



Position statement on priorities for artificial intelligence in GI endoscopy: a report by the ASGE Task Force

Tyler M. Berzin, MD,^{1,*} Sravanthi Parasa, MD,^{2,*} Michael B. Wallace, MD,³ Seth A. Gross, MD,⁴ Alessandro Repici, MD,⁵ Prateek Sharma, MD⁶

Boston, Massachusetts; Seattle, Washington; Jacksonville, Florida; New York, New York; Kansas City, Missouri, USA; Rozzano, Milan, Italy

Artificial intelligence (AI) in GI endoscopy holds tremendous promise to augment clinical performance, establish better treatment plans, and improve patient outcomes. Although there are promising initial applications and preliminary clinical data for AI in gastroenterology, the field is still in a very early phase, with limited clinical use. The American Society for Gastrointestinal Endoscopy has convened an AI Task Force to develop guidance around clinical implementation, testing/validating algorithms, and building pathways for successful implementation of AI in GI endoscopy. This White Paper focuses on 3 areas: (1) priority use cases for development of AI algorithms in GI, both for specific clinical scenarios and for streamlining clinical workflows, quality reporting, and practice management; (2) data science priorities, including development of image libraries, and standardization of methods for storing, sharing, and annotating endoscopic images/video; and (3) research priorities, focusing on the importance of high-quality, prospective trials measuring clinically meaningful patient outcomes. (Gastrointest Endosc 2020;92:951-9.)

Artificial intelligence (AI) is poised to play a crucial role throughout the practice of gastroenterology. Machine learning (ML) is a branch of AI focusing on computer algorithms that can learn from data and perform specific tasks and analyses, often without explicit human programming. Deep learning is an advanced subset of ML that relies on specific algorithms termed *artificial neural networks*. The enthusiasm for medical AI applications has centered

around expectations that these technologies may assist physicians by assimilating, triaging, and interpreting clinical data (laboratory results, electrocardiographic tracings, CTs, etc) to support physician performance in a consistent manner in all phases of clinical care and automate certain burdensome tasks, such as clinical documentation and billing, to allow physicians to refocus on direct patient care and other meaningful professional activities.¹

Abbreviations: AI, artificial intelligence; ASGE, American Society for Gastrointestinal Endoscopy; CADe, computer-aided polyp detection; CADx, computer-aided polyp diagnosis; EHR, electronic health record; ML, machine learning.

DISCLOSURE: The following authors disclosed financial relationships: T. M. Berzin: Consultant for Boston Scientific, Wision AI, and Medtronic; speaker for Medtronic; research support from Wision AI. M. B. Wallace: Consultant for Virgo Inc, Cosmo/Aries, Anx Robotica, Covidien, GI Supply, Endokey, Endostart, and Microtek; research support from Cosmo/Aries, Ninepoint Medical, Fujifilm, Boston Scientific, Olympus, and Medtronic; stock options in Virgo Inc; food/beverage reimbursement from Cook Medical and Synergy Pharmaceuticals. S. A. Gross: Consultant for Olympus. A. Repici: Consultant and advisory board for Medtronic and Fujifilm. P. Sharma: Consultant for Olympus, Boston Scientific, Lumendi, Lucid, Bausch, Fujifilm, and Medtronic; research support from Olympus, Fujifilm, Erbe, Medtronic, U.S. Endoscopy, Ironwood, Cosmo Pharmaceuticals, and CDx Labs. All other authors disclosed no financial relationships.

*Drs Berzin and Parasa contributed equally to this article.

Copyright © 2020 by the American Society for Gastrointestinal Endoscopy 0016-5107/\$36.00

<https://doi.org/10.1016/j.gie.2020.06.035>

Received April 16, 2020. Accepted June 14, 2020.

Current affiliations: Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA (1), Division of Gastroenterology, Swedish Medical Group, Seattle, Washington, USA (2), Department of Gastroenterology, Mayo Clinic Florida, Jacksonville, Florida, USA (3), Division of Gastroenterology and Hepatology, New York University Langone Health, New York, New York, USA (4), Digestive Endoscopy Unit, Humanitas Research Hospital and Department of Biomedical Sciences, Humanitas University, Rozzano, Milan (5), Gastroenterology and Hepatology, Veterans Affairs Medical Center and University of Kansas Medical Center, Kansas City, Missouri, USA (6).

Reprint requests: Tyler M. Berzin, MD, MS, FASGE, Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Ave, Boston, MA 02130.

