Could artificial intelligence replace fine-needle aspiration in endoscopic ultrasound?

Shirley X. Jiang[1]^, Nasim Parsa[2]^, Michael F. Byrne[3,4]

[1]Department of Medicine, University of British Columbia, Vancouver, Canada; [2]Division of Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, AZ, USA; [3]Division of Gastroenterology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada; [4]Vancouver General Hospital, Vancouver, BC, Canada

Correspondence to: Michael F. Byrne, MA, MD, MRCP, FRCPC. Clinical Professor of Medicine, Vancouver General Hospital, University of British Columbia, 5153, 2775 Laurel Street, Vancouver, BC, Canada; Division of Gastroenterology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada. Email: Michael.byrne@vch.ca.

Comment on: Yang X, Wang H, Dong Q, *et al.* An artificial intelligence system for distinguishing between gastrointestinal stromal tumors and leiomyomas using endoscopic ultrasonography. Endoscopy 2022;54:251-61.

Received: 29 April 2022; Accepted: 30 May 2022; Published online: 15 May 2023.
doi: 10.21037/gist-22-11
View this article at: https://dx.doi.org/10.21037/gist-22-11

The recent application of artificial intelligence (AI) in the field of gastroenterology has shown promising results in the diagnosis and management of digestive diseases (1-3). Solutions such as AI-powered detection and diagnosis systems are now commercially available for colorectal polyps (4). The backbone of AI systems for image classification is the convolutional neural network (CNN), a deep-learning algorithm that conducts multi-level image analysis through pattern recognition, and improves its own diagnostic ability by training with large datasets (5,6). With the ability to integrate pixel-level data, CNN is able to aid endoscopists in the rapid interpretation of seemingly ambiguous visual data. One such area of diagnostic dilemma is the evaluation of gastric subepithelial lesions (SELs). While endoscopic ultrasound (EUS) is the most accurate imaging modality, there are no definitive EUS features to differentiate gastrointestinal stromal tumors (GISTs) from the commonly encountered gastrointestinal leiomyomas (GILs) (7-9). Misdiagnosis of GISTs and GILs are thought to comprise the majority of incorrect EUS diagnoses (9). Given the malignant potential of GISTs, it is crucial to accurately diagnose these lesions and to differentiate them from GILs, which are benign. The current standard is to differentiate these two by obtaining tissue samples with fine-needle aspiration or biopsy (EUS-FNA/B). However, FNA/B is invasive and is reported to have a lower diagnostic rate for SELs smaller than 20 mm (9,10).

In this issue of *Endoscopy*, Yang *et al.* reported the result of their AI-powered EUS model for differentiation between GISTs and GILs (11). Using a CNN for image recognition, the AI model was trained, validated, and evaluated on a total of 10,439 EUS images from 752 patients with histologically confirmed GISTs and GILs from four endoscopic centers, collected in aggregate from 2013 to 2020. They reported a significantly higher diagnostic accuracy with the AI model compared with the expert endo-sonographer (94.0% *vs.* 70.2%, P value <0.001). More importantly, in the prospective evaluation of 508 consecutive patients with SELs, of whom 132 underwent histologic confirmation, the diagnostic accuracy remained significantly higher with the AI-powered EUS compared with the expert endo-sonographer (78.8% *vs.* 69.7%, P value =0.01). When examining only cases of histologically-confirmed GISTS or GILs, AI-joint diagnosis also had significantly higher accuracy, specificity, and positive predictive value (PPV) at 92.2%, 95.1%, and 94.1%, respectively, compared to individual diagnosis alone at 76.6% (P value =0.01), 65.9% (P value =0.002), and 69.6% (P value <0.01), respectively. The sensitivity and negative predictive value (NPV) of AI-joint diagnosis were similar to individual diagnosis. These

---

^ ORCID: Shirley X. Jiang, 0000-0003-0164-0619; Nasim Parsa, 0000-0003-3882-266X.

are promising results in applying deep learning to real-time EUS to distinguish between GISTs and GILs.

