UNITED EUROPEAN GASTROENTEROLOGY ueg journal

Original Article

United European Gastroenterology Journal
2019, Vol. 7(4) 538–547


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2050640619837443
journals.sagepub.com/home/ueg
SAGE

The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy

Jeroen de Groof[1], Fons van der Sommen[2], Joost van der Putten[2], Maarten R Struyvenberg[1], Sveta Zinger[2], Wouter L Curvers[3], Oliver Pech[4], Alexander Meining[5], Horst Neuhaus[6], Raf Bisschops[7], Erik J Schoon[3], Peter H de With[2] and Jacques J Bergman[1]

Abstract

Background: Computer-aided detection (CAD) systems might assist endoscopists in the recognition of Barrett's neoplasia.
Aim: To develop a CAD system using endoscopic images of Barrett's neoplasia.
Methods: White light endoscopy (WLE) overview images of 40 neoplastic Barrett's lesions and 20 non-dysplastic Barret's oesophagus (NDBO) patients were prospectively collected. Experts delineated all neoplastic images.
The overlap area of at least four delineations was labelled as the 'sweet spot'. The area with at least one delineation was labelled as the 'soft spot'. The CAD system was trained on colour and texture features. Positive features were taken from the sweet spot and negative features from NDBO images. Performance was evaluated using leave-one-out cross-validation. Outcome parameters were diagnostic accuracy of the CAD system per image, and localization of the expert soft spot by CAD delineation (localization score) and its indication of preferred biopsy location (red-flag indication score).
Results: Accuracy, sensitivity and specificity for detection were 92, 95 and 85%, respectively. The system localized and red-flagged the soft spot in 100 and 90%, respectively.
Conclusion: This uniquely trained and validated CAD system detected and localized early Barrett's neoplasia on WLE images with high accuracy. This is an important step towards real-time automated detection of Barrett's neoplasia.

Keywords

Barrett's oesophagus, endoscopy, computer-aided detection, Barrett's neoplasia, artificial intelligence

Received: 19 December 2018; accepted: 2 February 2019

Key summary

What is known on this subject?

- Endoscopic detection of Barrett's neoplasia is difficult.
- Computer-aided detection (CAD) systems could potentially assist endoscopists in detection of neoplasias.

What are the significant and/or new findings of this study?

- Our CAD-system detected and localized Barrett's neoplasia on endoscopic images with high accuracy.

[1]Department of Gastroenterology and Hepatology, University of Amsterdam, Amsterdam, The Netherlands
[2]Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
[3]Department of Gastroenterology and Hepatology, Catharina Hospital Eindhoven, Eindhoven, The Netherlands
[4]Gastroenterology and Interventional Endoscopy, Krankenhaus Barmherzige Brüder, Regensburg, Germany
[5]Center of Internal Medicine, Ulm University, Ulm, Germany
[6]Internal Medicine, Evangelisches Krankenhaus Düsseldorf, Düsseldorf, Germany
[7]Department of Gastroenterology and Hepatology, University Hospitals Leuven, Leuven, Belgium

Corresponding author:
Jacques J Bergman, Academic Medical Centre, Meibergdreef 9 1105 AZ, Amsterdam, The Netherlands.
Email: a.j.degroof@amc.uva.nl

Introduction

Barrett's oesophagus (BO) is a known precursor for esophageal adenocarcinoma (EAC). BO patients undergo regular endoscopic surveillance to detect EAC at an early stage to enable endoscopic treatment, which is associated with excellent outcomes.[1–3] However, endoscopic detection of early neoplasia is difficult and early lesions are therefore often missed.[4] Primarily, this is due to its subtle appearance; early Barrett's neoplasia is most often flat with only minimal changes in mucosal colour and texture. These subtle changes are generally visible on high-definition white light endoscopy (WLE) in expert hands; however, due to the low progression rate of Barrett's neoplasia ($<1\%$ per patient year), most general endoscopists rarely encounter early Barrett's neoplasia, and are thus unfamiliar with its endoscopic appearance and therefore do not recognize these lesions.[5,6]

Over the last decade, multiple computer-aided detection (CAD) systems have been developed for multiple applications in medical imaging.[7–13] The ability of modern-day computers to automatically recognize informative patterns in data sets can potentially improve endoscopic detection of early neoplastic BO. Ideally, such a CAD system would be incorporated in the endoscopy system to run real-time on the background during surveillance endoscopies. The development of such a system is structured in several steps. First the algorithm is trained on individual endoscopic still images, followed by incorporating video recordings and finally an algorithm for real-time analyses. Herein, we describe the first step of this structured approach by the ARGOS consortium. The ARGOS consortium consists of three international tertiary referral centres for Barrett's neoplasia, a leading academic image analysis group, two commercial enterprises, and is supported by the Dutch Cancer Society and Technology Foundation STW, as part of their joint strategic research programme 'Technology for Oncology'.