In GI endoscopy, a wide variety of applications for AI algorithms are being proposed, developed, and, in some cases, validated in clinical trials. Many current AI algorithms for gastroenterology are in the field of computer vision, which refers to technologies that can “see and interpret” visual data, for instance, a live video stream from the endoscope. For example, computer-aided polyp detection (CAdE) and computer-aided polyp diagnosis (CAdx) are 2 applications of AI computer vision for colonoscopy already extensively studied.²⁻⁴ Computer vision software is also being evaluated in GI endoscopy to automatically assess bowel preparation quality, to identify possible areas of bleeding during capsule endoscopy, and to facilitate detection of esophageal squamous cancers and dysplastic Barrett’s esophagus, among innumerable examples.⁵⁻⁹ Beyond endoscopy, computer vision is also being applied quite extensively in abdominal radiology, for instance, to classify and risk stratify pancreatic cysts identified on CT and magnetic resonance imaging, in some cases with algorithms that can integrate other clinical and laboratory data as well.^{10,11}

To focus only on computer vision, however, creates a too-narrow view of the potential for AI in gastroenterology. There is also enthusiasm and some early progress for AI technologies seeking to reduce the tremendous burden of documentation in electronic health records (EHRs), to automate repetitive tasks, to reduce administrative burdens, and to facilitate quality reporting and analytics.

With this in mind, the American Society for Gastrointestinal Endoscopy (ASGE) has formed a Task Force on Artificial Intelligence. A primary goal of the AI task force is to leverage the expertise of the ASGE to set clinical and research priorities for AI applications. The priorities outlined in this document (summarized in [Table 1](#)) are intended to encourage development of AI technology that will be both patient-centered, to improve meaningful clinical outcomes, and physician-centered, to support clinical workflow and efficiency.

PRIORITY USE CASES

Computer vision applications for GI endoscopy

CAdE and CAdx have been the areas of most rapid progress so far in applying AI/computer vision to GI endoscopy.^{2,3} Given the central importance of screening and surveillance colonoscopy for colorectal cancer prevention, continued development and research validation of CAdE and CAdx remains a top priority. Prioritization of additional use cases for computer vision in GI endoscopy must consider the prevalence of the targeted disease state, the potential clinical impact of the proposed algorithm, and the likely solvability by AI. Of the most common GI cancers, colorectal cancers, gastric cancers, and esophageal cancers all have precursor lesions that can be diagnosed by traditional endoscopic modalities. Additionally,

inflammatory bowel disease represents a relatively high-prevalence group of conditions with an elevated risk of colorectal cancer but with dysplastic precursor lesions that are difficult to recognize endoscopically. These priority clinical use cases, with clinical narrative, suggested workflow, key data points, and suggested outputs, are reviewed in [Table 2](#).

The Data Science Institute of the American College of Radiology has taken the step of structuring image libraries around similar use cases (eg, acute appendicitis and cholecystitis), which encourages algorithms to be developed around clinically relevant questions.¹² Each use case provides narrative descriptions and flow charts that specify the healthcare goal of the algorithm, the required clinical input, how it should integrate into the clinical workflow, and how it will interface with users and tools. The “use cases” included in [Table 2](#) represent a starting point for developing a more detailed set of recommendations, with expert input from each subspecialty of gastroenterology.

Tools to streamline endoscopy/clinical workflow

Currently, the work day of a practicing gastroenterologist is filled with hours of computer data entry in the endoscopy suite, outpatient clinic, and hospital. This data entry role has grown immeasurably in the era of electronic medical records and is recognized as an important contributor to physician stress, work dissatisfaction, and the phenomenon commonly labeled “burnout.”^{13,14} Efforts are currently underway to leverage AI to play a role in addressing these new challenges by automatically documenting, managing, and interpreting the vast amounts of data generated by routine medical encounters.¹ Many of these efforts will have direct impact on the practice of gastroenterology.

In the endoscopy suite, AI has the potential to automate endoscopy report generation. Already, computer vision applications have been developed that can recognize and label key endoscopic landmarks, automatically measure withdrawal time, record bowel preparation quality, and, in some cases, even recognize endoscopic tools and interventions.^{5,15,16} When such tools reach a high threshold for reliability, it is foreseeable that at the completion of a procedure, the endoscopist will be able to simply review and sign a nearly complete, automatically generated endoscopy report. Further, if CAdx technology is able to provide optical biopsy specimen data with high accuracy, follow-up interval recommendations can be immediately generated for certain cases. For example, although the current workflow for colonoscopy involves numerous steps ([Fig. 1A](#)), including awaiting pathology follow-up and subsequently generating results letters and reminders, automated reporting integrated with CAdx has the potential to vastly simplify workflow such that the entire care episode is completed and documented by the end of the procedure, including polyp optical biopsy sample results and follow-up recommendations ([Fig. 1B](#)).

