Accepted Manuscript online: 2022-07-22 Article published online: 2022-11-15

Artificial intelligence in gastroenterology: Where are we heading?

OPEN ACCESS

Authors

Glenn Koleth*[,1], James Emmanue*[,1,2], Marco Spadaccini[3,4], Pietro Mascagni[5,6], Kareem Khalaf[3,4], Yuichi Mori[7,8], Giulio Antonelli[9,10], Roberta Maselli[3,4], Silvia Carrara[4], Piera Alessia Galtieri[4], Gaia Pellegatta[4], Alessandro Fugazza[4], Andrea Anderloni[4], Carmelo Selvaggio[3], Michael Bretthauer[7], Alessio Aghemo[3,11], Antonino Spinelli[3,12], Victor Savevski[13], Prateek Sharma[14], Cesare Hassan**[,3,4], Alessandro Repici**[,3,4]

Institutions

1. Hospital Selayang, Department of Gastroenterology and Hepatology, Selangor, Malaysia
2. Queen Elizabeth Hospital, Department of Gastroenterology and Hepatology, Sabah, Malaysia
3. Humanitas University, Department of Biomedical Sciences, Pieve Emanuele, Italy
4. Humanitas Clinical and Research Center – IRCCS, Endoscopy Unit, Rozzano, Italy
5. Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
6. Institute of Image-Guided Surgery, IHU-Strasbourg, France
7. Clinical Effectiveness Research Group, Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway
8. Digestive Disease Center, Showa University Northern Yokohama Hospital, Yokohama, Japan
9. "Sapienza" University of Rome, Department of Anatomical, Histological, Forensic Medicine and Orthopedics Sciences, Rome, Italy
10. Ospedale dei Castelli Hospital, Gastroenterology and Digestive Endoscopy Unit, Ariccia, Rome, Italy
11. Humanitas Research Hospital – IRCCS, Internal Medicine and Hepatology Unit, Rozzano, Italy
12. Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy; Division of Colon and Rectal Surgery, Humanitas Clinical and Research Hospital IRCCS, Rozzano, Milan, Italy
13. Humanitas Clinical and Research Center – IRCCS, Artificial Intelligence Research, Rozzano, Italy
14. Kansas City VA Medical Center, Gastroenterology and Hepatology, Kansas City, United States

submitted 2.12.2021
accepted after revision 20.7.2022
published online 22.7.2022

Bibliography
Endosc Int Open 2022; 10: E1474–E1480
DOI 10.1055/a-1907-6569
ISSN 2364-3722

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

Corresponding author
Marco Spadaccini, Humanitas Research Hospital and University, Via Manzoni 56, 20089 Rozzano (Milano) Italy
Fax: +390282242595
marco.spadaccini@humanitas.it

Supplementary material is available under https://doi.org/10.1055/a-1907-6569

ABSTRACT

Background and study aims Artificial intelligence (AI) is set to impact several fields within gastroenterology. In gastrointestinal endoscopy, AI-based tools have translated into clinical practice faster than expected. We aimed to evaluate the status of research for AI in gastroenterology while predicting its future applications.

Methods All studies registered on Clinicaltrials.gov up to November 2021 were analyzed. The studies included used AI in gastrointestinal endoscopy, inflammatory bowel disease (IBD), hepatology, and pancreatobiliary diseases. Data regarding the study field, methodology, endpoints, and publication status were retrieved, pooled, and analyzed to observe underlying temporal and geographical trends.

Results Of the 103 study entries retrieved according to our inclusion/exclusion criteria, 76 (74 %) were based on AI application to gastrointestinal endoscopy, mainly for detec-

* These authors share first authorship.
** These authors share senior authorship.

tion and characterization of colorectal neoplasia (52/103, 50 %). Image analysis was also more frequently reported than data analysis for pancreaticobiliary (six of 10 [60 %]), liver diseases (eight of nine [89 %]), and IBD (six of eight [75 %]). Overall, 48 of 103 study entries (47 %) were interventional and 55 (53 %) observational. In 2018, one of eight studies (12.5 %) were interventional, while in 2021, 21 of 34 (61.8 %) were interventional, with an inverse ratio between observational and interventional studies during the study period. The majority of the studies were planned as single-center (74 of 103 [72 %]) and more were in Asia (45 of 103 [44 %]) and Europe (44 of 103 [43 %]).

