data adds further evidence to the long term benefit of eradication of dysplastic BE in reducing invasive OAC. Achieving and maintaining CRD in dysplastic BE patients reduces OAC risk to a level similar to patients with non-dysplastic BE.

O30 DEEP NEURAL NETWORK FOR THE DETECTION OF EARLY NEOPLASIA IN BARRETT'S OESOPHAGUS

[1,2]Mohamed Hussein*, [2,3]Juana González-Bueno Puyal, [2,3]Patrick Brandao, [3]Daniel Toth, [1,2]Vinay Sehgal, [1]Martin Everson, [1]Gideon Lipman, [1,2]Omer Ahmad, [1,2]Rawen Kader, [4]Juan Miguel Esteban, [5]Raff Bisschops, [2]Matthew Banks, [3]Peter Mountney, [1]Danail Stoyanov, [1,2]Laurence Lovat, [1,2]Rehan Haidry. *[1]University College London, UK; [2]University College London Hospital, UK; [3]Odin Vision, UK; [4]Hospital Clinico San Carlos, Spain; [5]UZ Leuven, Belgium*

10.1136/gutjnl-2020-bsgcampus.30

Introduction Missed oesophageal adenocarcinoma occurs within one year of index endoscopy in a quarter of patients undergoing surveillance for Barrett's oesophagus (BE). Increased efforts are necessary to improve early neoplasia detection in BE so that early curative endotherapy can be offered. The aim of this study was to develop a deep Neural network that can aid in the diagnosis of early neoplasia in BE.

Methods Videos were prospectively collected in high definition white light and i-scan Pentax (Hoya, Japan) imaging modes in patients with dysplastic lesions in BE (high grade dysplasia/intramucosal adenocarcinoma) and patients with non-dysplastic Barrett's oesophagus. The combination of endoscopic resection margins/targeted biopsy sites and histology served as the ground truth for areas of dysplasia in the videos. All videos with segments of dysplasia were annotated for definite presence of dysplasia. We trained a convolutional neural network with a Resnet101 architecture to classify images into dysplastic or non-dysplastic using randomly selected frames from annotated videos.

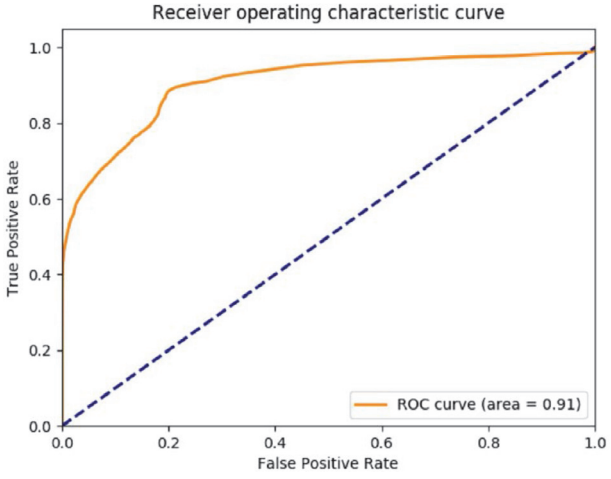
Results A total of 65 different patients each with a video of BE assessment (28 with HGD/Intramucosal cancer and 37 controls) and a total of 266,930 frames were included in the study. The cases were randomly divided into three independent groups: training (N = 39, 17 dysplasia, 22 controls), testing (N = 7, 3 dysplasia, 4 controls) and independent validation set (N =19, 8 dysplasia, 11 controls). For training, a balanced dataset was created with equal numbers of dysplastic and non-dysplastic frames. All available dysplastic frames were used whereas non-dysplastic frames were chosen randomly. The neural network was trained using a total of 73,266 frames. This was then tested on a balanced testing set of 4228 frames.

For the independent validation dataset, all frames in the video sequences were included. This came to a total of 189,436 frames (9194 positive). A dysplastic lesion was only diagnosed if it was deemed present in a sequence of consecutive frames.