TABLE 1. Development priorities for artificial intelligence applications in GI endoscopy practice

Category	Specific needs
Priority use cases	Computer vision applications for GI endoscopy Tools to streamline endoscopy/clinical workflow Decision support and quality assessment/reporting Practice management applications
Data science priorities	Endoscopy image/video library: Endonet Standardized methods for storing, sharing, and labeling endoscopic images/videos
Research priorities	Validation of computer vision tools for GI endoscopy in prospective clinical trials Identification of clinically relevant outcome measures for AI interventions

Voice-powered virtual assistants such as Apple's Siri, Amazon's Alexa, and Google's Assistant, use AI speech recognition that has now improved to the point of human accuracy in transcribing conversation. Already, AI technologies underpin many current dictation programs used in clinical medicine. An important step beyond voice-driven data entry is ambient clinical intelligence, which refers to technologies installed to monitor and react to visual or auditory input in the clinical environment. One application of ambient clinical intelligence is computer-aided documentation in which, for instance, a computer can hear, interpret, and document the clinical encounter without the need for additional physician dictation or documentation.¹⁷ An additional important step will be higher-order voice-driven interaction with EHR so that hands-free data retrieval and analysis becomes a clinical reality (eg, "Show me the hemoglobin trend over the last year" or "Pull up the last 2 colonoscopy reports on screen").

Decision support and quality assessment and reporting

Clinical decision support systems that collect data from EHRs, medical literature, guidelines, regulatory warnings, and other public information can aid in decision-making. AI-enabled clinical knowledge retrieval tools could also be accessed through the same conversational systems that allow clinicians to retrieve patient information from the EHR. Through techniques such as information extraction, natural language processing (NLP), automatic summarization, and deep learning, AI has the potential to transform static narrative articles into patient-specific, interactive visualizations of clinical evidence that could be seamlessly accessed within the EHR. Similarly, quality reporting (adenoma detection rate, bowel preparation quality) is a major practical challenge for many gastroenterology groups and hospital systems, because most EHRs struggle to deliver helpful outcome measures at the click of a button. AI tools, leveraging NLP, represent a potential pathway toward dramatically simplifying extraction and reporting of gastroenterology practice quality measures.^{18,19}

Practice management applications

The administrative complexity and cost of managing a gastroenterology practice, both in private practice and hospital-based settings, has increased substantially over time. Time-intensive administrative coordination occurs around scheduling, billing, and payment, and this cycle of work occurs at every phase: before, during, and after each patient encounter. ML models are needed that can work alongside administrative personnel to perform mundane, repetitive tasks in a more efficient, accurate, and unbiased fashion. AI software tools that automate many of these administrative activities have already been proposed and in some specific use-cases implemented successfully.²⁰

Most health plans and pharmacy benefit managers require prior authorization of devices, durable equipment, laboratory data, and procedures. Automating this process, with human review, can reduce biased decisions and improve the speed, consistency, and quality of decisions. AI could be applied several different ways in prior authorization. For example, AI could simply be used to sort cases to the appropriate level of reviewer (eg, nurse practitioner, physician advisor, medical director). Additionally, AI could identify and highlight the specific, relevant information in long documents or narratives to produce information regarding estimated costs or benefit-to-risk assessment for the physician or patient. To address the challenge of prior authorizations, AI with NLP could extract relevant data from the EHR to provided automated justification for a prior authorization claim.

Patient scheduling in most gastroenterology practices has generally followed a "1 size fits all" approach, with standardized appointment durations for initial visits, follow-up visits, and various procedure types. This approach may not take into account significant variations in complexity and medical needs for each patient. Risk-adjusted paneling, using an ML algorithm, has been trialed in primary care over the last several years.²¹ When this approach is used, the number of scheduled patients and visit length can be adjusted automatically based on predicted or known complexity. Similar methods may be helpful for scheduling a mixed panel of gastroenterology patients with widely varying complexity and medical needs, both

