Artificial intelligence (AI) has recently become an essential technology applied to various fields, including weather forecast, social networking, automatic driving, and video games. This evolution has also rapidly affected medicine and healthcare, starting with the dermatological classification of skin cancer with deep neural networks. In the field of gastrointestinal endoscopy, computer-aided diagnosis (CAD) is gaining increased attention as a reliable technological revolution that can not only compensate for human errors but also perform tasks that are beyond human ability. The major roles and expected advantages of CAD include computer-aided detection of lesions that are barely detectable and computer-aided histopathological diagnosis without biopsy or removal of tissue. Colorectal cancer is a major cause of cancer-related death worldwide, increasing the adenoma detection rate, one of the vital quality indicators in colonoscopy, is a topic of high priority, while optical biopsy is another tool that may help to reduce increasing medical expenses. Optical biopsy has gained more acceptance because it may be more cost-effective and time-efficient than traditional procedures, e.g. histopathology of removed polyps, especially small non-malignant polyps, do not need to be reviewed by a pathologist.

In an issue of Clinical Endoscopy, Racz et al. reported that the artificial intelligence-based polyp histology prediction (AIPHP) software could predict colorectal polyp histology with high accuracy in the larger polyp subgroup. In this study, a total of 373 colorectal polyps were removed by polypectomy or mucosectomy after detection, first by high-definition colonoscopy followed by observation with narrow-band imaging (NBI) at the maximum (×65) magnification. Then, the stored NBI images were analyzed according to the NBI International Colorectal (NICE) classification by three endoscopists with more than 20 years of experience, and AIPHP in a parallel fashion. They used histology as the gold standard based on the WHO classification, and two-class classifications were considered as hyperplastic or neoplastic, including sessile serrated lesions (SSL), adenomas, and invasive adenocarcinomas. The AIPHP software was based on the vascular pattern and color of the polyps, but human interaction is needed to select the area of interest because the present version cannot automatically find it. The main results are as follows: the accuracy of AIPHP was 86.6% overall, 82.1% for diminutive polyps, and 92.2% for non-diminutive polyps. The accuracy of the hyperplastic histology prediction by NICE was 95.2% in the diminutive polyps and 97.1% in all evaluated polyps, while the accuracy by the AIPHP method was 82.1% and 86.6%, thus significantly better with NICE in both categories. Linear regression correlation coefficient results indicate that the AIPHP method was significantly more accurate in the bigger polyps than in the smaller polyps, when a detailed calculation was...
performed.

The results demonstrated that presently the diagnostic performance of AIPHP did not surpass that of NICE, which is an already tested classification method by human endoscopists. Therefore, the conclusion they have drawn was relatively modest. However, this study has elucidated some latent and profound issues in developing AI technology in conjunction with knowledge and experience integrated by humans. First, the development of an AI diagnostic algorithm requires a human diagnosis and labeling to train the AI. The system used in this study involved human interaction in selecting good-quality areas, which may build an inherent bias into the system from the start. Second, AI will not be able to compensate for the risk of suboptimal colonoscopy in cases of inadequate cleaning that may affect the classification of the surface pattern of each polyp. Although smaller numbers of SSL were found in this study, this flat and isochromic polyp is frequently covered with a mucus cap or debris during routine colonoscopy, which may hinder the original character. Third, this system is a proprietary specification system developed in a single center, and there may be an uneven spectrum of diagnostic features, including pathology. Hence, further external validation of the accuracy of a large number of facilities are needed. Fourth, along with the evolution of endoscopic AI, high precision and negative predictive values should be warranted. We have to wait until multicenter, prospective randomized studies with a large number of cases have been conducted. Currently, we are only halfway through the distribution of AI; however, it will gain wider acceptance not only as a clinical tool in daily practice but also as an educational gadget in the near future.

Conflicts of Interest

The authors have no potential conflicts of interest.

Funding

None.

Author Contributions

Writing-original draft: Naoki Muguruma
Writing-review & editing: Tetsuji Takayama

ORCID

Naoki Muguruma: https://orcid.org/0000-0002-4740-4809
Tetsuji Takayama: https://orcid.org/0000-0002-0175-1573

REFERENCES