

Improving the endoscopic detection of early oesophageal neoplasia

**Dr Martin Anthony Everson
BSc(Hons), MBBS, MRCP(UK)**

**This thesis is submitted for the degree of
Doctor of Medicine (Research)**

Division of Surgery and Interventional Sciences,
National Medical Laser Centre,
University College London

September 2021

Abstract of this thesis

The endoscopic detection of oesophageal cancer is complex; largely owing to the subtle appearances of early oesophageal lesions on endoscopy, as well as clinician experience. Early detection is vital, since lesions confined to the mucosal or superficial layers of the submucosa can be treated with endoscopic eradication therapies to good effect. Conversely, patients presenting with late stage oesophageal cancer have very poor outcomes.

Improving the detection of oesophageal cancer requires a multifaceted approach. Since the symptoms patients present with are often vague until the disease has progressed beyond the point that it is curable, developing a way to risk stratify or rationalise patient access to endoscopy, based on objective markers of the presence of serious underlying pathology, is vital to allow adequate resource provision in the modern UK endoscopy unit. In patients who do undergo endoscopy there remains a significant mis-rate of cancers in those with de-novo oesophageal cancer as well as those enrolled in Barrett's oesophagus surveillance programs. We postulate that advanced imaging technologies, in combination with artificial intelligence systems, may improve the diagnostic performance of endoscopists assessing for oesophageal cancers.

This body of work presents a comprehensive review of the literature surrounding the epidemiology, detection, classification and endoscopic treatment modalities for both squamous cell and adenocarcinomas of the oesophagus. It also presents four studies undertaken with the overarching aim of improving the endoscopic detection of oesophageal cancer. The first study presents a methodology for the quantification of a biomarker from gastric aspirate samples and an assessment of whether differences in expression levels can be used to predict the presence of neoplasia in patients with or without Barrett's oesophagus. The second study investigates the role of a novel, advanced endoscopic imaging technology and whether it improves the diagnostic performance of expert and trainee endoscopists assessing Barrett's oesophagus for the presence of dysplasia or adenocarcinoma. The final two studies present a

significant body of work assessing the feasibility and diagnostic performance of a novel artificial intelligence system designed as part of this thesis, for the detection and characterisation of squamous cell cancer of the oesophagus based on microvascular patterns.

Impact statement

This thesis comprises a suite of four studies undertaken with the overarching theme of improving the endoscopic detection of early oesophageal cancer. Early detection is crucial, since when detected at an early stage curative therapies exist which can be delivered safely endoscopically.

The main academic value of this work comes from the studies focusing on the development of an artificial intelligence system to aid the classification of early squamous cell cancers of the oesophagus. At the time of writing, this work is entirely novel for this application. This thesis provided a proof of concept for this application of artificial intelligence, and furthermore, provided a detailed methodology and results, validated against expert clinicians, which form a benchmark for further clinical studies and academic work. The further refinement and clinical acceptance of this system would be of considerable clinical and commercial interest worldwide. It would be of particular utility in settings which are resource poor, or where clinicians undertaking endoscopy have low exposure or training in the assessment of early cancers. Another study in this thesis assesses, for the first time, a new advanced endoscopic imaging platform for use in the detection of neoplasia associated with Barrett's oesophagus. This work validates a simple endoscopic classification system, which should enable clinicians to better recognise these early lesions – thereby improving patient access to curative therapies. The study described here is novel and contributes to the academic literature which supports the potential that advanced endoscopic imaging may offer. Since completion of this work, each chapter has been peer reviewed and published in leading journals in the field of endoscopy and gastroenterology and has been the subject of numerous conference presentations and posters internationally.

Dedication

To my wife Renee for the unending support and kindness you have devoted to me in both my professional and personal life since we met.

To Dr Rehan Haidry and Prof. Laurence Lovat for taking a chance on a keen foundation doctor five years ago and the support that has continued since.

