 [DOI: 10.1016/j.dsp.2017.10.011]

85 **Yang JH**, Wright SN, Hamblin M, McCloskey D, Alcantar MA, Schrübbers L, Lopatkin AJ, Satish S, Nili A, Palsson BO, Walker GC, Collins JJ. A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action. *Cell* 2019; **177**: 1649-1661.e9 [PMID: 31080069 DOI: 10.1016/j.cell.2019.04.016]

86 **Tosun AB**, Pullara F, Becich MJ, Taylor DL, Fine JL, Chennubhotla SC. Explainable AI (xAI) for Anatomic Pathology. *Adv Anat Pathol* 2020; **27**: 241-250 [PMID: 32541594 DOI: 10.1097/PAP.0000000000000264]

87 **Flinner N**, Gretser S, Quaas A, Bankov K, Stoll A, Heckmann LE, Mayer RS, Doering C, Demes MC, Buettner R, Rueschoff J, Wild PJ. Deep learning based on hematoxylin-eosin staining outperforms immunohistochemistry in predicting molecular subtypes of gastric adenocarcinoma. *J Pathol* 2022; **257**: 218-226 [PMID: 35119111 DOI: 10.1002/path.5879]

88 **Jang HJ**, Lee A, Kang J, Song IH, Lee SH. Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach. *World J Gastroenterol* 2021; **27**: 7687-7704 [PMID: 34908807 DOI: 10.3748/wjg.v27.i44.7687]

89 **Guinney J**, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L, Tejpar S. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015; **21**: 1350-1356 [PMID: 26457759 DOI: 10.1038/nm.3967]

90 **Dienstmann R**, Vermeulen L, Guinney J, Kopetz S, Tejpar S, Tabernero J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer* 2017; **17**: 268 [PMID: 28332502 DOI: 10.1038/nrc.2017.24]

91 **Ganesh K**, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, Diaz LA Jr. Immunotherapy in colorectal cancer:

rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol* 2019; **16**: 361-375 [PMID: 30886395 DOI: 10.1038/s41575-019-0126-x]

92 **Li K**, Luo H, Huang L, Zhu X. Microsatellite instability: a review of what the oncologist should know. *Cancer Cell Int* 2020; **20**: 16 [PMID: 31956294 DOI: 10.1186/s12935-019-1091-8]

93 **André T**, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, Smith D, Garcia-Carbonero R, Benavides M, Gibbs P, de la Fouchardiere C, Rivera F, Elez E, Bendell J, Le DT, Yoshino T, Van Cutsem E, Yang P, Farooqui MZH, Marinello P, Diaz LA Jr; KEYNOTE-177 Investigators. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med* 2020; **383**: 2207-2218 [PMID: 33264544 DOI: 10.1056/NEJMoa2017699]

94 **De Smedt L**, Lemahieu J, Palmans S, Govaere O, Tousseyn T, Van Cutsem E, Prenen H, Tejpar S, Spaepen M, Matthijs G, Decaestecker C, Moles Lopez X, Demetter P, Salmon I, Sagaert X. Microsatellite instable vs stable colon carcinomas: analysis of tumour heterogeneity, inflammation and angiogenesis. *Br J Cancer* 2015; **113**: 500-509 [PMID: 26068398 DOI: 10.1038/bjc.2015.213]

95 **Sepulveda AR**, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, Kopetz SE, Lieu C, Lindor NM, Minsky BD, Monzon FA, Sargent DJ, Singh VM, Willis J, Clark J, Colasacco C, Bryan Rumble R, Temple-Smolkin R, B Ventura C, Nowak JA. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and American Society of Clinical Oncology. *Arch Pathol Lab Med* 2017; **141**: 625-657 [PMID: 28165284 DOI: 10.5858/arpa.2016-0554-CP]

96 **Punt CJ**, Koopman M, Vermeulen L. From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat Rev Clin Oncol* 2017; **14**: 235-246 [PMID: 27922044 DOI: 10.1038/nrclinonc.2016.171]

97 **Sirinukunwattana K**, Domingo E, Richman SD, Redmond KL, Blake A, Verrill C, Leedham SJ, Chatzipli A, Hardy C, Whalley CM, Wu CH, Beggs AD, McDermott U, Dunne PD, Meade A, Walker SM, Murray GI, Samuel L, Seymour M, Tomlinson I, Quirke P, Maughan T, Rittscher J, Koelzer VH; S:CORT consortium. Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. *Gut* 2021; **70**: 544-554 [PMID: 32690604 DOI: 10.1136/gutjnl-2019-319866]