The aim of this study was to validate an improved version of our CAD system on high-quality endoscopic images.

Methods

Study setting

This study was performed at the departments of Gastroenterology and Hepatology of the Amsterdam University Medical Centers (location Academic Medical Center), the Catharina Hospital Eindhoven and University Hospital Leuven, and at the department of Electrical Engineering of Eindhoven University of Technology. The Medical Research Involving Human Subjects Act did not apply to this study. Official approval of this study was therefore waived by the Medical Ethics Review Committees of the participating centres.

Image acquisition

In this study, both endoscopic images of early Barrett's lesions and endoscopic images of normal appearing, non-dysplastic Barrett's oesophagus (NDBO) were prospectively collected. All images were recorded via WLE in full high-definition format (1280×1024 pixels) with the ELUXEO^{TM} 7000 endoscopy system (FUJIFILM, Tokyo, Japan).

All procedures were performed by expert endoscopists (JB, ES, RB and WC) with extensive experience in the use of advanced imaging techniques and endoscopic treatment of Barrett's neoplasia. The endoscopic images with neoplastic Barrett's lesions were prospectively collected in a previous study.[14] The NDBO endoscopic images were collected prospectively for this study from NDBO patients undergoing standard surveillance endoscopy, performed by the same experts and using the same image acquisition protocol as the study mentioned above. In the absence of visible lesions, the endoscopist selected an area containing normal appearing Barrett's mucosa, from which a dedicated WLE image was obtained in overview. During imaging of these areas, the endoscopist tried to mimic the endoscopic positioning as if there was a visible lesion present, to minimize potential bias. Subsequently targeted biopsies of this area were obtained followed by random biopsies according to the Seattle protocol.

All endoscopic resection specimens and biopsies were reviewed by pathologists expert in early Barrett's neoplasia at the participating centres.

Image processing and development of ground truth for algorithm

All images of the neoplastic lesions were independently assessed by six international BO experts (JB, RB, OP, ES, AM and HN) who delineated the lesions using a proprietary online software module. The software of this module allowed endoscopic images to be delineated on a computer screen and subsequently enabled the calculation of surface overlap of delineations. Figure 1 shows exemplary overview images with expert delineations.

The expert delineations were used to establish a ground truth that could be used as input for the algorithm. An overlap area of at least four expert delineations, i.e. the area that $>50\%$ of experts assessed as neoplastic, was considered to have the highest suspicion of visible neoplasia and was labeled as the 'sweet spot'. This area was used to train the algorithm to recognize