Acknowledgement of individual contributions

This body of work would not have been possible without the support of many colleague, friends and supervisors; their individual contributions to the original content of this thesis are listed below. More specific acknowledgements of work undertaken by some team members are made within each of the chapters:

Dr Luis Carlos Garcia-Peraza Herrera: undertook the technical development of the convolutional neural network architecture used in the studies presented in chapters 7 and 8. The technical aspects of this work are presented in his doctoral thesis.

Dr Rehan Haidry: acted as my primary supervisor, undertaking and supervising a number of the endoscopic procedures required for this body of work. Dr Haidry provided assistance with study conception, design and analysis as well as revisions of published manuscripts and this thesis.

Prof. Laurence Lovat: acted as my secondary supervisor throughout this work. Prof. Lovat provided assistance with study conception, design and analysis as well as revisions of published manuscripts and this thesis.

Prof. Wen-Lun Wang: provided high quality videos of endoscopic assessments and clinical advice regarding the diagnosis and management of early squamous cell carcinoma of the oesophagus. Prof. Wang provided assistance with study conception, design and analysis as well as revisions of published manuscripts and this thesis.

Mr Paul Bassett: provided independent statistical appraisal, review and guidance for the work presented in chapters 6, 7 and 8.

Dr Stefan Mitrasinovic and Dr Yezen Sammaraiee for developing and providing the bespoke image assessment tool used by experts to assess images of squamous mucosa in chapter 8.

To the teams of expert and trainee endoscopists who undertook hundreds of image assessments for the studies reported in chapters 6, 7 and 8.

Publications arising from this body of work

Peer reviewed publications

Everson, M.A., Ragunath, K., Bhandari, P., Lovat, L. and Haidry, R., 2018. How to Perform a High-Quality Examination in Patients With Barrett's Esophagus. *Gastroenterology*, 154(5), pp.1222-1226.

Everson, M.A., Lovat, L.B., Graham, D.G., Bassett, P., Magee, C., Alzoubaidi, D., Fernández-Sordo, J.O., Sweis, R., Banks, M.R., Wani, S. and Esteban, J.M., 2019. Virtual chromoendoscopy by using optical enhancement improves the detection of Barrett's esophagus-associated neoplasia. *Gastrointestinal endoscopy*, 89(2), pp.247-256.

Everson, M., Magee, C., Alzoubaidi, D., Brogden, S., Graham, D., Lovat, L.B., Novelli, M. and Haidry, R., 2019. Minichromosomal Maintenance Component Complex 5 (MCM5) as a Marker of Barrett's Esophagus-Related Neoplasia: A Feasibility Study. *Digestive diseases and sciences*, 64(10), pp.2815-2822.

Everson, M., Herrera, L.G.P., Li, W., Luengo, I.M., Ahmad, O., Banks, M., Magee, C., Alzoubaidi, D., Hsu, H.M., Graham, D. and Vercauteren, T., 2019. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. *United European gastroenterology journal*, 7(2), pp.297-306.

García-Peraza-Herrera, L.C., **Everson, M.**, Lovat, L., Wang, H.P., Wang, W.L., Haidry, R., Stoyanov, D., Ourselin, S. and Vercauteren, T., 2020. Intrapapillary capillary loop classification in magnification endoscopy: open dataset and baseline methodology. *International journal of computer assisted radiology and surgery*, pp.1-9.

Everson, M., Garcia-Peraza-Herrera, L., Wang, H. P., Lee, C. T., Chung, C. S., Hsieh, P. H., ... & Haidry, R. J. A clinically interpretable convolutional neural network for the real-time prediction of early squamous cell cancer of the esophagus: comparing diagnostic performance with a panel of expert European and Asian endoscopists. *Gastrointestinal Endoscopy*

Garcia-Peraza-Herrera, L.C., **Everson, M.**, Li, W., Luengo, I., Berger, L., Ahmad, O., Lovat, L., Wang, H.P., Wang, W.L., Haidry, R. and Stoyanov, D., 2018. Interpretable fully convolutional classification of intrapapillary capillary loops for real-time detection of early squamous neoplasia. *arXiv preprint arXiv:1805.00632*.