Conclusions AI implementation in gastroenterology is dominated by computer-aided detection and characterization of colorectal neoplasia. The timeframe for translational research is characterized by a swift conversion of observational into interventional studies.

Introduction

The adoption and implementation of digital health in gastroenterology have occurred more quickly than was expected. This has been primarily related to the application of artificial intelligence (AI) utilizing deep learning for image analysis in the field of gastrointestinal endoscopy. AI for polyp detection has been proven effective in standalone performance studies as well as in randomized clinical trials. In a very short time, several products have been approved by regulatory agencies and introduced into clinical practice [1]. Similar algorithms are now available to detect early cancer in upper gastrointestinal endoscopy and small bowel endoscopy [2–4]. More recently, algorithms for neoplasia characterization, mainly in the lower gastrointestinal field, have become available, opening the door to AI-assisted optical diagnosis [5].

In addition to the technology's widespread use in gastrointestinal endoscopy, artificial neural network applications have expanded to data analytics that may lead to a more precise or personalized diagnostic and therapeutic process based on a multidimensional assessment of disease severity and predictors of either prognosis or therapeutic responses [6–8]. This may apply to inflammatory bowel disease (IBD), liver diseases, and other immune-related or oncological disorders of the gastrointestinal tract.

The primary reason for the subitaneous translation of AI in clinical practice is represented by the fact that these plug-in software algorithms only require the generation and validation of a mathematical algorithm. This represents a major difference with hardware innovations, such as new endoscopic platforms or advanced imaging, and new drugs that tend to appear relatively infrequently in a 5– to 10-year span. However, the development of AI algorithms still requires the collection of cases and supervised annotation, which is a time-consuming process that may require several years to be finalized. For this reason, it is relevant to understand how this translational process works. For instance, observational studies are needed to collect large and robust databases to train AI algorithms. On the other hand, prospective interventional trials are requested before such algorithms can be clinically implemented. Clusterization of studies on one specific topic may be expected to lead to rapid implementation in clinical practice, while some other fields may be marginalized due to paucity of data or challenges in algorithm generation.

To predict the future of AI in gastroenterology, we systematically reviewed all the studies registered on the clinicaltrials.gov platform, including the main field and purposes.