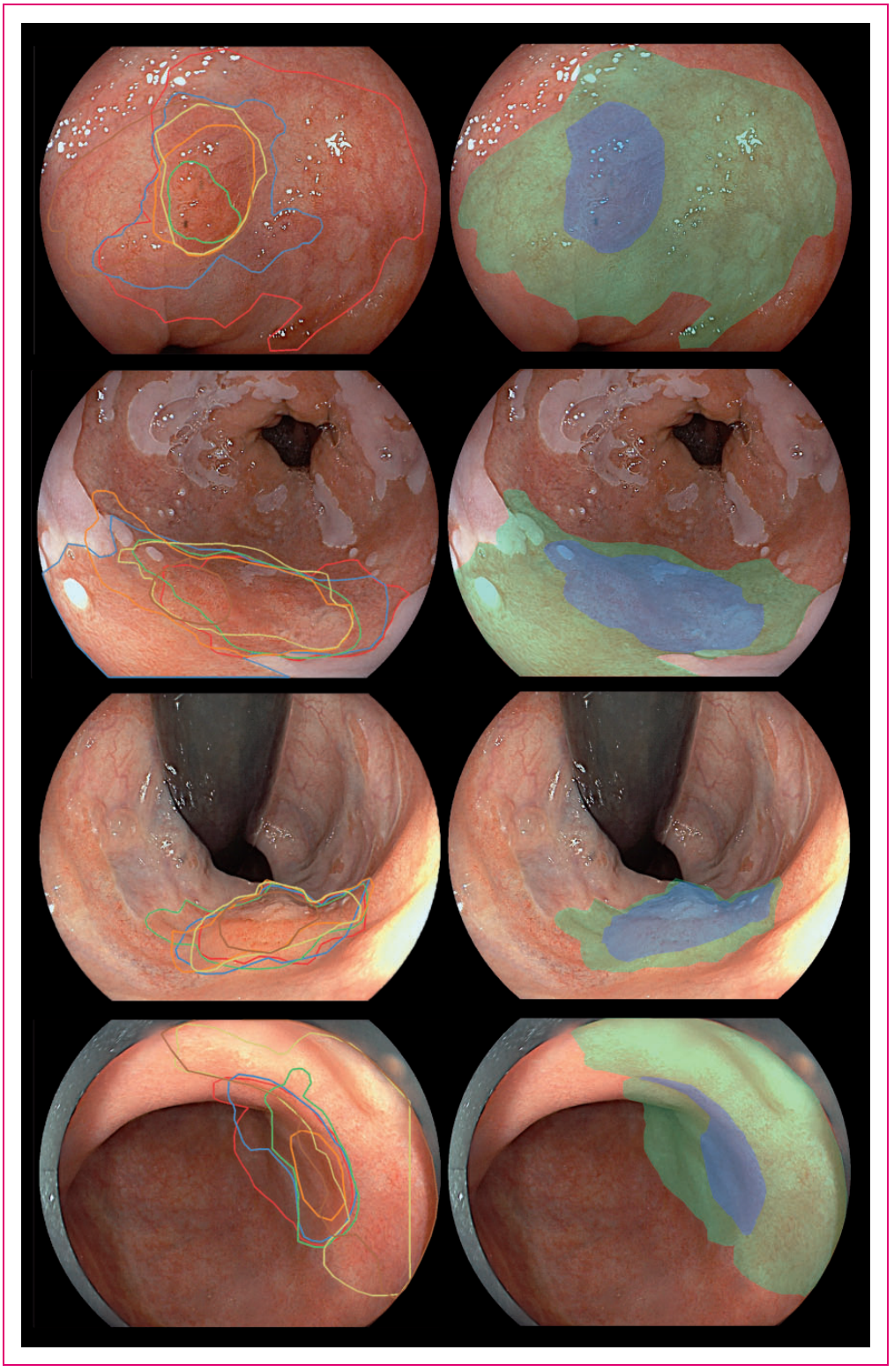

Figure 1. Illustration of six international expert delineations (left), and the creation of the sweet spot in blue and the soft spot in green (right).

neoplasia. The area with at least one expert delineation was labelled as the 'soft spot'. This larger area was considered to potentially contain neoplasia, since at least one expert assessed this area as neoplastic. The area outside the soft spot was considered to be non-neoplastic, since none of the experts assessed this area as neoplastic. Figure 1 shows a graphical display of these areas.

NDBO images were included when both targeted and random biopsies showed no dysplasia, and review of all images by two experts (JB and WC) confirmed the absence of any visible lesions.

Computer algorithm design

The primary goal of this algorithm is to serve as a red-flag detection technique by the creation of delineations of neoplastic areas, thereby assisting endoscopists to detect areas of interest during surveillance endoscopies. Although the algorithm thus provides a delineated area, the focus is on allowing detection, not on exact delineation, of the lesion since this is generally done using a combination of optical chromoscopy and magnified view.

This CAD system uses supervised learning techniques and is designed to follow a stepwise workflow, described briefly below. An extensive, technical explanation of the baseline system has been described previously.[15]

First, the CAD system detects regions of interest. The system is designed to automatically detect the lumen, intestinal juices and specular reflections in the endoscopic images, and excludes these areas from analyses.

Subsequently, the regions of interest are divided in blocks, each block encompassing an area of 60×60 pixels. Each block is labelled as being 'neoplastic' or 'non-neoplastic' based on the combined expert delineations. The neoplastic blocks are obtained from the above-mentioned sweet spot. The non-neoplastic blocks are obtained from the NDBO images as well as the area outside the soft spot of the neoplastic images.

Third, informative image features are extracted from each block, on which the algorithm discriminates between 'normal' versus 'abnormal' tissue. As we have learned from previous studies, early neoplasia is associated with changes in colour and texture.[15] To quantify both features, statistical information about the colour values is computed and special filters are applied to capture the relevant texture patterns.[10]

The features are then used as input for a support vector machine classifier, which is first trained to discriminate between neoplastic and non-neoplastic features, and subsequently employed for classification of the blocks into either category. The CAD system first decides whether the image is suspicious for neoplasia. This is done by combining all individual block predictions, resulting in an image-based confidence score. When this confidence score meets its threshold, the CAD system labels the image as being 'neoplastic' and produces a delineation on the image, encircling the region suspicious for neoplasia (Figure 2). After delineating this region, ideally capturing the entire neoplastic lesion, the algorithm subsequently indicates the most abnormal part of the lesion by calculating which block within its delineation is most abnormal. This is then displayed on the image as a cross-hair visualization, thereby 'red-flagging' the most suspicious area (Figure 3).

Outcome measurements

Performance of the algorithm was evaluated using a leave-one-out cross-validation.

Primary outcome measurements. Detection scores, per image analysis: detection was considered correct when the algorithm correctly identified an image as neoplastic or non-neoplastic.

Secondary outcome measurements

1. Localization scores: number of images in which the delineation produced by the algorithm overlapped with parts of the sweet or soft spot of experts;
2. Red-flag indication scores: percentage of recognized neoplastic images where the algorithm red-flagged the sweet or soft spot of experts;
3. Time required for analysis of an image by the algorithm.