98 **Echle A**, Grabsch HI, Quirke P, van den Brandt PA, West NP, Hutchins GGA, Heij LR, Tan X, Richman SD, Krause J, Alwers E, Jenniskens J, Offermans K, Gray R, Brenner H, Chang-Claude J, Trautwein C, Pearson AT, Boor P, Luedde T, Gaisa NT, Hoffmeister M, Kather JN. Clinical-Grade Detection of Microsatellite Instability in Colorectal Tumors by Deep Learning. *Gastroenterology* 2020; **159**: 1406-1416.e11 [PMID: 32562722 DOI: 10.1053/j.gastro.2020.06.021]

99 **Kruger AJ**, Sha L, Kannan M, Joshi RP, Leibowitz BD, Zhang R, Khan AA, Stumpe M. H&E Image-based Consensus Molecular Subtype Classification of Colorectal Cancer Using Weak Labeling. *J Clin Oncol* 2020; **38**: e16097 [DOI: 10.1200/JCO.2020.38.15_suppl.e16097]

100 **Popovici V**, Budinská E, Dušek L, Kozubek M, Bosman F. Image-based surrogate biomarkers for molecular subtypes of colorectal cancer. *Bioinformatics* 2017; **33**: 2002-2009 [PMID: 28158480 DOI: 10.1093/bioinformatics/btx027]

101 **Cao R**, Yang F, Ma SC, Liu L, Zhao Y, Li Y, Wu DH, Wang T, Lu WJ, Cai WJ, Zhu HB, Guo XJ, Lu YW, Kuang JJ, Huan WJ, Tang WM, Huang K, Huang J, Yao J, Dong ZY. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. *Theranostics* 2020; **10**: 11080-11091 [PMID: 33042271 DOI: 10.7150/thno.49864]

102 **Bilal M**, Raza SEA, Azam A, Graham S, Ilyas M, Cree IA, Snead D, Minhas F, Rajpoot NM. Development and validation of a weakly supervised deep learning framework to predict the status of molecular pathways and key mutations in colorectal cancer from routine histology images: a retrospective study. *Lancet Digit Health* 2021; **3**: e763-e772 [PMID: 34686474 DOI: 10.1016/S2589-7500(21)00180-1]

103 **Sun F**, Liu S, Song P, Zhang C, Liu Z, Guan W, Wang M. Impact of retrieved lymph node count on short-term complications in patients with gastric cancer. *World J Surg Oncol* 2020; **18**: 224 [PMID: 32838799 DOI: 10.1186/s12957-020-02000-9]

104 **Veta M**, Pluim JP, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: a review. *IEEE Trans Biomed Eng* 2014; **61**: 1400-1411 [PMID: 24759275 DOI: 10.1109/TBME.2014.2303852]

105 **Bhargava R**, Madabhushi A. Emerging Themes in Image Informatics and Molecular Analysis for Digital Pathology. *Annu Rev Biomed Eng* 2016; **18**: 387-412 [PMID: 27420575 DOI: 10.1146/annurev-bioeng-112415-114722]

106 **Wang X**, Chen Y, Gao Y, Zhang H, Guan Z, Dong Z, Zheng Y, Jiang J, Yang H, Wang L, Huang X, Ai L, Yu W, Li H, Dong C, Zhou Z, Liu X, Yu G. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun* 2021; **12**: 1637 [PMID: 33712598 DOI: 10.1038/s41467-021-21674-7]

107 **Hu Y**, Su F, Dong K, Wang X, Zhao X, Jiang Y, Li J, Ji J, Sun Y. Deep learning system for lymph node quantification and metastatic cancer identification from whole-slide pathology images. *Gastric Cancer* 2021; **24**: 868-877 [PMID: 33484355 DOI: 10.1007/s10120-021-01158-9]

108 **Matsushima J**, Sato T, Ohnishi T, Yoshimura Y, Mizutani H, Koto S, Ikeda JI, Kano M, Matsubara H, Hayashi H. The Use of Deep Learning-Based Computer Diagnostic Algorithm for Detection of Lymph Node Metastases of Gastric Adenocarcinoma. *Int J Surg Pathol* 2022; 10668969221113475 [PMID: 35898183 DOI: 10.1177/10668969221113475]