Methods

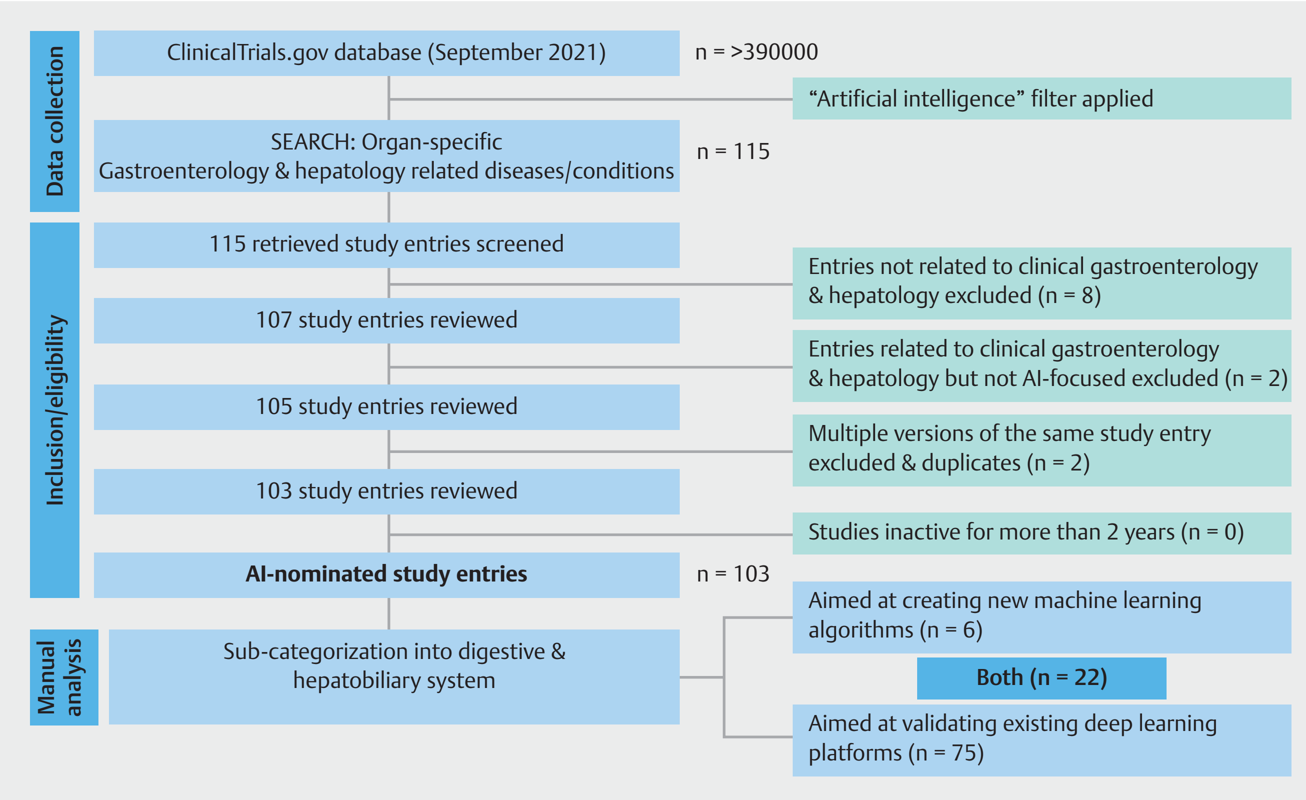
This systematic review was based on data retrieved from Clinicaltrials.gov (https://clinicaltrials.gov), a free database listing 390,330 studies (November 2021). This database has been widely used in the past when conducting systematic reviews to analyze research trends in gastroenterology (such as in IBD) [9, 10] and daily endoscopic practice [11–13], forecasting future research efforts and drawing patient-centered management plans or guideline proposals. Search-filter strategy, data extraction sub-classifications, and study nomination process are reported in the **Appendix**.

Inclusion and exclusion criteria

For our study, only study entries that focused on use of AI-based applications in the field of gastroenterology/hepatology were selected. The term AI, formally defined by Medical Subject Headings (MeSH) in 1986 and elaborated upon further in the **Appendix**, refers to any task completed by a computer system that would ordinarily be achieved using human intellect and reasoning. The terms machine learning and deep learning fall under the broad class of AI as per MeSH definitions. Study entries that analyzed machine learning algorithms to produce models that predict disease progression, prognosis, medical treatment outcomes, and/or improve the quality of current clinical practices in the field were included. Study entries that utilized AI but had no relevance to the field and vice versa were excluded. If multiple versions of the same entry existed, only the most recent version of it was included, and those prior excluded. Those studies that remained inactive for more than 2 years were automatically excluded. No language restrictions were needed during the selection process as all study entries registered to ClinicalTrials.gov were in English.

Data extraction

Each study entry was searched for the following information: 1) geographical setting; 2) funding (academic/hospital); 3) year of registration; 4) field of application (gastrointestinal endoscopy/IBD/liver disease/other); 5) type and design of study (interventional/observational, single-center/multicenter); 6) patient



▶ **Fig. 1** Flowchart of the systematic review.

population (number, age); 7) state of recruitment (completed/ongoing/not started); 8) study-endpoints; and 9) study status (published/unpublished). One-dimensional frequency distributions (absolute, relative) were determined for the analyzed study characteristics. Dataset quantitative acquisition, processing, and statistical evaluation were carried out using Microsoft Excel software for Microsoft Windows.