TABLE 2. Priority clinical use cases for AI algorithm development

Use case 1: Colon polyp detection Rationale: Adenomatous colon polyps are precursor lesions to colorectal cancer. Identification and removal of adenomatous colon polyps is an effective approach for colorectal cancer prevention. Some colon polyps may not be seen or recognized by the endoscopist. An AI tool may help visually identify colon polyps to improve polyp detection during colonoscopy. Workflow description: During colonoscopy, the live video image is sent to the AI application and analyzed in real time. If a colon polyp is present on the screen, the AI application delivers a signal (screen alert, location box etc) to an endoscopy monitor, alerting the endoscopist to the location of the suspected polyp. The endoscopist then determines if a polyp is present and if removal is indicated. Procedure type: Colonoscopy Required data elements: Imaging/light type [white light, narrow-band imaging, LCI, chromoendoscopy], indication [screening or surveillance], age [years], pathologic diagnosis for each polyp [adenoma, hyperplastic, sessile serrated, inflammatory, other], size for each polyp [mm]. Additional suggested data elements: History of colon surgery [yes or no], location for each polyp [cecum, right-sided colon, transverse colon, left-sided colon, rectosigmoid], morphology for each polyp [eg, Paris classification], colonoscope manufacturer and model. Primary outputs Definition: polyp detection Data type: categorical Value set: 1, present; 2, absent
Use case 2: Colon polyp diagnosis Rationale: During screening or surveillance colonoscopy, endoscopists identify and remove nearly all colon polyps that are identified. The type of polyp, adenoma vs hyperplastic for instance, determines clinical decision-making, including surveillance colonoscopy intervals. It is difficult to diagnose polyp type with a high level of certainty by visual appearance; therefore, polyps are sent for histopathologic analysis for definitive identification. Diagnosing polyp type by AI optical biopsy sample would permit a resect and discard approach for some polyps, obviating the need for histopathologic analysis. Workflow description: During colonoscopy, the live video image of a polyp is sent to the AI application and analyzed in real time. The AI application then displays an immediate prediction of polyp type. Procedure type: Colonoscopy Required data elements: Imaging/light type [white light, narrow-band imaging, blue-light imaging, i-SCAN, LCI, chromoendoscopy, magnification], pathologic diagnosis for each polyp [adenoma, hyperplastic, sessile serrated, inflammatory, other]. Additional suggested data elements: Morphology for each polyp [eg, Paris classification], colonoscope manufacturer and model. Primary outputs Definition: polyp subtype Data type: categorical Value set: 1, hyperplastic; 2, adenoma; 3, sessile serrated polyp; 4, inflammatory; 5, other
Use case 3: Detection of dysplasia in Barrett's esophagus Rationale: Barrett's esophagus is a precursor lesion to esophageal adenocarcinoma. The presence of regions of dysplasia (low grade or high grade) in Barrett's esophagus increases the risk of progression and guides treatment considerations. The appearance of dysplasia may be subtle, and segmental biopsy samples may not detect patchy dysplasia. Optical identification and diagnosis of dysplasia would guide treatment decisions during endoscopy for Barrett's esophagus. Workflow description: During upper endoscopy, the live video image is sent to the AI application and analyzed in real time. The AI application delivers a signal (screen alert, location box, etc) to an endoscopy monitor, alerting the endoscopist to the area of suspected dysplasia. This area can then be biopsy sampled or directly treated (endoscopic resection or ablation therapy). Procedure type: Upper endoscopy Required data elements: Imaging/light type [white light, narrow-band imaging, blue-light imaging, i-SCAN, LCI, chromoendoscopy, magnification], pathologic diagnosis for each visualized segment of Barrett's esophagus. Additional suggested data elements: Size of lesion (mm), location of lesion, morphology of lesion, endoscope manufacturer and model. Primary outputs Definition: Barrett's esophagus subtype Data type: categorical Value set: 1, nondysplastic; 2, low-grade dysplasia; 3, high-grade dysplasia; 4, adenocarcinoma

(continued on the next page)

TABLE 2. Continued

Use case 4: Detection of gastric cancer precursor lesions and early gastric cancer
Rationale: Gastric cancer is among the most prevalent GI cancers worldwide. Development of gastric cancer is a multistep process that involves the development of gastritis, atrophy, intestinal metaplasia, dysplasia, and finally adenocarcinoma.
Workflow description: During upper endoscopy, the live video image is sent to the AI application and analyzed in real time. The AI application delivers a signal (screen alert, location box, etc) to an endoscopy monitor, alerting the endoscopist to the area of suspected intestinal metaplasia, dysplasia, or early gastric cancer. The area of interest can then be biopsy sampled and/or resected endoscopically.
Procedure type: Upper endoscopy
Required data elements: Imaging/light type [white light, narrow-band imaging, blue-light imaging, i-SCAN, LCI, chromoendoscopy, magnification], pathologic biopsy sample diagnosis for each targeted area in the stomach.
Additional suggested data elements: Size of lesion (mm), location of lesion, morphology of lesion, endoscope manufacturer and model.
Primary outputs Definition: gastric cancer sequence Data type: categorical Value set: 1, gastritis; 2, gastric atrophy; 3, intestinal metaplasia; 4, gastric dysplasia; 5, adenocarcinoma.
Use case 5: Detection of dysplasia in IBD
Rationale: Patients with long-standing IBD have a higher risk of colorectal cancer than the general population. Colonoscopy surveillance to evaluate for precancerous changes (dysplasia) is recommended. The appearance of dysplasia may be subtle, and segmental or targeted biopsy samples may not detect patchy dysplasia. Optical identification and diagnosis of dysplasia would guide treatment decisions during surveillance colonoscopy for IBD.
Workflow description: During colonoscopy, the live video image is sent to the AI application and analyzed in real time. The AI application delivers a signal (screen alert, location box, etc) to an endoscopy monitor, alerting the endoscopist to the area of suspected dysplasia and predicted diagnosis. This area can then be biopsy sampled and/or resected endoscopically.
Procedure type: Colonoscopy
Required data elements: Imaging/light type [white light, narrow-band imaging, blue-light imaging, i-SCAN, LCI, chromoendoscopy, magnification], pathologic diagnosis for each visualized segment of colon tissue.
Additional suggested data elements: Size of lesion (mm), location of lesion, morphology of lesion, colonoscope manufacturer and model.
Primary outputs Definition: IBD dysplasia Data type: categorical Value set: 1, nondysplastic; 2, low-grade dysplasia; 3, high-grade dysplasia; 4, adenocarcinoma.