109 **Huang SC**, Chen CC, Lan J, Hsieh TY, Chuang HC, Chien MY, Ou TS, Chen KH, Wu RC, Liu YJ, Cheng CT, Huang YJ, Tao LW, Hwu AF, Lin IC, Hung SH, Yeh CY, Chen TC. Deep neural network trained on gigapixel images improves lymph node metastasis detection in clinical settings. *Nat Commun* 2022; **13**: 3347 [PMID: 35688834 DOI: 10.1038/s41467-022-30746-1]

110 **Kwak MS**, Lee HH, Yang JM, Cha JM, Jeon JW, Yoon JY, Kim HI. Deep Convolutional Neural Network-Based Lymph Node Metastasis Prediction for Colon Cancer Using Histopathological Images. *Front Oncol* 2020; **10**: 619803 [PMID: 33520727 DOI: 10.3389/fonc.2020.619803]

111 **Pai RK**, Hartman D, Schaeffer DF, Rosty C, Shivji S, Kirsch R, Pai RK. Development and initial validation of a deep learning algorithm to quantify histological features in colorectal carcinoma including tumour budding/poorly differentiated clusters. *Histopathology* 2021; **79**: 391-405 [PMID: 33590485 DOI: 10.1111/his.14353]

112 **Kiehl L**, Kuntz S, Höhn J, Jutzi T, Krieghoff-Henning E, Kather JN, Holland-Letz T, Kopp-Schneider A, Chang-Claude J,

Brobeil A, von Kalle C, Fröhling S, Alwers E, Brenner H, Hoffmeister M, Brinker TJ. Deep learning can predict lymph node status directly from histology in colorectal cancer. *Eur J Cancer* 2021; **157**: 464-473 [PMID: 34649117 DOI: 10.1016/j.ejca.2021.08.039]

113 **van Pelt GW**, Kjær-Frifeldt S, van Krieken JHJM, Al Dieri R, Morreau H, Tollenaar RAEM, Sørensen FB, Mesker WE. Scoring the tumor-stroma ratio in colon cancer: procedure and recommendations. *Virchows Arch* 2018; **473**: 405-412 [PMID: 30030621 DOI: 10.1007/s00428-018-2408-z]

114 **Fu M**, Chen D, Luo F, Li M, Wang Y, Chen J, Li A, Liu S. Association of the tumour stroma percentage in the preoperative biopsies with lymph node metastasis in colorectal cancer. *Br J Cancer* 2020; **122**: 388-396 [PMID: 31787749 DOI: 10.1038/s41416-019-0671-7]

115 **Geessink OGF**, Baidoshvili A, Klaase JM, Ehteshami Bejnordi B, Litjens GJS, van Pelt GW, Mesker WE, Nagtegaal ID, Ciompi F, van der Laak JAWM. Computer aided quantification of intratumoral stroma yields an independent prognosticator in rectal cancer. *Cell Oncol (Dordr)* 2019; **42**: 331-341 [PMID: 30825182 DOI: 10.1007/s13402-019-00429-z]

116 **Hong Y**, Heo YJ, Kim B, Lee D, Ahn S, Ha SY, Sohn I, Kim KM. Deep learning-based virtual cytokeratin staining of gastric carcinomas to measure tumor-stroma ratio. *Sci Rep* 2021; **11**: 19255 [PMID: 34584193 DOI: 10.1038/s41598-021-98857-1]

117 **Kather JN**, Poleszczuk J, Suarez-Carmona M, Krisam J, Charoentong P, Valous NA, Weis CA, Tavernar L, Leiss F, Herpel E, Klupp F, Ulrich A, Schneider M, Marx A, Jäger D, Halama N. In Silico Modeling of Immunotherapy and Stroma-Targeting Therapies in Human Colorectal Cancer. *Cancer Res* 2017; **77**: 6442-6452 [PMID: 28923860 DOI: 10.1158/0008-5472.CAN-17-2006]

118 **Scheer R**, Baidoshvili A, Zoidze S, Elferink MAG, Berkel AEM, Klaase JM, van Diest PJ. Tumor-stroma ratio as prognostic factor for survival in rectal adenocarcinoma: A retrospective cohort study. *World J Gastrointest Oncol* 2017; **9**: 466-474 [PMID: 29290917 DOI: 10.4251/wjgo.v9.i12.466]