Results

Study characteristics

An initial literature search using the Clinicaltrials.gov results database identified 115 related study records. Eight study entries were excluded as unrelated to gastroenterology, two due to lack of AI technology, while one study was excluded as a duplicate. Thus, 103 study entries were included in our analysis. ▶ **Fig. 1** shows the methodological process used by the authors to select study entries. Twenty-one of 103 (20 %) and 27 of 103 (26 %) had already closed or completed enrollment at the time of the search, while 49 of 103 (48 %) were open (status not reported for six study entries). The study end points and area of focus of the individual AI studies are summarized in ▶ **Table 1** and **Supplementary Table 2** respectively. **Supplementary Table 1** shows the publication status of completed studies. **Supplementary Fig. 1** represents the recruitment age bracket for the selected study entries.

Study type and design

Among the 103 study entries, 48 (47 %) were planned as interventional and 55 (53 %) as observational studies. In detail, 30 of 48 (63 %) interventional trials were randomized, 16 of 48 (33 %) were N/A (non-applicable) and two of 48 (4 %) were non-randomized trials. Of the 30 randomized trials, 29 (97 %) had a parallel design, while one of 30 (3 %) was based on a factorial assignment. In the non-randomized interventional group, both studies had a parallel assignment. The observational studies were predominantly cohort-type (34/55 [62 %]), case-only (5/55 [9 %]), and three of 55 (5 %) were case control, followed by other methodologies in 13 of 55 (24 %). The time frame of the observational studies was prospective in 38 of 55 (69 %), retrospective in nine (16 %), cross-sectional in one (2 %), and other in seven (13 %). The study types and designs are illustrated in **Supplementary Fig. 2.**

Study location and funding source

According to the retrieved records, the study was based in Asia in 45 of 103 cases (44 %), in Europe in 44 of 103 (43 %), United States 10 of 103 (10 %), and North America four of 103 (4 %). ▶ **Fig. 2** depicts the geographical distribution of registered study protocols across various countries worldwide. The majority of the studies (74/103 [72 %]) were planned as single-center. Hospitals and universities were the primary funding sources for trials conducted in Asia 43 of 45 (96 %), Europe 35 of 44 (80 %), and North America four of four (100 %), whereas the primary

▶ **Table 1** End points of the included studies.

Main topics of the study entries		No. studies
Upper gastrointestinal endoscopy		
Esophagus	Esophageal cancer detection: Squamous cell neoplasia-focused; incorporating probe based confocal laser endomicroscopy (pCLE)	2
	Barrett's early detection-focused; incorporating image retrieval methods	1
	Esophageal cancer treatment response prediction: Predicting a complete pathological response to neoadjuvant therapy; incorporating radiomics	1
	GERD assessment: Incorporating near focus narrow band imaging AI	1
Stomach	Gastric precancer/polyp/neoplasm detection	10
	Magnetic controlled capsule endoscopy: AI platforms developed for MCCE, quality assessments	3
	Clinical decision support system: for upper gastrointestinal cancer care	1
	Gastric cancer (GC) multiomics: AI diagnostic algorithm; incorporating multiomics to characterize advanced GC in Europe, Latin American and Carribean populations	1
Lower gastrointestinal endoscopy		
	Detection polyp/neoplasm	22
	Detection & characterization polyp/neoplasm	21
	Characterization polyp/neoplasm	9
	Bowel preparation quality	3
	Colorectal cancer biomarkers and multiomics evaluation application of AI to identify blood and stool biomarkers to detect early colorectal cancer (Freenome)	2
Small bowel	Small bowel bleeding – capsule endoscopy: AI assistance in video reading	1
Hepatology		
	Hepatocellular carcinoma detection and characterization: Incorporating imaging phenomics to create signatures for various HCC types (iBiopsy)	2
	Hepatocellular carcinoma prognostication: Incorporating MRI radiomics to predict the prognosis of early-stage HCC after minimally invasive treatment	2
	Hepatocellular carcinoma screening: Incorporating DL to clinical, biological, elastographic and ultrasonic parameters to risk stratify hepatocarcinogenesis in non-tumor liver parenchyma	1
	Hepatobiliary disease – ocular association: Incorporating DL to predict hepatobiliary diseases from ocular images	1
	Metastatic liver disease in colon cancer: Developing AI based software to predict metastatic liver nodules in patients with colorectal cancer	1
	Non-alcoholic fatty liver disease & non-alcoholic steatohepatitis: Incorporating AI to differentiate both and to stage fibrosis	1
	Polycystic liver disease: Developing a CNN for automated liver contour; segment detection in polycystic liver	1