AI, Artificial intelligence; LCI, linked color imaging; IBD, inflammatory bowel disease.

in the office and in the endoscopy suite or ambulatory surgical center.

Routine communication with patients outside of the medical encounter, from triaging phone inquiries to delivering reminder phone calls, represents another area ripe for innovation. In many other fields involving customer phone inquiries, “chatbots” have been deployed successfully to provide initial phone triage and basic information. Chatbots are NLP systems that can respond and interact with human-like conversation.²² Although this technology is still in a developmental phase, health chatbots could potentially increase access to health care, improve doctor–patient and clinic–patient communication, and/or assist with managing the increasing demand for health services such as via remote testing, medication compliance monitoring, or teleconsultations.^{23,24} Health chatbot capabilities will include activities such as setting up health-related reminders and booking appointments, and for gastroenterology could play a role in answering routine questions such as bowel preparation queries.²⁵

Finally, supply chain management is an important challenge for gastroenterology practices. Successful supply chain management integrates a variety of data streams including product life cycle data, ordering pattern, patients scheduled, and so on. AI has already been successfully applied to streamline supply chain management in other fields.²⁶ Similar applications applied to ambulatory surgical centers and hospitals could assist by forecasting clinical demand and automating optimization and stocking of the tools and equipment required to operate an endoscopy suite.

DATA SCIENCE PRIORITIES

Endoscopy image and video library: Endonet

The development of publicly available labeled imaging libraries of chest radiographs, CTs, and other modalities has helped spur research and rapidly propel the field of radiology AI forward.^{27,28} There appears to be a need for a comprehensive Endonet image library for GI endoscopy. An image library maintained by a professional

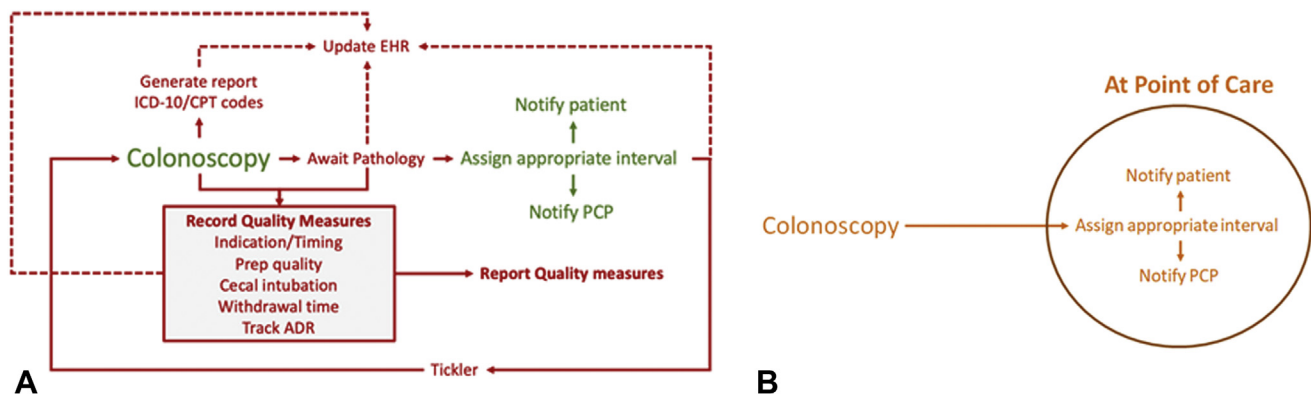


Figure 1. A, Current workflow for screening colonoscopy (Source, W. Karnes, used with permission). **B,** Proposed workflow for screening colonoscopy in the setting of computer-aided polyp diagnosis with automation of data entry and analytics (Source, W. Karnes). *EHR*, Electronic health record; *ICD-10*, International Statistical Classification of Diseases and Related Health Problems, 10th Revision; *CPT*, current procedural terminology; *PCP*, primary care provider; *ADR*, adenoma detection rate.

society (such as the ASGE) or institutional partners could have 2 key components: 1), a freely available image library to accelerate computer vision innovation for the field of GI endoscopy and 2), embargoed image libraries available specifically for product performance testing and validation of algorithms by companies, regulatory bodies, or physician groups and hospitals. Curated libraries of this type may be particularly important for facilitating premarket and postmarket regulatory review pathways for AI algorithms in GI endoscopy. The use cases included in Table 2 provide a framework for how such a library could be organized. The following are some of the additional steps needed for such an Endonet dataset.

Standardized methods for storing, sharing, and labeling endoscopic images and videos. A major barrier to progress in ML in GI endoscopy is the relative absence of high-quality, labeled images for training of AI algorithms.²⁹ To overcome this barrier, there is an urgent need to develop standards and recommendations regarding appropriate formats and quality standards for endoscopy image and video, along with accepted protocols for categorizing such images, storing metadata, and transferring and storing images while protecting private health information (Table 3). Recently, the Heart Rhythm Society and Consumer Technology Association released a collaborative document providing guidance on wearables that monitor cardiovascular biometrics.³⁰ Development of equivalent standards for GI endoscopy will require similar collaboration between gastroenterologists, ML scientists, and industry experts.

