

A human-machine collaboration: bringing artificial intelligence into colonoscopy

Omer F. Ahmad¹, Danail Stoyanov¹, Laurence B. Lovat^{1,2}

Author Affiliations:

- 1) Wellcome/EPSRC Centre for Interventional & Surgical Sciences (WEISS), University College London, London, UK.
- 2) Division of Surgery & Interventional Science, University College London, London, UK.

Correspondence to:

Dr Omer F. Ahmad; o.ahmad@ucl.ac.uk

Colorectal cancer (CRC) is the third commonest malignancy worldwide. Colonoscopy offers protection against the development of CRC by detection and resection of neoplastic lesions. Unfortunately, the procedure remains highly operator dependent. A meta-analysis of tandem colonoscopy studies revealed a pooled miss rate of 22% for polyps of any size.¹ Post-colonoscopy CRCs are associated with low adenoma detection rates (ADRs) and incompletely resected or missed lesions are recognised as key contributory factors.² International efforts to improve quality must be commended, particularly those led by the Joint Advisory Group on Gastrointestinal Endoscopy in the United Kingdom, where individual colonoscopy performance is assessed by quality assurance using key performance indicators. Despite these efforts further significant improvement is needed.

Numerous strategies have been utilised in attempts to improve ADRs including educational interventions, enhanced imaging techniques and mechanical devices to improve mucosal exposure. Computer aided detection and diagnosis (CAD) systems using advanced artificial intelligence (AI) techniques represent an emerging technology that will likely lead to a paradigm shift in the field.³

Machine learning is a type of AI applied to systems, that allows for automatic learning and improved performance on datasets without the need for explicit programming of prediction rules. Advanced deep-learning based approaches have revolutionized the area by utilising artificial neural networks.⁴ These systems are biologically inspired by the concept of neurons and synapses in the human brain to discover image features that optimally represent the data for a specific task. Examples of successful applications include demonstrations that deep-learning algorithms are able to match the performance of expert dermatologists in differentiating benign from malignant skin lesions and are better than pathologists at assessing slide images for the presence of breast cancer metastases within axillary node specimens during a timed exercise designed to simulate clinical practice.^{5,6}

CAD systems are being developed rapidly for real-time polyp detection. A recent study by Urban *et al.* describes the use of a deep convolutional neural network (CNN) trained on 8641 images from 2000 patients, leading to a polyp detection accuracy of 96.4% in an independent set of 1330 images containing 672 polyps operating at real-time video rate.⁷ More importantly, when further evaluated on 9 retrospectively collected colonoscopy videos, an

additional 17 polyps that were not removed at the index procedure were identified using CNN assistance with a relatively low false positive rate.

Furthermore, some CAD systems can now match human experts in performing ‘optical biopsies’.⁸ In 2017, the National Institute for Clinical and Healthcare Excellence (NICE) recommended that virtual chromoendoscopy can be used instead of histopathology to determine if diminutive ($\leq 5\text{mm}$) polyps are adenomatous or hyperplastic under certain conditions which includes high confidence assessments.⁹ Studies have demonstrated that the Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) standards set for such a ‘resect and discard’ strategy can be achieved in academic settings but not always in community-based practice, limiting its incorporation into routine care.¹⁰ CAD offers a promising timely solution to overcome this barrier by providing decision support. Byrne *et al.* used a CNN to differentiate diminutive adenomas from hyperplastic polyps on unaltered NBI videos collected from a previous study using standard colonoscopy.¹¹ The algorithm provided an associated probability score for predictions. Using histopathology as gold standard, the model provided real-time high confidence decisions for 106 diminutive polyp videos, the overall accuracy achieved was 94% (sensitivity 98%, specificity 83%, positive predictive value 90%, negative predictive value 97%). More recently, Mori *et al.* conducted a prospective study using a CAD system for endocytoscopy, which produces ultra-high magnification images that allow for in vivo cellular or microvascular assessment with staining or NBI respectively.¹² The system achieved the PIVI threshold for a ‘diagnose and leave’ strategy for diminutive, non-neoplastic recto-sigmoid polyps. The authors commented on a future clinical trial to investigate the CAD system for a ‘resect and discard’ strategy.

Clinical trials utilising AI powered software during endoscopy are imminent and it is vital that clinical endoscopists are engaged at this time of rapid technological development. It is possible that new quality metrics will emerge based on computational advances. An advantage of CAD is the potential for relatively simple integration with standard endoscopy equipment by using ‘plug-in’ software. However, before implementation into routine NHS practice or bowel cancer screening programmes can even be considered there must be further evaluation of clinician acceptance and assessment of potential impact on workflow. There are also regulatory hurdles that must be overcome including addressing concerns about the lack of transparency and complexity involved in deep-learning based decision making which is often considered a ‘black box’. Despite these challenges, we should take this opportunity to reflect on the quality of our own current endoscopy practice and embrace a future ‘human-machine collaboration’ that will pave the way for the very highest quality endoscopy for our patients.

Contributors:

All authors contributed to the intellectual content, drafting and final approval of the work.

Competing interests: None declared.

Acknowledgements

This work was undertaken at UCL/UCLH who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health. This work was also supported by the CRUK Experimental Cancer Medicine Centre at UCL and the Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS) at UCL; [203145Z/16/Z].

REFERENCES

- 1 Van Rijn JC, Reitsma JB, Stoker J, Bossuyt PM, Van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: A systematic review. *Am J Gastroenterol* 2006; **101**: 343–50.
- 2 Kaminski MF, Regula J, Kraszewska E, *et al.* Quality Indicators for Colonoscopy and the Risk of Interval Cancer. *N Engl J Med* 2010; **362**: 1795–803.
- 3 Byrne MF, Shahidi N, Rex DK. Will Computer-Aided Detection and Diagnosis Revolutionize Colonoscopy? *Gastroenterology* 2017; **153**: 1460–1464.e1.
- 4 Chartrand G, Cheng PM, Vorontsov E, *et al.* Deep Learning: A Primer for Radiologists. *RadioGraphics* 2017; **37**: 2113–31.
- 5 Esteva A, Kuprel B, Novoa RA, *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; **542**: 115–8.
- 6 Ehteshami Bejnordi B, Veta M, Johannes van Diest P *et al.* Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 2017; **318**: 2199–210.
- 7 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; published online July 12. DOI:10.1053/j.gastro.2018.06.037.
- 8 Chen P-J, Lin M-C, Lai M-J, Lin J-C, Lu HH-S, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568–75.
- 9 NICE. Virtual chromoendoscopy to assess colorectal polyps during colonoscopy. National Institute for Health Care Excellence. May, 2017. <https://www.nice.org.uk/guidance/dg28>.
- 10 Rees CJ, Rajasekhar PT, Wilson A, *et al.* Narrow band imaging optical diagnosis of small colorectal polyps in routine clinical practice: the Detect Inspect Characterise Resect and Discard 2 (DISCARD 2) study. *Gut* 2017; **66**: 887–95.
- 11 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* 2017. DOI:10.1136/gutjnl-2017-314547.
- 12 Mori Y, Kudo S, Misawa M, *al et.* Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: A prospective study. *Ann Intern Med* 2018; published online Aug 14. <http://dx.doi.org/10.7326/M18-0249>.