Selected abstracts and oral presentations

Everson, M., Bisschops, R., Wani, S., Graham, D., Banks, M., Magee, C., Alzoubaidi, D., Ahmad, O.O., Ortiz-Fernandez-Sordo, J., Lovat, L. and Sweis, R., 2018. OWE-004 Iscan OE improves detection of early barretts oesophagus associated neoplasia in trainee and expert endoscopists. – oral presentation and abstract of distinction

Everson, M., Bisschops, R., Wani, S., Sordo, J.O.F., Esteban, J.M., Sweis, R., Banks, M., Graham, D., Lovat, L., Ragunath, K. and Haidry, R., 2018. PTH-069 Validating a classification system using ISCAN optical enhancement for detection of early barrett's oesophagus neoplasia.

Everson, M., Herrera, L.G.P., Li, W., Muntion, I.L., Ahmad, O., Graham, D., Banks, M., Lovat, L., Vercauteren, T., Ourselin, S. and Wang, H.P., 2018. ADTH-07 Deep learning based classification of intrapapillary capillary loops for detection of early oesophageal squamous neoplasia.

Everson, M., Herrera, L.C.G.P., Lee, C.T., Chung, C.S., Hsieh, P.H., Chen, C.C., Tseng, C.H., Sammaraiee, Y., Mitrasinovic, S., Lovat, L. and Bergman, J., 2020. Su1019 Predicting histologic invasion depth of esophageal early squamous cell neoplasia using the Japanese Endoscopic Society intrapapillary capillary loop classification; a multicentre comparison of the diagnostic performance in European and Asian endoscopists. *Gastrointestinal Endoscopy*, 91(6),

Everson, M., Herrera, L. G. P., Sammaraiee, Y., Mitrasinovic, S., Hussein, M., Lee, C. T., ... & Haidry, R. (2021). O2 Diagnostic performance of a neural network for the prediction of oesophageal squamous cell cancer. *Gut*

Statement of originality

I, Martin Anthony Everson, confirm that the work presented in this thesis is my own. Any contributions made to the research by colleagues, with whom I have worked at UCL or elsewhere during my candidature, are fully acknowledged.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

Table of Contents

IMPROVING THE ENDOSCOPIC DETECTION OF EARLY OESOPHAGEAL NEOPLASIA	1
ABSTRACT OF THIS THESIS	2
IMPACT STATEMENT	4
DEDICATION	5
ACKNOWLEDGEMENT OF INDIVIDUAL CONTRIBUTIONS	6
PUBLICATIONS ARISING FROM THIS BODY OF WORK.....	7
STATEMENT OF ORIGINALITY	9