Image storage and sharing. Although some progress has been made to overcome many technical challenges of cloud-based sharing of clinical images and reports between care teams at disparate facilities, additional challenges need to be addressed regarding logistical, operational, and regulatory barriers (Table 3). Source data from clinical images captured during endoscopy and ancillary tools used for image enhancement should be used for

image analysis. If compression of the images is performed by the algorithm, then benchmarks are needed for quality ratings that address the suitability of the images for specific tasks. In certain cases, the source data for raw images can be difficult to obtain because these are kept in a proprietary format that impedes data sharing and can raise intellectual property concerns if shared, so industry cooperation often is required.

Of note, there are several challenges specific to the different classes of endoscopic imaging and image-based clinical reports that impact various technical factors for development of an image library (as below). These factors warrant clarity regarding the process of image source, storage, and sharing:

1. The fidelity of preserving the original data and the meta-data (endoscopy report, pathology report) that describes the context of acquisition
2. The way the images were compressed during and after acquisition, for distribution and viewing or review, and for archiving in the database
3. The suitability of devices and software used for display in endoscopy suite, where such AI algorithms will be deployed

In the early 1990s, the field of radiology settled on DICOM (Digital Imaging and Communications in Medicine) as a standard for storing, compressing, and exchanging image data sets and associated metadata. Metadata refers to information included in the DICOM file regarding the image size, dimensions, equipment used, equipment settings, and private health information including clinical history. This standardization, interoperability, and rich metadata was an early foundational element for the recent leaps made by ML applications for diagnostic imaging, and similar approaches are required for GI endoscopy.³¹

Communication protocols. Protocols for seamlessly communicating images and video files between different institutions need to be clarified. For instance, the following scenarios need to have standardized approaches:

TABLE 3. Research opportunities and infrastructure development requirements for collection, storage, and sharing of endoscopy images for multi-institutional collaboration

Area	Current state	Gap in knowledge and infrastructure	Needs
Data availability/data needs for machine learning research	Few large endoscopy image datasets are available. Algorithms are developed most often with single-institution data with no evidence that algorithm outputs will be generalizable to routine clinical practice and free of unintended bias.	Infrastructure changes are needed to allow availability of distributed data sets for use in training, testing, and validating AI algorithms.	Create data sets for AI training and testing based on structured use cases so that similar data sets can be created at multiple institutions and used centralized or on premises; ensure geographic, technical, and patient demographic diversity. Create standards for data use agreements to facilitate data sharing among organizations.
Standards for image and video formats, quality, labeling/annotations, and deidentification	Few annotation models specific to endoscopic images are available; however, standardization is lacking.	No current standards have been established for structured data capture, segmentation, or labeling for endoscopy.	Industry will need to lead the way in developing more robust computing platforms for annotations specific to GI endoscopy (especially for videos) and transfer of annotations, deidentification of the image data, and cybersecurity.
Standards for data sharing	No platforms for data sharing of high-quality standard endoscopic images exist.	Identification and use of the proper standards for AI interoperability have not occurred.	Develop recommendations for identifying proper standards for AI interoperability. Federated learning could be an option to explore.

AI, Artificial intelligence.

1. Sending, querying, and retrieving DICOM objects
2. Sending, querying, and retrieving non-DICOM objects
3. Extracting compressed pixel data or video payload from DICOM objects

Annotation and labeling of endoscopic images and videos. Once high-quality images and videos are collected, current approaches to computer vision require that images are categorized and hand-labeled. For instance, the development of an algorithm for CAdE could require a hand-drawn outline around each colon polyp in an image library numbering tens of thousands of images.^{32,33} There are 2 major hurdles for labeling such large datasets of GI endoscopy images or videos. The first is a lack of standardized approach toward image labeling. Several tools have been proposed to assist in ground truth generation for specific image annotation tools and include Label Me (MIT, Computer Science and Artificial Intelligence Laboratory, Cambridge, Mass, USA),³⁴ VGG image annotator (Visual Geometry Group, Department of Engineering Science, University of Oxford, Oxford, UK),³⁵ and GTCreator.³⁶ However, none of the presently available labeling tools is purpose-built with a physician, let alone a GI endoscopist, in mind. The second challenge is that accurate labeling requires an expert, in this case a gastroenterologist, to review the images and provide the ground truth by hand-annotating and describing the key areas of interest, using standardized terminology and

labels. Labeling still images can be difficult and labor-intensive, but video represents an even greater challenge. For example, in the case of a colonoscopy video analysis, ground truth definition is a highly time-consuming task, requiring clinicians to precisely annotate thousands of images that may be generated by even a short video clip. Finally, the quality of such annotation should be reported, and tools that allow transfer of metadata between images would be helpful.