funding source for the United States was from industry in five of 10 (50 %). The various funding sources for each location are shown in **Supplementary Fig. 3**.

AI field of application

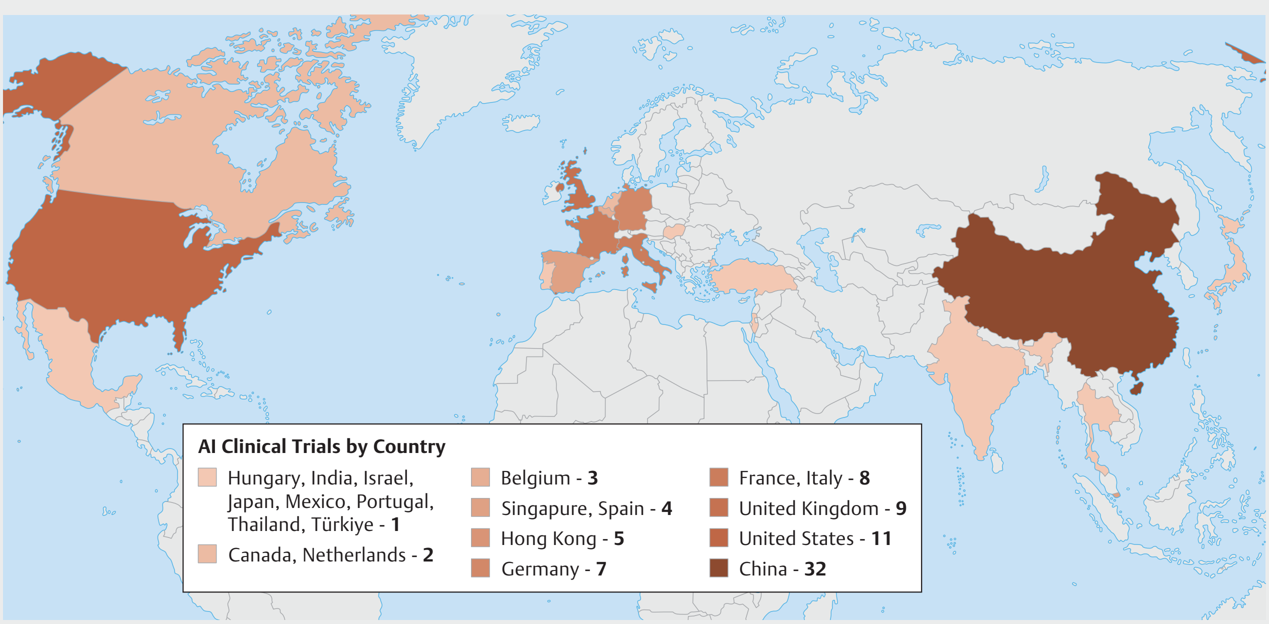
The area of focus of AI research appeared to be gastrointestinal endoscopy in 76 of 103 study entries (74 %), followed by pancreato-biliary diseases in 10 of 103 (10 %), hepatology in nine of 103 (9 %), and IBD in eight of 103 (8 %). Among the records in gastrointestinal endoscopy, colorectal polyp detection was the most prevalent researched area in 43 of 76 (57 %), followed by colorectal polyp characterization in 30 of 76 (39 %). ▶ **Fig. 3** shows the main field of the proposed application for AI in gastroenterology.