TABLE OF FIGURES AND TABLES	13
LIST OF ABBREVIATIONS USED IN THIS THESIS	16
CHAPTER 1 – INTRODUCTION	18
1.1 An overview of oesophageal cancer	18
1.2 Morbidity and mortality associated with oesophageal cancer	20
1.3 The global epidemiology of oesophageal cancer	21
CHAPTER 2 - SQUAMOUS CELL CARCINOMA OF THE OESOPHAGUS	23
2.1 Introduction	23
2.2 Epidemiology.....	23
2.3 The dysplasia – carcinoma sequence in oesophageal squamous cell cancers	24
2.4 Improving squamous dysplasia detection – chromoendoscopy	27
2.5 Improving squamous dysplasia detection – virtual chromoendoscopy	30
2.6 Improving squamous dysplasia detection – magnification endoscopy.....	31
2.7 IPCL classification systems for the recognition of ESCN.....	32
2.7.1 Inoue classification (2001)	33
2.7.2 Arima classification (2005).....	34
2.7.3 The Japanese Endoscopic Society IPCL classification system (2011)	35
2.8 Predicting ESCN invasion depth using the JES classification system	38
2.9 Treatment modalities for oesophageal squamous cell cancers	40
2.9.1 Endoscopic mucosal resection	40
2.9.2 Radiofrequency ablation (RFA)	40
2.9.3 Endoscopic submucosal dissection (ESD)	41
CHAPTER 3 - ADENOCARCINOMA AND BARRETT’S OESOPHAGUS	43
3.1 Introduction	43
3.2 Epidemiology.....	43
3.3 Barrett’s oesophagus	44
3.3.1 Diagnosing Barrett’s oesophagus	45
3.3.2 Barrett’s oesophagus as a risk factor for oesophageal adenocarcinoma	46
3.3.3 The dysplasia – adenocarcinoma sequence in Barrett’s oesophagus	47
3.3.4 Endoscopic surveillance of Barrett’s oesophagus – the Seattle Protocol	51
3.3.5 Limitations of the Seattle Protocol	54
3.4 Adjuncts to improve Barrett’s dysplasia detection - chromoendoscopy	56
3.5 Adjuncts to improve Barrett’s dysplasia detection – virtual chromoendoscopy	57
3.5.1 Narrow band imaging	58
3.5.2 iScan Optical Enhancement	58
3.5.3 Blue laser imaging.....	59
3.6 Endoscopic eradication therapy (EET) of early oesophageal neoplasia	60
3.6.1 Photodynamic therapy (PDT).....	60
3.6.2 Radiofrequency ablation (RFA)	61
3.6.3 Cryoablation.....	62
3.6.4 Endoscopic mucosal resection	62
3.6.5 Endoscopic submucosal dissection	63
3.6.6 UK guidelines for the assessment and endoscopic eradication therapy of Barrett’s neoplasia	64