Statistical analysis

Diagnostic accuracy of the algorithm per image was displayed as area under the curve and in terms of accuracy, sensitivity and specificity. Software package MATLAB 2018a (MathWorks, Inc., Natick, Massachusetts, USA) was used to perform statistical tests.

Results

In this study, 60 patients in total were included for prospective image acquisition. Forty patients presented with a neoplastic lesion and 20 patients presented with NDBO. From each patient, one endoscopic image was included. Histological evaluation of all endoscopic resection specimens showed high-grade dysplasia or EAC for all neoplastic cases. All biopsies from NDBO patients were shown to contain non-dysplastic Barrett's mucosa.

Primary outcome measurements

Detection scores. In per-image analyses, accuracy, sensitivity and specificity of the algorithm were 91.7, 95 and 85%, respectively. Figure 4 shows the corresponding receiver operating characteristic.

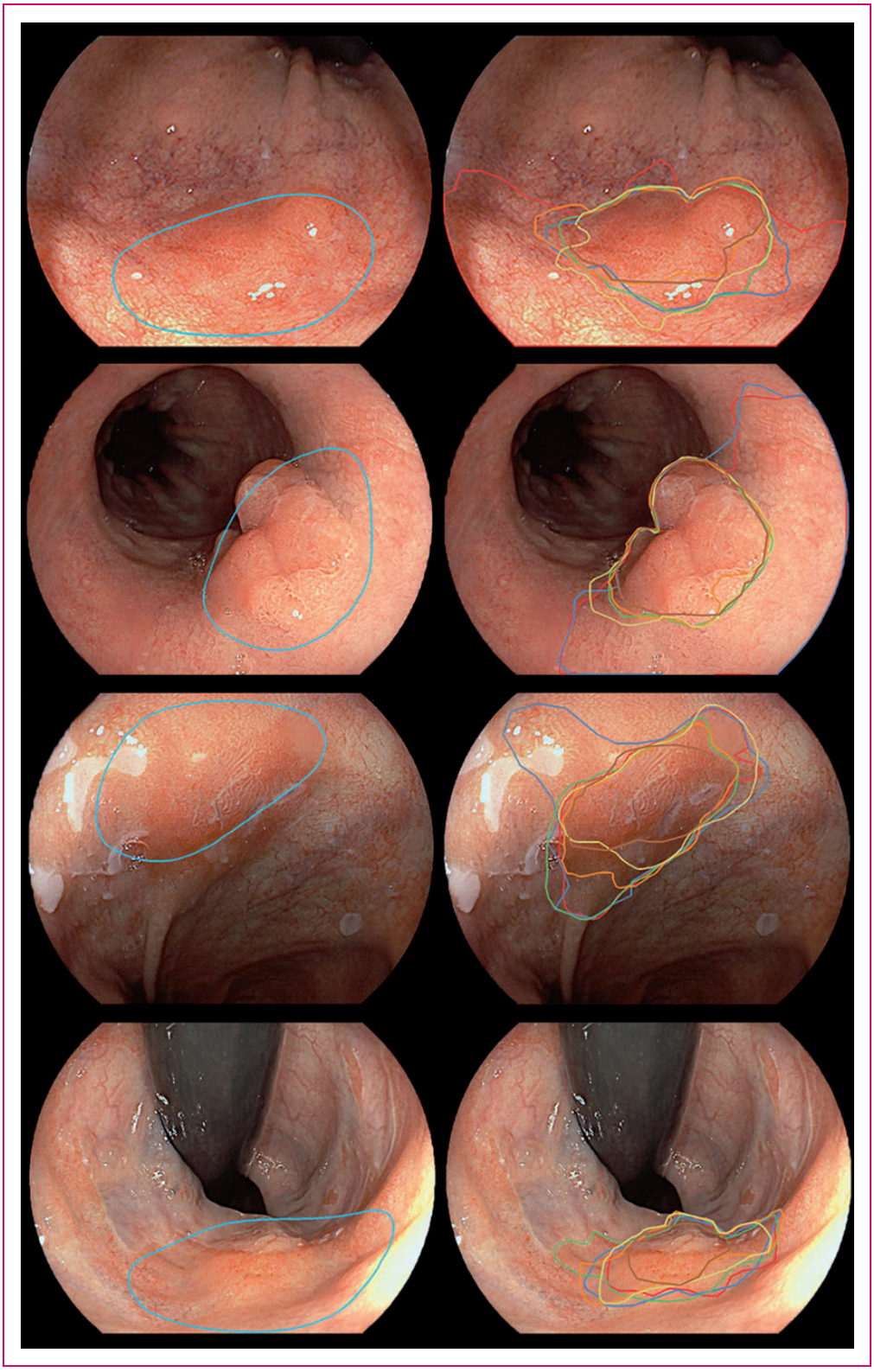

Figure 2. Illustration of the delineation tool of the algorithm (left), compared with the six international experts (right).