119 **Kather JN**, Halama N, Jaeger D. Genomics and emerging biomarkers for immunotherapy of colorectal cancer. *Semin Cancer Biol* 2018; **52**: 189-197 [PMID: 29501787 DOI: 10.1016/j.semcancer.2018.02.010]

120 **Weis CA**, Kather JN, Melchers S, Al-Ahmdi H, Pollheimer MJ, Langner C, Gaiser T. Automatic evaluation of tumor budding in immunohistochemically stained colorectal carcinomas and correlation to clinical outcome. *Diagn Pathol* 2018; **13**: 64 [PMID: 30153844 DOI: 10.1186/s13000-018-0739-3]

121 **Kather JN**, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA, Gaiser T, Marx A, Valous NA, Ferber D, Jansen L, Reyes-Aldasoro CC, Zörnig I, Jäger D, Brenner H, Chang-Claude J, Hoffmeister M, Halama N. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS Med* 2019; **16**: e1002730 [PMID: 30677016 DOI: 10.1371/journal.pmed.1002730]

122 **Shapcott M**, Hewitt KJ, Rajpoot N. Deep Learning With Sampling in Colon Cancer Histology. *Front Bioeng Biotechnol* 2019; **7**: 52 [PMID: 30972333 DOI: 10.3389/fbioe.2019.00052]

123 **Sirinukunwattana K**, Snead D, Epstein D, Aftab Z, Mujeeb I, Tsang YW, Cree I, Rajpoot N. Novel digital signatures of tissue phenotypes for predicting distant metastasis in colorectal cancer. *Sci Rep* 2018; **8**: 13692 [PMID: 30209315 DOI: 10.1038/s41598-018-31799-3]

124 **Swiderska-Chadaj Z**, Pinckaers H, van Rijthoven M, Balkenhol M, Melnikova M, Geessink O, Manson Q, Sherman M, Polonia A, Parry J, Abubakar M, Litjens G, van der Laak J, Ciompi F. Learning to detect lymphocytes in immunohistochemistry with deep learning. *Med Image Anal* 2019; **58**: 101547 [PMID: 31476576 DOI: 10.1016/j.media.2019.101547]

125 **Zhao K**, Li Z, Yao S, Wang Y, Wu X, Xu Z, Wu L, Huang Y, Liang C, Liu Z. Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer. *EBioMedicine* 2020; **61**: 103054 [PMID: 33039706 DOI: 10.1016/j.ebiom.2020.103054]

126 **Zhao K**, Wu L, Huang Y, Yao S, Xu Z, Lin H, Wang H, Liang Y, Xu Y, Chen X, Zhao M, Peng J, Liang C, Li Z, Li Y, Liu Z. Deep learning quantified mucus-tumor ratio predicting survival of patients with colorectal cancer using whole-slide images. *Precis Clin Med* 2021; **4**: 17-24 [PMID: 35693123 DOI: 10.1093/pcmedi/pbab002]

127 **Meier A**, Nekolla K, Hewitt LC, Earle S, Yoshikawa T, Oshima T, Miyagi Y, Huss R, Schmidt G, Grabsch HI. Hypothesis-free deep survival learning applied to the tumour microenvironment in gastric cancer. *J Pathol Clin Res* 2020; **6**: 273-282 [PMID: 32592447 DOI: 10.1002/cjp2.170]

128 **Huang B**, Tian S, Zhan N, Ma J, Huang Z, Zhang C, Zhang H, Ming F, Liao F, Ji M, Zhang J, Liu Y, He P, Deng B, Hu J, Dong W. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: A retrospective multicentre study. *EBioMedicine* 2021; **73**: 103631 [PMID: 34678610 DOI: 10.1016/j.ebiom.2021.103631]

129 **Jiang Y**, Xie J, Han Z, Liu W, Xi S, Huang L, Huang W, Lin T, Zhao L, Hu Y, Yu J, Zhang Q, Li T, Cai S, Li G. Immunomarker Support Vector Machine Classifier for Prediction of Gastric Cancer Survival and Adjuvant Chemotherapeutic Benefit. *Clin Cancer Res* 2018; **24**: 5574-5584 [PMID: 30042208 DOI: 10.1158/1078-0432.CCR-18-0848]