Previous studies have reported an improved diagnostic accuracy for differentiating between GISTs and GILs from non-GIST SELs with the use of AI-powered EUS (12-16). Hirai *et al.* developed an AI system to distinguish GISTs from other gastric SELs using a total of 16,110 images from 631 lesions in a multicentre retrospective dataset for testing and training; the AI system had higher accuracy for diagnosing GIST at 86.1% compared to experts at 27.0–68.0% (P value <0.001) (13). Seven *et al.* applied an AI system to differentiate GISTs from GILs, again trained and tested on 978 retrospectively-collected images from 145 lesions. Similarly, the diagnostic accuracy of the AI system (86.98%) was superior to experts (63.0%, P value <0.05). These earlier AI systems tended to be trained and tested on retrospective datasets, limiting their external validity. Yang *et al.* are the first to report a higher diagnostic accuracy using prospective diagnostic testing, in addition to retrospective testing, from real patients at multiple centers. Reflecting the difficulty in discerning smaller lesions, a previous AI model by Minoda *et al.* had a higher diagnostic yield in SELs larger than 20 mm of 90.0%, compared to 86.3% in SELs less than 20 mm (12). Comparatively, Yang *et al.* found similarly high levels of diagnostic accuracy of SELs larger and smaller than 20 mm (96.3% and 90%, respectively) when experts were assisted by AI (11). Yang *et al.*'s study advances the ability of AI to not only improve diagnostic accuracy compared to experts, but particularly with smaller SELs that are the most challenging to evaluate.

The understated strength of this study is the application of a high-performing AI system to the real-world setting. For the prospective evaluation, the AI system was installed on EUS workstations to assist with diagnoses during endoscopic examinations. Once the endoscopist determined the gastric lesion to be suspicious of GIST or GIL, the AI system was applied to at least five images for image classification, which required about one minute to capture and frame. The AI system then produced a classification in real time to assist the endoscopist with diagnosis. This evaluation provides generalisability that is frequently lacking in early AI development studies, which typically utilize carefully curated testing images that have been pre-processed. There is also the practicality of using the AI system to assist endoscopists to make joint diagnoses as the human operator retains a higher level of abstract thinking. Finally, while Yang *et al.* designed the AI system using a powerful CNN, it was able to perform on entry-level computers and be accessible to endoscopists through a graphical user interface (17). As users did not require any equipment upgrades or additional operator training, the AI system was low in cost relative to its potential benefit. Overall, this AI system was designed with high applicability to support clinicians and could be particularly useful in areas where resources and experts may be scarce.

There are several limitations of this study to note. As lesions were only biopsied when clinically indicated, such as for suspicion of GIST, the histologic diagnosis was only obtained in approximately 30% of the patients in the prospective cohort. This leads to significant verification bias and overestimates sensitivity while underestimating specificity. Yang *et al.* also noted significant discrepancies in AI performance with different EUS probes, corresponding with variation noted in mean pixel values at each study site, limiting the applicability of their AI system to certain imaging devices. Lastly, the model is only able to distinguish between GIST and GIL; authors acknowledge that they were unable to enroll other types of SELs, as their low incidence did not allow for sufficient training images in the dataset. This limitation was similarly seen in a previous AI system that found high accuracy overall to diagnose GISTs, but misdiagnosed a schwannoma as few training images were available (12). As the performance of an AI model is predicated upon gathering sufficient heterogeneity in training data, this is a common challenge in AI development.