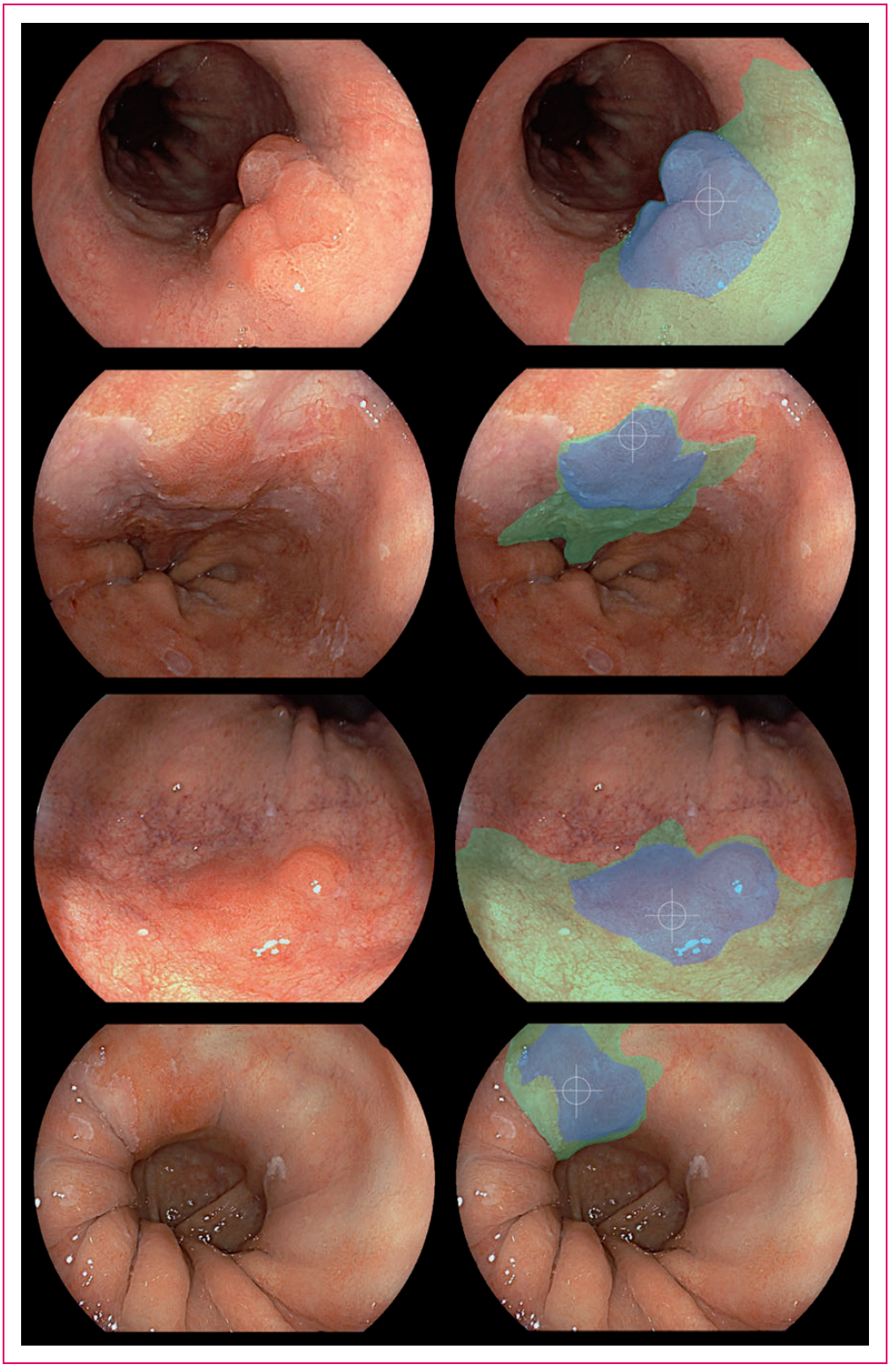

Figure 3. Illustration of the red-flag tool of the algorithm with background visualization of the sweet spot (blue) and soft spot (green).

Secondary outcome measurements

Localization scores. In the 38 images that the CAD system correctly identified as neoplastic, in 100% (38/38) of the images the delineation of the algorithm recognized the soft spot as neoplastic. In 97.4% (37/38) images, the system also recognized parts of the sweet