The neural network classified dysplasia in BE with a sensitivity of 88.26% and specificity of 80.13% in the independent validation set. The area under the ROC was 0.91.

Conclusion Preliminary data suggest that a convolutional neural network can be trained to detect dysplastic BE mucosa with accuracies similar to expert endoscopists. The algorithm was created from annotations using whole videos which minimised selection bias.

ROC curve analysis for detection of Barrett's dysplasia (figure 1).



Abstract O30 Figure 1

O31 PIXEL-BASED DEEP LEARNING MODELS, A STEP CLOSER TO REAL-TIME COMPUTER-ASSISTED DETECTION OF BARRETT'S NEOPLASIA

[1]Mohamed Abdelrahim*, [2]Masahiro Saiko, [3]Yukiko Masaike, [1]Sophie Arndtz, [1]Ejaz Hossain, [1]Pradeep Bhandari. *[1]Queen Alexandra Hospital, Portsmouth, UK; [2]Biometrics Research Laboratories NEC, Kanagawa, Japan; [3]NEC Europe Ltd, London, UK*

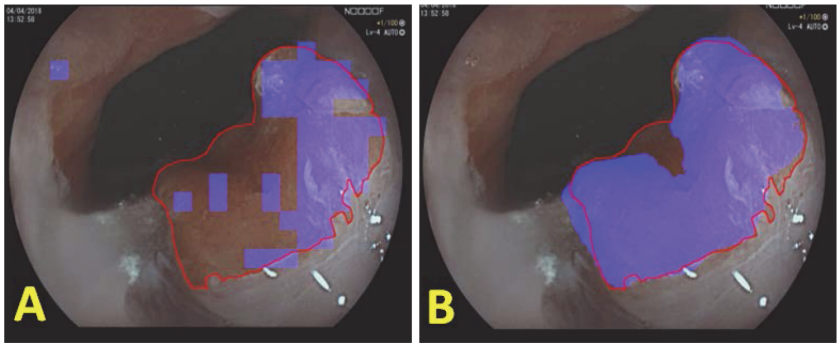
10.1136/gutjnl-2020-bsgcampus.31

Initial results from ***THE PAIGE PROJECT*** ***P***ortsmouth's Project on ***A***rtificial ***I***ntelligence in ***G***astrointestinal ***E***ndoscopy

Introduction Early detection of Barrett's neoplasia is challenging. Computer aided detection through deep learning (DL) is proposed to play a role, with limited recent reports showing encouraging results. However, comparative data on the best methods to develop and implement this technology is lacking. We aim to compare two different DL Artificial Intelligence models for detection of Barrett's neoplasia, a classical patch-based, and a pixel-based model.

Methods We collected 76 anonymous, HDWLE, histologically-confirmed images from our database, including adenocarcinoma, HGD and LGD. For patch-based model, LeNet-5 architecture was used. Each image is divided into patches of 48 × 48 pixels, each patch had a confidence score and label (neoplastic or non-neoplastic). For pixel-based, we used SegNet architecture. Each pixel in the image was given a label and confidence score. Validation performed using 4 fold leave-one-out cross-validations. Graphic processing unit used was 'GeForce RTX 2080 Ti. Processing speed, global accuracy (how often is the model prediction right), F-score (harmonic mean of sensitivity and precision), and IoU (overlap between model prediction and expert marking) were calculated and compared using paired sample t-test.

Results Average lesion detection speed with pixel-based was 33 ms/image, compared with 102.6 ms/image for patch-based model. At a score threshold of 0.8, pixel and patch-based models showed mean values of global accuracy 88% and 84% (P-value 0.00002), IoU 0.40 and 0.21 (P< 0.0001), and F-



Abstract O31 Figure 1 Shows an example of a patch-based prediction (A, left side, Recall 46.3%) and a Pixel-based prediction (B, right side, recall 97.5%). The red line represents the expert endoscopist marking of the lesion, while the blue patch is the DL model prediction

score (for correctly predicted images, at IoU 0.5) 0.81 and 0.69 (P value < 0.0001), respectively.