RESEARCH PRIORITIES

In recent years, a large ecosystem of academic research groups, technology start-ups, and established companies such as Google and Amazon has developed around potential applications of AI in clinical medicine. U.S. regulatory approval has already been achieved for numerous AI algorithms, most focused on radiology but none as yet in the field of GI endoscopy.³⁷

However, despite intense interest and investment in medical AI, high-quality prospective clinical trials are notably lacking to assess the clinical impact of these technologies. In fact, in all of clinical medicine only 6 prospective randomized controlled trials have been published studying the effects of an AI/deep learning intervention in clinical care. The field of gastroenterology has been an early leader in

the completion of randomized prospective trials evaluating AI technologies, and 5 of these 6 randomized controlled trials are in the field of GI endoscopy.^{2,5,38-40}

Optimal study design approaches for clinical trials of AI in colonoscopy have been outlined by Vinsard et al,⁴¹ and the recommendations have important implications for GI endoscopy AI trials in general. Clinically relevant outcome measures must be prespecified based on the AI application investigated. For instance, CAdE studies should use validated outcome measures such as adenoma detection rate, adenomas per colonoscopy, or, in the case of tandem studies, adenoma miss rate. For CADx, the performance thresholds defined by the ASGE Preservation and Incorporation of Valuable Innovations are appropriate to determine if CADx performance is adequate to allow for a resect-and-discard approach for colon polyps.⁴² Other computer vision applications for GI endoscopy will need to follow a similar approach, and future communications by the ASGE Task Force on Artificial Intelligence will seek to consolidate and summarize recommended outcome measures for a variety of potential trial designs in this arena. Guidance for clinical trial designs using AI interventions, relevant to all fields of medicine, is also currently being developed by the SPIRIT group.⁴³

Additionally, it is crucial that the performance of AI algorithms is tested in diverse patient groups and practice settings. In particular, study populations should ideally be distinct (eg, considering geography and ethnicity) from the original algorithm development cohort whenever possible.⁴⁴ AI applications must be validated in real-world clinical settings with varying expertise and experience and using a range of endoscopic equipment made by different manufacturers.

CONCLUSIONS

The next several years will likely be a period of rapid development for AI tools in gastroenterology. It is crucial that as the field advances, there is a focus on technologies that provide real clinical benefits that have been validated in high-quality clinical trials. Equally important is investment in AI technologies that are physician-centered, with an eye toward reversing the steady creep of documentation and administrative work that threatens to move the focus of medicine further away from our patients. The AI applications that are most likely to see rapid adoption in clinical practice are those that ease clinical workflow, simplify or eliminate burdensome tasks, and improve efficiency.

The vast amount of data produced every day in GI endoscopy practice is currently largely inaccessible for the purposes of AI algorithm development and refinement. Reliable, standardized methods to store, organize, and label endoscopic images and video are sorely lacking. Current iterations of EHR technology, including those specific to the endoscopy suite, manage image and video

data in different ways, which is an impediment to the creation of the large, multi-institutional data sets that may be required to make more rapid progress in AI algorithm development for GI endoscopy. The ASGE intends to support the successful integration of clinically relevant AI technologies into GI practice by facilitating the development of standardized approaches for endoscopy data sets and image libraries, by providing guidance for regulatory bodies and industry regarding how best to assess and monitor AI interventions in GI endoscopy, and by directly supporting important research in this field.

REFERENCES

1. Topol EJ. Deep medicine: how artificial intelligence can make health-care human again. 1st ed. New York, NY: Basic Books; 2019.
2. Wang P, Berzin TM, Glissen Brown JR, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019;68:1813-9.
3. Byrne MF, Chapados N, Soudan F, et al. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2019;68:94-100.
4. Mori Y, Kudo SE, Misawa M, et al. Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Ann Intern Med* 2018;169:357-66.
5. Wu L, Zhang J, Zhou W, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut* 2019;68:2161-9.
6. Tsuboi A, Oka S, Aoyama K, et al. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig Endosc* 2020;32:382-90.
7. Thakkar S, Carleton NM, Rao B, et al. Use of artificial intelligence-based analytics from live colonoscopies to optimize the quality of the colonoscopy exam in real-time: proof of concept. *Gastroenterology* 2020;158:1219-22.
8. Hashimoto R, Requa J, Tyler D, et al. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). *Gastrointest Endosc* 2020;91:1264-71.
9. Guo L, Xiao X, Wu C, et al. 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.
10. Corral JE, Hussein S, Kandel P, et al. Deep learning to classify intraductal papillary mucinous neoplasms using magnetic resonance imaging. *Pancreas* 2019;48:805-10.
11. Springer S, Masica DL, Dai Molin M, et al. A multimodality test to guide the management of patients with a pancreatic cyst. *Sci Transl Med* 2019;11:eaav4772.
12. Allen B, Dreyer K. The role of the ACR Data Science Institute in advancing health equity in radiology. *J Am Coll Radiol* 2019;16(4 Pt B):644-8.
13. Gardner RL, Cooper E, Haskell J, et al. Physician stress and burnout: the impact of health information technology. *J Am Med Inform Assoc* 2019;26:106-14.
14. Overhage JM, McCallie D Jr. Physician time spent using the electronic health record during outpatient encounters: a descriptive study. *Ann Intern Med* 2020;172:169-74.
15. Zhou J, Wu L, Wan X, et al. A novel artificial intelligence system for the assessment of bowel preparation (with video). *Gastrointest Endosc* 2020;91:428-35.
16. Samarasekera J, Yu AR, Torralba EJ, et al. Artificial intelligence can accurately detect tools used during colonoscopy: another step forward toward autonomous report writing: Presidential Poster Award: 1075. *Am J Gastroenterol* 2018;113:S619-20.