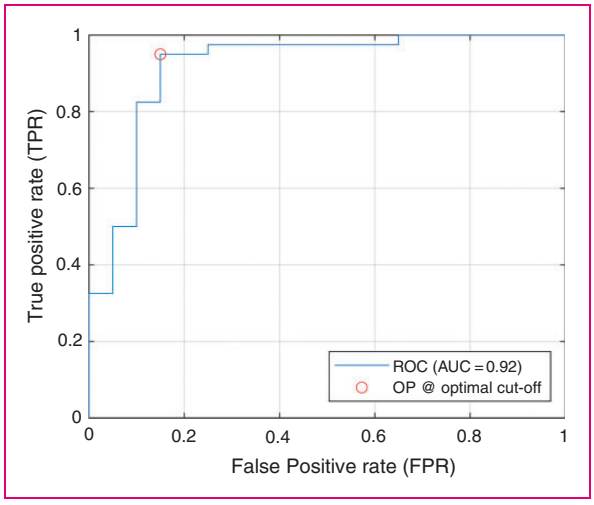

Figure 4. Receiver operating characteristic of detection performance.
AUC: area under the curve; FPR: false-positive rate; OP: [insert definition]; ROC: receiver operating characteristic; TPR: true-positive rate.

Table 1. Secondary outcome measurements.

	Soft spot	Sweet spot
Localization score algorithm (%)	100 (38/38)	97.4 (37/38)
Red-flag indication score algorithm (%)	89.5 (34/38)	76.3 (29/38)

spot as neoplastic. Taking into account the undetected lesions, this leads to localization scores of 95 (38/40) and 92.5% (37/40) for soft and sweet spot recognition, respectively (see Table 1).

Red-flag indication scores. In 89.5% (34/38) of the correctly identified images, the algorithm placed the red flag within the soft spot. In 76.3% (29/38) of images, the algorithm placed the red flag within the sweet spot. Taking into account the undetected lesions, this leads to red-flag indication scores of 85 (34/40) and 72.5% (29/40) for soft and sweet spot recognition, respectively (see Table 1). Figure 3 shows an illustration of the red-flag tool.

Figure 2 shows exemplary cases of detection performance of the algorithm, compared with the six international experts.

Time analyses. The total time it took the algorithm to analyse all images and produce lesion delineations was 61.8 seconds. Mean time per image was 1.051 seconds (SD 0.041). Mean time required for region of interest detection was 194 milliseconds (SD 19), for feature extraction 790 milliseconds (SD 32) and for delineation 31 milliseconds (SD 12).

Discussion

This paper describes the development of a WLE-based CAD system for real-time endoscopic detection of early Barrett's neoplasia. To our knowledge, we are the first group to develop such a CAD system.

Early Barrett's neoplasia can easily be missed during surveillance endoscopies, even with high-definition WLE. CAD systems have the potential to assist endoscopists in the recognition of Barrett's neoplasia. The first step in the development of such a CAD system is to develop an algorithm designed for endoscopic still images. In this study, our CAD system detected early neoplastic Barrett's lesions on a selection of endoscopic WLE images with high accuracy.

In 38/40 images, the algorithm correctly classified an image as containing neoplasia. Since we aim to develop an algorithm that can not only classify an image but also localize the neoplastic lesion, we developed two additional algorithm functions and corresponding outcome parameters. Our CAD algorithm depicts a delineation of the entire lesion to localize the lesion and subsequently indicates the most abnormal area within that delineation to indicate the most appropriate position to obtain a targeted biopsy. During the development of the corresponding outcome parameters (i.e. the localization score and the red-flag indication score), we reasoned that, in order to equal expert performance, these algorithm features should at least recognize parts of the soft spot, since one or more experts assessed this area as being neoplastic.

In all 38 detected images, the delineation produced by the algorithm recognized the soft spot as being neoplastic, while in 37/38 images the algorithm also recognized the sweet spot as neoplastic. Upon reviewing the two images not recognized as neoplastic by the algorithm, we noticed that these lesions were flat and had a very subtle appearance, with relatively lower image quality when compared to the other images, as shown in Figure 5. In the single image where only the soft spot was recognized by the algorithm, the sweet spot was not recognized as a region of interest. As mentioned, the algorithm excludes areas that are not suitable for analysis. When the largest part of a neoplastic lesion contains specular reflections, that area is excluded from further analyses by the algorithm and the lesion is therefore missed. This problem will likely be minimized when we apply our CAD to real-time imaging instead of still images, since this will provide more dynamic footage and thereby minimize any temporary effects of specular reflections.