Image analysis was more commonly studied as opposed to data analysis in the area of pancreaticobiliary diseases (6/10 [60 %]), hepatology (8/9 [89 %]), and IBD (6/8 [75 %]), as shown in **Appendix Figure.** ▶ **Fig. 4** shows the number of study entries by year for the entire period of 2007 to 2021. In 2018, seven of eight studies (87.5 %) were observational, while in 2021, 21 of

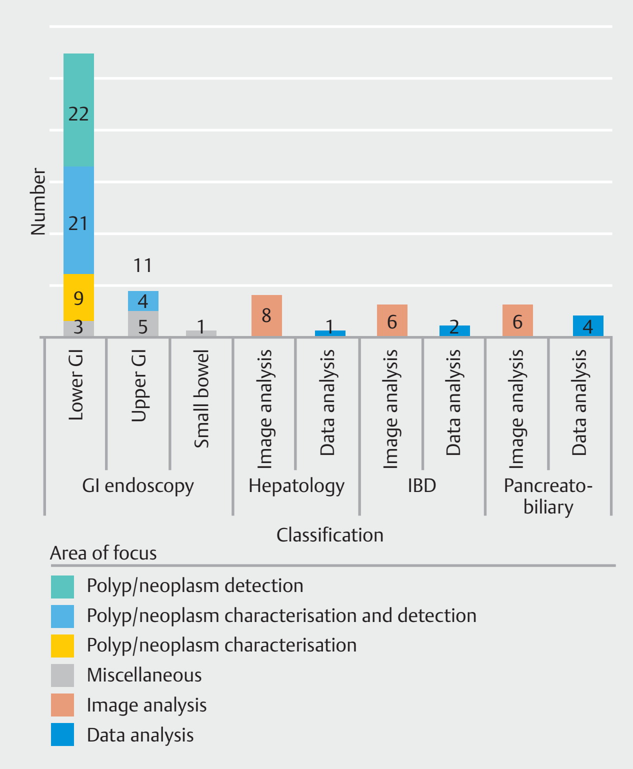
▶ **Table 1** (Continuation)

Main topics of the study entries		No. studies
Pancreaticobiliary		
	Pancreatic neoplasm histology: AI use for rapid on-site evaluation & automated counting of Ki-67 in biopsy samples of neuroendocrine neoplasias respectively.	2
	ERCP navigation system: AI application to biliary stricture navigational instructions for guidewire direction and stent placement	2
	Acute pancreatitis (AP): AI application to determining severity of AP	1
	Pancreatic disease biomarker evaluation: Pairing AI with biomolecular analyses of markers (Berg's Interrogative Biology Platform)	1
	Pancreatic neoplasm screening: AI-based surveillance to predict early pancreatic cancer using health records and big data	1
	Endoscopic ultrasound for pancreatic cancer staging: Lymph node metastases detection and characterization DL algorithm	1
	EUS navigation system: DL-based real-time scanning of the pancreas for lesion detection	1
	Choledocholithiasis prediction model: Symbolic regression applied to symptomatology, biochemical and imaging parameters	1
IBD		
	Crohn's and ulcerative colitis: AI application to assess disease severity via endoscopic images, Raman spectroscopy, chronic pain profiling, histology and radiomics	4
	Ulcerative colitis: DL application to automate evaluation of inflammatory activity using pCLE, red density and other big data, respectively	3
	Crohn's disease: AI applied to signals obtained from digital wearables to forecast transition of symptoms when in stress.	1