17. Coiera E, Kocaballi B, Halamka J, et al. The digital scribe. *NPJ Digit Med* 2018;1:58.
18. Raju GS, Lum PJ, Slack RS, et al. Natural language processing as an alternative to manual reporting of colonoscopy quality metrics. *Gastrointest Endosc* 2015;82:512-9.
19. Imler TD, Morea J, Kahi C, et al. Multi-center colonoscopy quality measurement utilizing natural language processing. *Am J Gastroenterol* 2015;110:543-52.
20. Lin SY, Mahoney MR, Sinsky CA. Ten ways artificial intelligence will transform primary care. *J Gen Intern Med* 2019;34:1626-30.
21. Rajkumar A, Yim JW, Grumbach K, et al. Weighting primary care patient panel size: a novel electronic health record-derived measure using machine learning. *JMIR Med Inform* 2016;4:e29.
22. Ivanovic M, Semnic M. The role of agent technologies in personalized medicine: 2018 5th International Conference on Systems and Informatics (ICSAI). 2018:299-304.
23. Hoermann S, McCabe KL, Milne DN, et al. Application of synchronous text-based dialogue systems in mental health interventions: systematic review. *J Med Internet Res* 2017;19:e267.
24. Fadhil A, Gabrielli S. Addressing challenges in promoting healthy lifestyles: the ai-chatbot approach. *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. Barcelona, Spain: ACM; 2017:261-5.
25. Abashev A, Grigoryev R, Grigorian K, et al. Programming tools for messenger-based chatbot system organization: implication for outpatient and translational medicines. *BioNanoScience* 2016;7:403-7.
26. Min H. Artificial intelligence in supply chain management: theory and applications. *Int J Log Res Appl* 2009;13:13-39.
27. Oakden-Rayner L. Exploring large-scale public medical image datasets. *Acad Radiol* 2020;27:106-12.
28. Johnson AEW, Pollard TJ, Berkowitz SJ, et al. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci Data* 2019;6:317.
29. Borgli H, Thambawita V, Smedsrud PS, et al. Hyper-kvasir: a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *OSF Preprints*. Available at: <https://osf.io/mkzqc>, Accessed June 1, 2020.
30. Guidance for wearable health solutions. Available at: <https://www.hronline.org/documents/CTA-Guidance-for-Wearable-Health-Solutions/download>. Accessed July 13, 2020.
31. Langlotz CP, Allen B, Erickson BJ, et al. A roadmap for foundational research on artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019;291:781-91.
32. Wang P, Xiao X, Glissen Brown JR, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2018;2:741-8.
33. Urban G, Tripathi P, Alkayali T, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 2018;155:1069-78.
34. Russell BC, Torralba A, Murphy KP, et al. LabelMe: a database and web-based tool for image annotation. *Int J Comput Vision* 2008;77:157-73.
35. Dutta A, Zisserman A. The VIA annotation software for images. 2019. arXiv:1904.10699. Available at: <https://arxiv.org/abs/1904.10699>. Accessed June 1, 2020.
36. Bernal J, Histace A, Masana M, et al. GTCreator: a flexible annotation tool for image-based datasets. *Int J Comput Assist Radiol Surg* 2019;14:191-201.
37. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44-56.
38. Chen D, Wu L, Li Y, et al. Comparing blind spots of unsedated ultrafine, sedated, and unsedated conventional gastroscopy with and without artificial intelligence: a prospective, single-blind, 3-parallel-group, randomized, single-center trial. *Gastrointest Endosc* 2020;91:332-9.
39. Gong D, Wu L, Zhang J, et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020;5:352-61.
40. Wang P, Liu X, Berzin TM, et al. 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;5:343-51.
41. Vinsard DG, Mori Y, Misawa M, et al. Quality assurance of computer-aided detection and diagnosis in colonoscopy. *Gastrointest Endosc* 2019;90:55-63.
42. Rex DK, Kahi C, O'Brien M, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc* 2011;73:419-22.
43. Liu X, Faes L, Calvert MJ, et al. Extension of the CONSORT and SPIRIT statements. *Lancet* 2019;394:1225.
44. Faes L, Liu X, Wagner SK, et al. A clinician's guide to artificial intelligence: how to critically appraise machine learning studies. *Translat Vis Sci Technol* 2020;9:1-9.