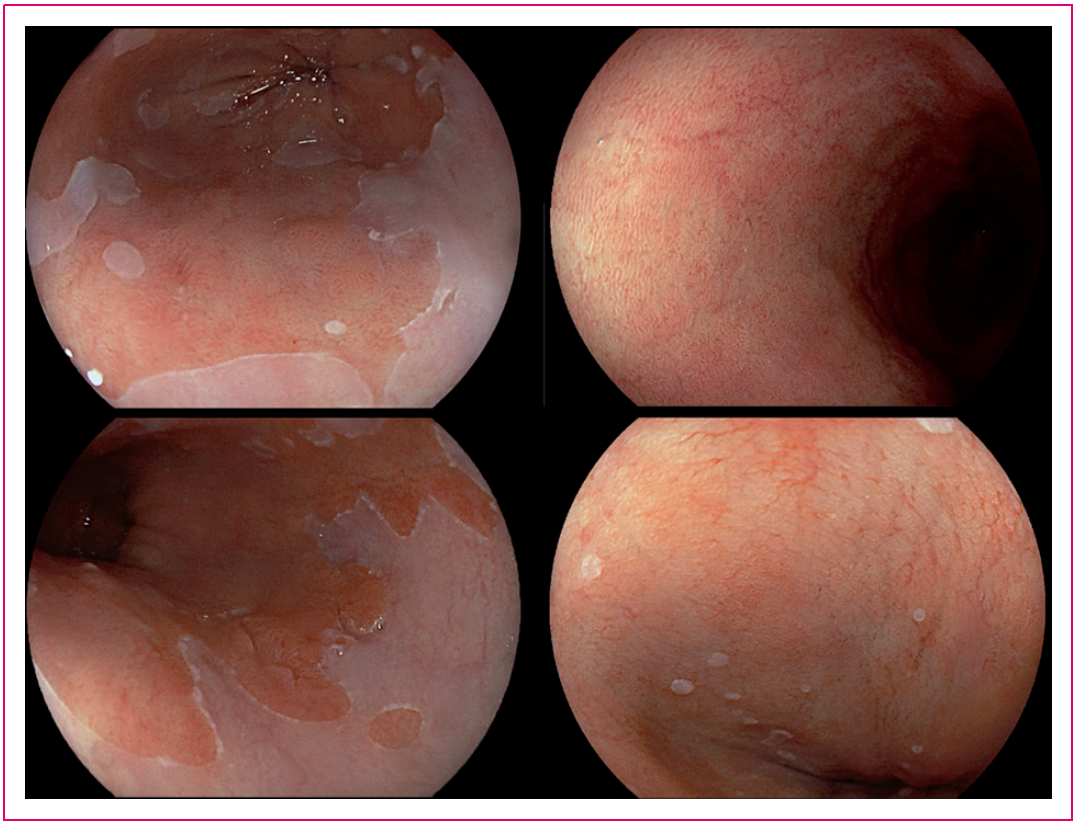

Figure 5. False-negative cases (left) and false-positive cases (right).

The red-flag indication score is a parameter that reflects the accuracy of our CAD algorithm to correctly indicate the most appropriate position for obtaining a targeted biopsy. The red-flag indicator was positioned in the soft spot in 34/38 detected cases and was placed in the sweet spot in 29/38 cases.

In 3/20 cases, the algorithm incorrectly detected and delineated a lesion on NDBO images. Upon reviewing these images, the 'detected' areas appeared to have increased vascularization and might therefore be misclassified as having been suspicious for neoplasia. Figure 5 displays two false-positive cases. The clinical relevance of false-positive detections, particularly when occurring as infrequently as in our study, is however much lower than that of false-negative detections.

On average, it took our algorithm 1 second to analyse an endoscopic image and subsequently produce its lesion delineation. Needless to say, this speed outperforms endoscopists. Given the current execution speed in the employed MATLAB development environment on a standard desktop personal computer, expanding our algorithm to allow real-time performance will not be a problem.

In this study, we expanded on our previous work.[16] However, several key elements were improved. First, technical improvements such as execution speed, efficiency and improved post-processing were made. Second, the red-flag indicator was implemented to guide the endoscopist in taking targeted biopsies. Third, the quality of the endoscopic images was vastly improved by the use of the latest version of the FUJIFILM ELUXEO system. Fourth, a gold standard based on the combined input of multiple international experts was created using the proprietary delineation tool. This enabled training of the algorithm with more reliable information. Finally, the neoplastic lesions on the images used in this study were subtler when compared to our previous work. We reasoned that, in order to enable true assistance during surveillance endoscopies, the CAD system should in particular be able to recognize relatively subtle lesions. This study shows that our CAD system is now capable of detecting most subtle lesions on endoscopic images, which would make it a value attribute in a surveillance setting. Such a system could be applied in a clinical setting in which multiple endoscopic overview images are obtained following a structured protocol, which are then directly analysed by the CAD system. In our stepwise approach towards a real-time video-based algorithm, we anticipate encountering new challenges, such as the presence of non-informative frames and the computational power of the CAD system.

This study has several limitations. Our data set comprises only a limited number of images. We have therefore chosen to both train and validate the algorithm on the same data set, via the leave-one-out methodology, instead of validating the algorithm on a separate data set. However, this methodology is often used and is well recognized for machine learning techniques. Furthermore, each image was divided into numerous blocks, thereby increasing our data set to 11.484 data points. In subsequent studies, we will expand our data set and validate the algorithm on a variety of data sets.

The images in this study are of superb quality, collected by expert endoscopists. In a community practice setting, images and videos might be of lower quality.