CHAPTER 4 - ARTIFICIAL INTELLIGENCE IN THE DIAGNOSIS OF OESOPHAGEAL CANCER.....	65
4.1 Defining artificial intelligence.....	65
4.2 Convolutional neural networks (CNNs).....	66
4.3 Previous work on the role of neural networks in early squamous neoplastic lesions	74
STUDY OVERVIEW AND AIMS	77
Study 1 - Minichromosomal maintenance component complex 5 (MCM5) as a marker of Barrett's oesophagus related neoplasia – a feasibility study	77
Study 2 – Assessing whether virtual chromoendoscopy using iScan Optical Enhancement improves the detection of Barrett's oesophagus associated neoplasia in expert and non expert endoscopists.	77
Study 3– Developing a clinically interpretable convolutional neural network to aid in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof of concept study.....	78
Study 4 – Validating a clinically interpretable convolutional neural network for the prediction of early squamous cell neoplasia of the oesophagus; comparing diagnostic performance with a panel of expert European and Asian endoscopists.....	78
CHAPTER 5– MINICHROMOSOMAL MAINTENANCE COMPONENT COMPLEX 5 (MCM5) AS A MARKER OF BARRETT'S OESOPHAGUS RELATED NEOPLASIA – A FEASIBILITY STUDY	80
5.1 Introduction	80
5.2 Aims of this study.....	84
5.3 Methods.....	85
5.4 Results.....	89
5.5 Discussion.....	94
5.6 Summary of this chapter	98
CHAPTER 6– ASSESSING WHETHER VIRTUAL CHROMOENDOSCOPY USING ISCAN OPTICAL ENHANCEMENT IMPROVES THE DETECTION OF BARRETT'S OESOPHAGUS ASSOCIATED NEOPLASIA IN EXPERT AND NON EXPERT ENDOSCOPISTS.	99
6.1 Introduction	99
6.2 Aims of this study.....	102
6.3 Methods.....	103
6.4 Results.....	113
6.5 Discussion.....	119
6.6 Summary of this chapter	124
CHAPTER 7– DEVELOPING A CLINICALLY INTERPRETABLE CONVOLUTIONAL NEURAL NETWORK TO AID IN THE ENDOSCOPIC DIAGNOSIS OF EARLY OESOPHAGEAL SQUAMOUS CELL CARCINOMA: A PROOF OF CONCEPT STUDY.....	125
7.1 Introduction	125
7.2 Aims of this study.....	132
7.3 Methods.....	133
7.4 Results.....	141
7.5 Discussion.....	146
7.6 Summary of this chapter	152
CHAPTER 8 – VALIDATING A CLINICALLY INTERPRETABLE CONVOLUTIONAL NEURAL NETWORK FOR THE PREDICTION OF EARLY SQUAMOUS CELL NEOPLASIA OF THE OESOPHAGUS; COMPARING DIAGNOSTIC PERFORMANCE WITH A PANEL OF EXPERT EUROPEAN AND ASIAN ENDOSCOPISTS.....	153
8.1 Introduction	153
8.2 Aims of this study.....	156
8.3 Methods.....	157
8.4 Results.....	165
8.5 Discussion.....	173
CHAPTER 9 - DISCUSSION AND FURTHER WORK	177
9.1 Minichromosomal maintenance component complex 5 (MCM5) as a marker of Barrett's oesophagus related neoplasia – a feasibility study.....	177
9.2 Assessing whether virtual chromoendoscopy using iScan Optical Enhancement improves the detection of Barrett's oesophagus associated neoplasia in expert and non-expert endoscopists...	179