The results of this study add to growing literature on the superior accuracy of AI systems to diagnose GISTs on EUS. Yang *et al.* have demonstrated the potential role of AI systems in real-time and real-world EUS diagnosis of GISTs. This exciting study illustrates the ability of AI systems to identify features that may not be visible or discernable to endoscopists, empowering clinicians with more information than ever before. To increase the external validity of AI systems for EUS, further research is needed with large, prospective, and multi-center data of histologically proven lesions and different EUS models. Future studies should strive to apply the same pragmatism as Yang *et al.* by emphasizing the importance of AI systems being user-friendly, compatible with existing infrastructure, and performing in real time without incurring excessive costs. Current challenges in gastrointestinal endoscopy represent opportunities for AI solutions to empower clinical decision making, optimize healthcare resources, and improve patient care in the not-so-distant future.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Gastrointestinal Stromal Tumor*. The article did not undergo external peer review.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://gist.amegroups.com/article/view/10.21037/gist-22-11/coif). MFB is the CEO, founder, and shareholder of Satisfai Health, and founder of AI4GI joint venture. There is a co-development agreement between Olympus America and AI4GI in artificial intelligence and colorectal polyps. The other authors have no conflicts of interest.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Kröner PT, Engels MM, Glicksberg BS, et al. Artificial intelligence in gastroenterology: A state-of-the-art review. World J Gastroenterol 2021;27:6794-824.
2. Le Berre C, Sandborn WJ, Aridhi S, et al. Application of Artificial Intelligence to Gastroenterology and Hepatology. Gastroenterology 2020;158:76-94.e2.
3. Calderaro J, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. Gut 2021;70:1183-93.
4. Taghiakbari M, Mori Y, von Renteln D. Artificial intelligence-assisted colonoscopy: A review of current state of practice and research. World J Gastroenterol 2021;27:8103-22.
5. Ruffle JK, Farmer AD, Aziz Q. Artificial Intelligence-Assisted Gastroenterology- Promises and Pitfalls. Am J Gastroenterol 2019;114:422-8.
6. Choi J, Shin K, Jung J, et al. Convolutional Neural Network Technology in Endoscopic Imaging: Artificial Intelligence for Endoscopy. Clin Endosc 2020;53:117-26.
7. Dumonceau JM, Deprez PH, Jenssen C, et al. Indications, results, and clinical impact of endoscopic ultrasound (EUS)-guided sampling in gastroenterology: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline - Updated January 2017. Endoscopy 2017;49:695-714.
8. Standards of Practice Committee; Faulx AL, Kothari S, et al. The role of endoscopy in subepithelial lesions of the GI tract. Gastrointest Endosc 2017;85:1117-32.
9. Deprez PH, Moons LMG, O'Toole D, et al. Endoscopic management of subepithelial lesions including neuroendocrine neoplasms: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy 2022;54:412-29.
10. Minoda Y, Chinen T, Osoegawa T, et al. Superiority of mucosal incision-assisted biopsy over ultrasound-guided fine needle aspiration biopsy in diagnosing small gastric subepithelial lesions: a propensity score matching analysis. BMC Gastroenterol 2020;20:19.
11. Yang X, Wang H, Dong Q, et al. An artificial intelligence system for distinguishing between gastrointestinal stromal tumors and leiomyomas using endoscopic ultrasonography. Endoscopy 2022;54:251-61.
12. Minoda Y, Ihara E, Komori K, et al. Efficacy of endoscopic ultrasound with artificial intelligence for the diagnosis of gastrointestinal stromal tumors. J Gastroenterol 2020;55:1119-26.
13. Hirai K, Kuwahara T, Furukawa K, et al. Artificial intelligence-based diagnosis of upper gastrointestinal subepithelial lesions on endoscopic ultrasonography images. Gastric Cancer 2022;25:382-91.
14. Seven G, Silahtaroglu G, Seven OO, et al. Differentiating Gastrointestinal Stromal Tumors from Leiomyomas Using a Neural Network Trained on Endoscopic Ultrasonography Images. Dig Dis 2021. [Epub ahead of print]. doi: 10.1159/000520032.
15. Kim YH, Kim GH, Kim KB, et al. Application of A Convolutional Neural Network in The Diagnosis of Gastric Mesenchymal Tumors on Endoscopic

Ultrasonography Images. J Clin Med 2020;9:3162.
16. Oh CK, Kim T, Cho YK, et al. Convolutional neural network-based object detection model to identify gastrointestinal stromal tumors in endoscopic ultrasound images. J Gastroenterol Hepatol 2021;36:3387-94.
17. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 27-30 June 2016; Las Vegas, NV, USA. IEEE, 2016:770-8.

doi: 10.21037/gist-22-11
Cite this article as: Jiang SX, Parsa N, Byrne MF. Could artificial intelligence replace fine-needle aspiration in endoscopic ultrasound? Gastrointest Stromal Tumor 2023;6:6.