Finally, it should be noted that our approach, using supervised learning techniques with clinically inspired features, allowed us to closely monitor performance and decisions made by the algorithm. This led to a certain understanding of why the algorithm made its decisions, which allows us to recognize potential pitfalls in the development of a video algorithm. However, this approach was restricted to quantifying colour and texture, in our opinion the main features used by the human eye to discriminate between normal and abnormal tissue. Nevertheless, it is possible that, by following a less strictly supervised approach, the algorithm might identify alternative discriminative features for neoplasia. This methodology is usually referred to as 'deep learning'. As part of the ARGOS project, our group will therefore also focus on the development of a deep learning algorithm in the near future.

In conclusion, in this paper we describe the development of a unique supervised CAD algorithm that detects early neoplastic Barrett's lesions on high-quality endoscopic WLE images with high accuracy. It is therefore an important step towards real-time automated detection of early Barrett's neoplasia. Future work of the ARGOS consortium will focus on improving localization performance and further development of the algorithm towards video analyses, and the development of a deep learning algorithm.

Acknowledgements

Author contributions: Study design: A.J. de Groof, F. van der Sommen, S. Zinger, W.L. Curvers, E.J. Schoon, P.H. de With and J.J. Bergman; patient recruitment and data collection: A.J. de Groof, M.R. Struyvenberg, W.L. Curvers, R. Bisschops, E.J. Schoon and J.J. Bergman; image delineation: O. Pech, A. Meining, H. Neuhaus, R. Bisschops, E.J. Schoon and J.J. Bergman; algorithm development: F. van der Sommen, J. van der Putten, S. Zinger and P.H. de With; and writing of manuscript: A.J. de Groof, F. van der Sommen, W.L. Curvers, E.J. Schoon, P.H. de With and J.J. Bergman. All authors reviewed the final manuscript for important intellectual content and agreed to submit.

Declaration of conflicting interests

All authors declared that they have no disclosures relevant to this manuscript.

Funding

This research is supported by the Dutch Cancer Society and Technology Foundation STW, as part of their joint strategic research programme 'Technology for Oncology'.

Ethics approval

The Netherlands National Trials Registry (number: NTR7072)

Informed consent

Informed consent was obtained from all patients.

References

1. Pech O, May A, Manner H, et al. Long-term efficacy and safety of endoscopic resection for patients with mucosal adenocarcinoma of the esophagus. *Gastroenterology* 2014; 146: 652–660.e1.
2. Spechler SJ, Sharma P, Souza RF, et al. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. *Gastroenterology* 2011; 140: 1084–1091.
3. Weusten B, Bisschops R, Coron E, et al. Endoscopic management of Barrett's esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy* 2017; 49: 191–198.
4. Scholvinck DW, van der Meulen K, Bergman JJ, et al. Detection of lesions in dysplastic Barrett's esophagus by community and expert endoscopists. *Endoscopy* 2017; 49: 113–120.
5. Hvid-Jensen F, Pedersen L, Drewes AM, et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N Engl J Med* 2011; 365: 1375–1383.
6. Sikkema M, de Jonge PJ, Steyerberg EW, et al. Risk of esophageal adenocarcinoma and mortality in patients with Barrett's esophagus: A systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2010; 8: 235–244; quiz e232.
7. Iakovidis DK, Maroulis DE and Karkanis SA. An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. *Comput Biol Med* 2006; 36: 1084–1103.
8. Kominami Y, Yoshida S, Tanaka S, et al. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; 83: 643–649.
9. Maroulis DE, Iakovidis DK, Karkanis SA, et al. CoLD: A versatile detection system for colorectal lesions in endoscopy video-frames. *Comput Methods Programs Biomed* 2003; 70: 151–166.
10. Misawa M, Kudo SE, Mori Y, et al. Characterization of colorectal lesions using a computer-aided diagnostic

system for narrow-band imaging endocytoscopy. *Gastroenterology* 2016; 150: 1531–1532.e3.
11. Takemura Y, Yoshida S, Tanaka S, et al. Computer-aided system for predicting the histology of colorectal tumors by using narrow-band imaging magnifying colonoscopy (with video). *Gastrointest Endosc* 2012; 75: 179–185.
12. Tischendorf JJ, Gross S, Winograd R, et al. Computer-aided classification of colorectal polyps based on vascular patterns: A pilot study. *Endoscopy* 2010; 42: 203–207.
13. 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: 1069–1078.e8.
14. de Groof J, Swager A-F, Pouw RE, et al. 1079 - Blue light imaging (Bli) has an additional value to white light endoscopy (Wle) in visualization of early Barrett's neoplasia. An international multicenter cohort study. *Gastroenterology* 2018; 154: S-209.
15. van der Sommen F, Zinger S, Schoon EJ, et al. Supportive automatic annotation of early esophageal cancer using local gabor and color features. *Neurocomputing* 2014; 144: 92–106.
16. 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.