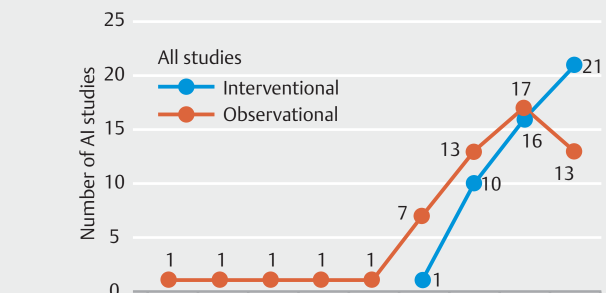
GERD, gastroesophageal reflux disease; AI, artificial intelligence; HCC, hepatocellular carcinoma; MRI, magnetic resonance imaging; DL, deep learning; CNN, convolutional neural network; ERCP endoscopic retrograde cholangiopancreatography; EUS, endoscopic ultrasound; IBD, inflammatory bowel disease.



▶ **Fig. 2** Geographical location of the study entries. The map was created with the help of free software https://www.mapchart.net/world.html



▶ **Fig. 3** Main field of proposed application for AI in gastroenterology.



	2007	2008	2014	2016	2017	2018	2019	2020	2021
Interventional						1	10	16	21
Observational	1	1	1	1	1	7	13	17	13

▶ **Fig. 4** AI studies in the field of gastroenterology by year of registration.

34 (61.8 %) were interventional, with a progressive increase of the study entries by year and inversion of the ratio between observational and interventional studies over time. **Supplementary Table 3** shows the number of studies in which such roles were created, validated or performed in their objectives. Of the 103 studies, 73 % validated existing AI-based platforms and algorithms (75/103), while only 6 % set out to create new ones (6/103). Twenty-one percent (22 of 103) of the studies aimed to do both. Their trend from 2007 to 2021 is illustrated in **Supplementary Fig. 4**.

Discussion

According to our study, research efforts surrounding AI in gastroenterology were primarily focused on detecting and characterizing colorectal neoplasia. At the same time, the use of machine learning for personalized medicine received far less attention and was found to be applied primarily in hepatology. In addition, the time of conversion between observational and interventional study entries was approximately 2 years, suggesting a fast translation of AI research in the clinical setting.

There may be three main reasons for the dominance of colorectal neoplasia as the primary driver of AI-based research. First, deep learning developed for non-medical imaging, i. e., face/object-recognition software, coupled with high-performing graphical interfaces, was technologically optimized for real-time endoscopy. However, it could be argued that this also applies to upper gastrointestinal neoplasia. Thus, the second reason may be the high prevalence of colorectal polyps with a polyp-to-patient ratio up to 3:1, facilitating the collection of a critical volume of cases needed for training an adequate deep learning system. We can estimate that thousands of frames must be annotated to train and thoroughly test a reliable deep learning model for polyp detection (i. e., computer-aided diagnosis for detection [CADe]) and characterization (computer-aided diagnosis for characterization [CADx]). In this regard, upper gastrointestinal neoplasia, such as Barrett's esophagus or early gastric cancer, may be penalized by its very low prevalence outside a tertiary center. Of note, there was a clear tendency for a single-center research setting, preventing the coalescence of multicenter/national databases that may be required for less prevalent disorders. Third and finally, CADe implementation exploits the clinical relevance of the conversion of any increase in adenoma detection rate in additional colorectal cancer prevention, as well as the progressive expansion of population-based colorectal cancer screening programs in several Western and Eastern countries.

The progressive inversion of the gradient between observational and interventional studies was expected. In the observational phase, databases are prospectively collected to train and test AI algorithms in an artificial setting, i. e., standalone performance. In the interventional phase, such algorithms are validated against the reference standard in a clinical setting through randomized or sequential trials. On the other hand, what was somewhat unexpected was the very short time of conversion between the two phases of AI translation in clinical practice, which is well in line with the subitaneous appearance in the gastrointestinal endoscopy market of several devices approved by regulatory agencies for colorectal polyp detection/characterization. Thus, AI research appears as one of the fastest channels to shift innovation from bench to bedside.