9.3 Developing a clinically interpretable convolutional neural network to aid in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof of concept study.	182
9.4 Validating a clinically interpretable convolutional neural network for the prediction of early squamous cell neoplasia of the oesophagus; comparing diagnostic performance with a panel of expert European and Asian endoscopists	183
9.5 Conclusions	186
REFERENCES	187
APPENDIX 1 – DATA COLLECTION PROFORMA USED FOR THE MCM5 STUDY	216
APPENDIX 2 – DATA COLLECTION PROFORMA USED FOR THE OE STUDY	219
APPENDIX 3 – PROTOCOL USED FOR IMAGE CAPTURE AND HISTOLOGY RECORDING IN OE STUDY.....	220
APPENDIX 4 – PEER REVIEWED PUBLICATIONS ARISING FROM THIS THESIS.....	228

Table of figures and tables

FIGURE 1 THE MAIN HISTOLOGIC SUBTYPES OF OESOPHAGEAL CANCER	18
FIGURE 2 SCHEMATIC REPRESENTATION (LEFT) AND HISTOLOGIC SPECIMEN DEMONSTRATING THE LAYERS OF THE OESOPHAGUS SEEN MACROSCOPICALLY. HISTOLOGIC SLIDE (RIGHT) DEMONSTRATING THESE LAYERS IN SECTION (FROM UNIVERSITY OF LEEDS ¹). 19	
FIGURE 3 HISTOLOGY SLIDE DEMONSTRATING THE MICROSCOPIC APPEARANCE OF THE OESOPHAGEAL LAYERS. THE LUMINAL EPITHELIUM (EP) COVERS THE MUCOSAL LAYER COMPRISING THE LAMINA PROPRIA (LP) AND MUCULARIS MUCOSA (MM). BENEATH THESE LAYERS ARE THE SUBMUCOSA (SM) AND MUSCULARIS PROPRIA (MP) (FROM DARTMOUTH COLLEGE ¹).....	19
FIGURE 4 AN OVERVIEW OF THE TNM STAGING SYSTEM FOR OESOPHAGEAL CANCER	20
FIGURE 5 UK OFFICE FOR NATIONAL STATISTICS (ONS) 1 YEAR AND 5 YEAR SURVIVAL STATISTICS FOR UK OESOPHAGEAL CANCER DIAGNOSIS BY STAGE	21
FIGURE 6 HISTOLOGIC PROGRESSION OF LGD (TOP), HGD (MIDDLE), CARCINOMA-IN-SITU (BOTTOM) FOR ESCN (FROM PATHOLOGY OUTLINES ¹).....	26
FIGURE 7 MACROSCOPICALLY NORMAL SQUAMOUS MUCOSA PRIOR TO THE APPLICATION OF LUGOL'S IODINE	28
FIGURE 8 LUGOL'S IODINE APPLIED VIA A SPRAY CATHETER	28
FIGURE 9 THE APPEARANCE OF UNSTAINED LESIONS FOLLOWING LUGOL'S APPLICATION	28
FIGURE 10 HD-WLE IMAGING OF THE OESOPHAGEAL SQUAMOUS MUCOSA (LEFT) COMPARED TO NBI (RIGHT).....	30
FIGURE 11 REPRESENTATIVE EXAMPLE OF IPCLs SEEN ON ME-NBI ENDOSCOPY	32
FIGURE 12 AN OVERVIEW OF THE JAPANESE ENDOSCOPIC SOCIETY CLASSIFICATION FOR INTRAPAPILLARY CAPILLARY LOOP PATTERNS.	37
FIGURE 13 DIAGNOSTIC PERFORMANCE OF ENDOSCOPISTS PREDICTING ESCN INVASION DEPTH USING THE JES CLASSIFICATION, ADAPTED FROM ⁶⁹	38
FIGURE 14 ENDOSCOPIC APPEARANCE OF NON-DYSPLASTIC BARRETT'S OESOPHAGUS.....	45
FIGURE 15 HISTOLOGIC APPEARANCE OF NON-DYSPLASTIC BARRETT'S OESOPHAGUS (FROM PATHOLOGY OUTLINES ¹)	48
FIGURE 16 HISTOLOGIC APPEARANCE OF BARRETT'S OESOPHAGUS WITH LOW GRADE DYSPLASIA (FROM PATHOLOGY OUTLINES ¹) ..	49
FIGURE 17 HISTOLOGY APPEARANCE OF BARRETT'S OESOPHAGUS WITH HIGH GRADE DYSPLASIA (FROM PATHOLOGY OUTLINES ¹)...	50