Conclusions Pixel-based model is significantly faster, and performed better than patch-based model. Given average human visual response latency is estimated at 70–100 ms, this data suggest our pixel-based model with image processing speed of 33 ms/image could potentially detect neoplasia faster than human eye so it will be best suited for real time detection. To our knowledge, this is the first report comparing these two different approaches in Barrett's neoplasia and suggests that all future work should be done with Pixel based model.

O32 DEFINING NORMAL RANGES FOR 96-HOUR WIRELESS PH MONITORING USING HEALTHY CONTROLS

[1]Radu Rusu*, [2]Mark Fox, [3]Emily Tucker, [1]Sebastian Zeki, [1]Jason Dunn, [1]Jafar Jafari, [1]Terence Wong. *[1]Guy's and St Thomas' NHS Foundation Trust, London, UK; [2]Center for Integrative Gastroenterology, Arlesheim, Switzerland; [3]Royal Derby Hospital, Derby, UK*

10.1136/gutjnl-2020-bsgcampus.32

Introduction Technology now allows up to 96-hour wireless pH-recording in clinical practice; however, no normal values exist for this methodology. This study acquired 96-hour esophageal pH recordings in healthy controls (HC) and compared these values against measurements in patients with endoscopic evidence of gastroesophageal reflux disease (GORD). Results were used to validate the Lyon Consensus Classification for GORD Diagnosis.

Methods HC had wireless pH monitoring (Bravo) over 96 hr at two tertiary centers. Bravo was inserted under sedation at endoscopy. Median and 95th percentile values were calculated for the acid exposure time (AET) over 24, 48, 72 and 96-hrs and compared against the 'worst 24-hrs' (i.e. most pathological day). The same analysis was applied to results from a clinical database of consecutive patients with erosive esophagitis (Los Angeles (LA) classification) that completed 96-hr monitoring (acid suppressants stopped ≥5-days previously). A receiver operating curve (ROC) analysis with the area under the ROC curve (AUC) and a Youden's index was also performed to define the optimal cut-off of AET.

Results 71 asymptomatic HC were studied, of whom 47 (age 28±9 years, 66% F) completed 96-hr pH recording. Median (upper 95th percentile) AET was 1.7% (3.0%) for any study day and 2.9% (4.5%) for worst day. 136 patients with reflux esophagitis completed the 96-hr study (61 LA A, 60 LA B, 12 LA C, 3 LA D) - table 1.

Abstract O32 Table 1

HC	96 h Average Day	96 h Worst Day
n	47	47
Median% AET (95th Percentile)	**1.7** (3.0)	**2.9** (4.5)
LA A		
	61	61
	6.1 (7.1)	**9.8** (11.0)
LA B		
	60	60
	8.2 (10.4)	**11.65** (15.1)
LA C		
	12	12
	9.6 (12.1)	**13.2** (16.4)
LA D		
	3	3
	23.4 (38.0)	**27.9** (51.6)

Linear regression analysis revealed a correlation ($p<0.0001$, $R2=33\%$) between endoscopic findings and AET after adjusting for gender, age and duration of the test. ROC analysis for the average AET over 96 hrs differentiated the group with endoscopic evidence of GORD (LA B, C, D) from HC with sensitivity 92%, specificity 75%, positive predictive value (PPV) 77%, negative predictive value (NPV) 91% for a cut-off AET of 3.1%, with AUC 0.91. Similar results were present for the 'worst 24-hr' analysis (sensitivity 91%, specificity 79%, PPV 80%, NPV 90% for a cut-off AET of 5.8%; AUC 0.92).

Conclusions This study defines the normal range for 96-hr ambulatory wireless pH monitoring. These measurements discriminated between HC and patients with conclusive GORD diagnosis, based on endoscopic findings. The findings also validate the diagnostic criteria proposed by the Lyon Consensus. The optimal upper limit of normal in HC was ~4% AET (3.0% average AET, 4.5% worst day AET) and the cut-off to define pathological AET in patients with reflux esophagitis (LA B, C, D) was approximately ~6% AET (worst day 5.8% AET).