130 **Jiang Y**, Xie J, Huang W, Chen H, Xi S, Han Z, Huang L, Lin T, Zhao LY, Hu YF, Yu J, Cai SR, Li T, Li G. Tumor Immune Microenvironment and Chemosensitivity Signature for Predicting Response to Chemotherapy in Gastric Cancer. *Cancer Immunol Res* 2019; **7**: 2065-2073 [PMID: 31615816 DOI: 10.1158/2326-6066.CIR-19-0311]

131 **Wang Y**, Wang YG, Hu C, Li M, Fan Y, Otter N, Sam I, Gou H, Hu Y, Kwok T, Zalcberg J, Boussioutas A, Daly RJ, Montúfar G, Liò P, Xu D, Webb GI, Song J. Cell graph neural networks enable the precise prediction of patient survival in gastric cancer. *NPJ Precis Oncol* 2022; **6**: 45 [PMID: 35739342 DOI: 10.1038/s41698-022-00285-5]

132 **Bychkov D**, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 2018; **8**: 3395 [PMID: 29467373 DOI: 10.1038/s41598-018-21758-3]

133 **Skrede OJ**, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, Askautrud HA, Pradhan M, Nesheim JA, Albregtsen F, Farstad IN, Domingo E, Church DN, Nesbakken A, Shepherd NA, Tomlinson I, Kerr R, Novelli M, Kerr

DJ, Danielsen HE. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* 2020; **395**: 350-360 [PMID: 32007170 DOI: 10.1016/S0140-6736(19)32998-8]

134 **Jiang D**, Liao J, Duan H, Wu Q, Owen G, Shu C, Chen L, He Y, Wu Z, He D, Zhang W, Wang Z. A machine learning-based prognostic predictor for stage III colon cancer. *Sci Rep* 2020; **10**: 10333 [PMID: 32587295 DOI: 10.1038/s41598-020-67178-0]

135 **Kotsenas AL**, Balthazar P, Andrews D, Geis JR, Cook TS. Rethinking Patient Consent in the Era of Artificial Intelligence and Big Data. *J Am Coll Radiol* 2021; **18**: 180-184 [PMID: 33413897 DOI: 10.1016/j.jacr.2020.09.022]

136 **Inoue T**, Yagi Y. Color standardization and optimization in whole slide imaging. *Clin Diagn Pathol* 2020; **4** [PMID: 33088926 DOI: 10.15761/cdp.1000139]

137 **Yoshida H**, Yokota H, Singh R, Kiyuna T, Yamaguchi M, Kikuchi S, Yagi Y, Ochiai A. Meeting Report: The International Workshop on Harmonization and Standardization of Digital Pathology Image, Held on April 4, 2019 in Tokyo. *Pathobiology* 2019; **86**: 322-324 [PMID: 31707388 DOI: 10.1159/000502718]

138 **Janowczyk A**, Basavanhally A, Madabhushi A. Stain Normalization using Sparse AutoEncoders (StaNoSA): Application to digital pathology. *Comput Med Imaging Graph* 2017; **57**: 50-61 [PMID: 27373749 DOI: 10.1016/j.compmedimag.2016.05.003]

139 **Vahadane A**, Peng T, Sethi A, Albarqouni S, Wang L, Baust M, Steiger K, Schlitter AM, Esposito I, Navab N. Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images. *IEEE Trans Med Imaging* 2016; **35**: 1962-1971 [PMID: 27164577 DOI: 10.1109/TMI.2016.2529665]

140 **Janowczyk A**, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistoQC: An Open-Source Quality Control Tool for Digital Pathology Slides. *JCO Clin Cancer Inform* 2019; **3**: 1-7 [PMID: 30990737 DOI: 10.1200/CCI.18.00157]

141 **Senaras C**, Niazi MKK, Lozanski G, Gurcan MN. DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. *PLoS One* 2018; **13**: e0205387 [PMID: 30359393 DOI: 10.1371/journal.pone.0205387]

142 **Arora A**, Arora A. Pathology training in the age of artificial intelligence. *J Clin Pathol* 2021; **74**: 73-75 [PMID: 33020175 DOI: 10.1136/jclinpath-2020-207110]

143 **Hekler A**, Utikal JS, Enk AH, Solass W, Schmitt M, Klode J, Schadendorf D, Sondermann W, Franklin C, Bestvater F, Flaig MJ, Krahl D, von Kalle C, Fröhling S, Brinker TJ. Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. *Eur J Cancer* 2019; **118**: 91-96 [PMID: 31325876 DOI: 10.1016/j.ejca.2019.06.012]

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