FIGURE 18 HISTOLOGY APPEARANCE OF BARRETT'S ASSOCIATED ADENOCARCINOMA (FROM PATHOLOGY OUTLINES ¹).....	51
FIGURE 19 SCHEMATIC OF HOW THE CM PRAGUE CLASSIFICATION IS RECORDED. CIRCUMFERENTIAL LENGTH (A) AND MAXIMAL EXTENT (B). ADAPTED FROM ¹⁴⁰	53
FIGURE 20 SCHEMATIC OF THE PARIS CLASSIFICATION USED TO DESCRIBE OESOPHAGEAL LESIONS (INDICATED IN GREY).....	54
FIGURE 21 A TYPICAL SUBTLE, FLAT LESION SEEN IN EARLY BARRETT'S ASSOCIATED NEOPLASIA	55
FIGURE 22: PROCEDURE FOR COLLECTING GASTRIC ASPIRATES INTRA-ENDOSCOPY. CLOCKWISE FROM TOP LEFT. 1) GASTRIC FLUID FREE OF FOOD OR BLOOD IDENTIFIED WITHIN THE STOMACH 2) A STERILE PLASTIC SUCTION CATHETER IS PASSED DOWN CLEAN, DRY WORKING CHANNEL OF ENDOSCOPE INTO THE GASTRIC FLUID 3) 5-10ML OF GASTRIC FLUID ASPIRATED FROM STOMACH AND IMMEDIATELY REFRIGERATED BEFORE ANALYSIS. 4) HISTOLOGIC CONFIRMATION OF NEOPLASIA IF PRESENT BY FORCEPS BIOPSY OR EMR.....	87
FIGURE 23 SUMMARY TABLE SHOWING PATIENT DEMOGRAPHICS AND HISTOLOGICAL SUBGROUPS RECRUITED IN THE MCM5 STUDY	89
FIGURE 24 SUMMARY TABLE SHOWING HISTOLOGICAL CHARACTERISTICS OF PATIENTS WITH CONFIRMED ADENOCARCINOMA.	90
FIGURE 25 COMPARISON OF MCM5 EXPRESSION LEVELS IN GASTRIC ASPIRATE SAMPLES (PG/ML) BETWEEN PATIENTS OF EACH HISTOLOGICAL SUBGROUP (AR: MACROSCOPICALLY NORMAL/ACID REFLUX WITHOUT VISIBLE OESOPHAGITIS ONLY, NDBE: NON-DYSPLASTIC BARRETT'S OESOPHAGUS, HGD: HIGH GRADE DYSPLASIA AND CANCER).....	91
FIGURE 26 COMPARISON OF DIFFERENCES BETWEEN MEAN MCM5 EXPRESSION LEVELS AND HISTOLOGICAL SUBTYPES.....	92
FIGURE 27 (LEFT) ROC CURVE (BLUE) FOR THE DIAGNOSTIC PERFORMANCE OF OUR MCM5 EXPRESSION ASSAY FOR THE CHARACTERISATION OF EITHER A MACROSCOPICALLY NORMAL OESOPHAGUS OR ADENOCARCINOMA COMPARED TO THE NULL HYPOTHESIS REFERENCE LINE (RED). (RIGHT) ROC CURVE (BLUE) FOR THE DIAGNOSTIC PERFORMANCE OF OUR MCM5 EXPRESSION ASSAY FOR THE CHARACTERISATION OF PATIENTS AS HAVING NEOPLASTIC HISTOLOGY (ADENOCARCINOMA OR HGD) COMPARED TO NON-NEOPLASTIC HISTOLOGY (NDBE OR NORMAL HISTOLOGY). THE NULL HYPOTHESIS REFERENCE LINE IS SHOWN IN RED.	93
FIGURE 28 SCHEMATIC DIAGRAM OF THE IMAGE PRE AND POST PROCESSING TECHNOLOGY INCORPORATED WITHIN iSCAN OE ENDOSCOPIC IMAGING TECHNOLOGY.	101
FIGURE 29 REPRESENTATIVE EXAMPLE OF HOW IMAGES WERE GENERATED THROUGHOUT THE BE SEGMENT BY CARRYING OUT A STEADY "PULL THROUGH" SEQUENCE TO SIMULATE THE NORMAL ENDOSCOPE WITHDRAWAL MANOEUVRE PERFORMED DURING BE SURVEILLANCE ENDOSCOPY. FROM LEFT TO RIGHT: AT THE DISTAL OESOPHAGUS/GASTRO-OESOPHAGEAL JUNCTION, MID-SECTION OF THE BARRETT'S MUCOSA, AT THE PROXIMAL SQUAMOCOLUMNAR JUNCTION OF THE BARRETT'S SEGMENT. TOP ROW: iSCAN OE, BOTTOM ROW HD-WLE.....	104