Disappointingly, only a limited number of studies on personalized medicine based on the use of machine learning were documented during the timeframe of our analysis. This represents a substantial difference between oncology and genetics, where most AI models are aimed at patient prognosis or response prediction. A possible reason for the small number is the difficulties faced in collecting large enough databases for rare and infrequently encountered diseases, such as liver or pancreatic cancer or IBD [14]. In fact, most of the study entries in these fields were related to image analysis, irrespective of whether that was endoscopy-based or using ultrasound or cross-sectional imaging.

The main limitation of our analysis is represented by the fact that a study entry, simply put, does not necessarily indicate that the study will be executed, finalized and published. However, given that a substantial proportion of databases were already completed, a relevant proportion of published studies would support the validity of our data.

Conclusions

In conclusion, our analysis shows the dominance of CADe/CADx for colorectal neoplasia for AI research in gastroenterology, as well as the limited time span required for its conversion into clinical practice, mirroring what is happening in the gastrointestinal endoscopy market. A different research approach, including a possible lead of scientific societies, is required for AI application to rarer and/or clinically oriented fields.

Competing interets

The authors declare that they have no conflict of interest.

References

[1] Hann A, Troya J, Fitting D. Current status and limitations of artificial intelligence in colonoscopy. United Eur Gastroenterol J 2021; 9: 527–533

[2] Zippel C, Bohnet-Joschko S. Rise of clinical studies in the field of machine learning: a review of data registered in ClinicalTrials.gov. Int J Environ Res Public Health 2021; 18: 5072

[3] Chen H, Sung JJY. Potentials of AI in medical image analysis in Gastroenterology and Hepatology. J Gastroenterol Hepatol 2021; 36: 31–38

[4] Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. World J Gastroenterol 2019; 25: 1666–1683

[5] Bisschops R, East JE, Hassan C et al. Advanced imaging for detection and differentiation of colorectal neoplasia: European Society of Gastrointestinal Endoscopy (ESGE) Guideline – Update 2019. Endoscopy 2019; 51: 1155–1179

[6] van der Sommen F, Zinger S, Curvers WL et al. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. Endoscopy 2016; 48: 617–624

[7] Swager A-F, van der Sommen F, Klomp SR et al. Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy. Gastrointest Endosc 2017; 86: 839–846

[8] Hirasawa T, Aoyama K, Tanimoto T et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. Gastric cancer Off J Int Gastric Cancer Assoc Japanese Gastric Cancer Assoc 2018; 21: 653–660

[9] Caballol B, Gudiño V, Panes J et al. Ulcerative colitis: shedding light on emerging agents and strategies in preclinical and early clinical development. Expert Opin Investig Drugs 2021; 30: 931–946

[10] Bonovas S, Fiorino G, Allocca M et al. Biologic therapies and risk of infection and malignancy in patients with inflammatory bowel disease: a systematic review and network meta-analysis. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc 2016; 14: 1385–1397.e10

[11] Hafner S, Zolk K, Radaelli F et al. Water infusion versus air insufflation for colonoscopy. Cochrane Database Syst Rev 2015; 5: CD009863

[12] Wei Y, Shen G, Yang Y et al. Inspection and polypectomy during both insertion and withdrawal or only during withdrawal of colonoscopy? A protocol for systematic review and meta analysis Medicine (Baltimore) 2020; 99: e20775

[13] Lim S, Hammond S, Park J et al. Training interventions to improve adenoma detection rates during colonoscopy: a systematic review and meta-analysis. Surg Endosc 2020; 34: 3870–3882

[14] Chen B-B. Artificial intelligence in pancreatic disease. World J Gastroenterol 2020; 1: 19–30

Koleth Glenn et al. Artificial intelligence in … Endosc Int Open 2022; 10: E1474–E1480 | © 2022. The Author(s).