FIGURE 30 REPRESENTATIVE EXAMPLE SHOWING HOW THE GOLD STANDARD DELINEATION (YELLOW, RIGHT) WAS GENERATED FROM THE TWO EXPERT DELINEATIONS SHOWN IN THE MIDDLE COLUMN (RED AND BLUE) OF A SUSPICIOUS AREA SEEN HERE IN IScan OE (LEFT). THIS PROCESS WAS REPEATED FOR EACH INDIVIDUAL IMAGE USED IN THE STUDY. THE AREA MARKED IN YELLOW REPRESENTS THE AREA DEEMED POSITIVE FOR DYSPLASIA WHEN THE BLINDED ENDOSCOPISTS MADE THEIR ASSESSMENTS OF IMAGES FOR DYSPLASIA.	106
FIGURE 31 ILLUSTRATIVE EXAMPLE OF THE EXPERT DELINEATED CONSENSUS AREA CONSIDERED POSITIVE FOR DYSPLASIA (RED). ASSESSOR BIOPSY SITES CONSIDERED A TRUE POSITIVE (WHITE) AND FALSE NEGATIVE (BLACK). IMAGES WHERE THE ASSESSOR MADE NO MARK WERE DEEMED TO BE CLASSIFIED AS NORMAL	108
FIGURE 32 REPRESENTATIVE MAGNIFICATION IMAGES OF BARRETT'S OESOPHAGUS CLASSIFIED WITH THE MV CLASSIFICATION AS NON-DYSPLASTIC (TOP RIGHT) AND DYSPLASTIC (OTHER IMAGES), USING HD-WLE	109
FIGURE 33 REPRESENTATIVE MAGNIFICATION IMAGES OF BARRETT'S OESOPHAGUS CLASSIFIED WITH THE MV CLASSIFICATION AS NON-DYSPLASTIC (TOP RIGHT) AND DYSPLASTIC (OTHER IMAGES), USING IScan OE	110
FIGURE 34 AN OVERVIEW OF THE IScan MV CLASSIFICATION FOR USE WITH THE IScan MAGNIFICATION ENDOSCOPY SYSTEM. A SCORE OF M1V1 REPRESENTS NON-DYSPLASTIC BARRETT'S OESOPHAGUS. ANY CLASSIFICATION CONTAINING AT LEAST ONE OF AN M2 OR A V2 REPRESENTS BARRETT'S TISSUE SUSPICIOUS FOR DYSPLASIA.	110
FIGURE 35 STUDY WORKFLOW FOR INVESTIGATING THE UTILITY OF IScan OE IN THE DETECTION OF BARRETT'S NEOPLASIA	112
FIGURE 36 SUMMARY OF LESION HISTOLOGY FOR PATIENTS RECRUITED TO THE IScan OE STUDY	113
FIGURE 37 TRAINEE DIAGNOSTIC PERFORMANCE USING OE COMPARED TO HD-WLE FOR DYSPLASIA DETECTION	114
FIGURE 38 EXPERT DIAGNOSTIC PERFORMANCE USING OE COMPARED TO HD-WLE FOR DYSPLASIA DETECTION	115
FIGURE 39: DIAGNOSTIC PERFORMANCE FOR POOLED ENDOSCOPISTS USING HD-WLE COMPARED TO OE FOR BARRETT'S ASSOCIATED DYSPLASIA DETECTION	116
FIGURE 42 DIAGNOSTIC PERFORMANCE OF EXPERT ENDOSCOPISTS USING MAGNIFICATION ENDOSCOPY AND THE MV CLASSIFICATION TO DETECT BARRETT'S ASSOCIATED NEOPLASIA USING HD-WLE AND OE	117
FIGURE 43 INTEROBSERVER AGREEMENT FOR DYSPLASIA DETECTION USING A MUCOSAL PATTERN ASSESSMENT AND VASCULAR PATTERN ASSESSMENT USING HD-WLE COMPARED TO OE.	118
FIGURE 44 REPRESENTATIVE EXAMPLES OF THE DIFFERENT IPCL MORPHOLOGIES AND THE ASSOCIATED ESCN INVASION DEPTHS AS CLASSIFIED BY THE JAPANESE ENDOSCOPIC SOCIETY ^{69,192}	129
FIGURE 45 HOW THE FINDINGS OF PARTICULAR IPCL MORPHOLOGIES AT ENDOSCOPY CORRELATE WITH THE PRESENCE OF NEOPLASIA AND WHETHER ENDOSCOPIC ERADICATION THERAPY IS INDICATED. ADAPTED FROM ¹	129
FIGURE 46 SCHEMATIC REPRESENTATION OF HOW PATIENTS WERE ASSIGNED TO ALLOW FIVE-FOLD CROSS VALIDATION WITH TRAINING, VALIDATION AND TESTING IMAGE DATASETS IN THIS STUDY	136
FIGURE 47 NUMBER OF FRAMES IN EACH FOLD USED FOR CROSS VALIDATION CONTAINING NORMAL IPCL PATTERNS	137
FIGURE 48 NUMBER OF FRAMES IN EACH FOLD USED FOR CROSS VALIDATION CONTAINING ABNORMAL IPCL PATTERNS	137
FIGURE 49 SCHEMATIC REPRESENTATION OF THE WORKFLOW AND CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE USED IN THIS STUDY	139
FIGURE 50 SUMMARY OF DEMOGRAPHICS AND LESION INFORMATION FOR PATIENTS RECRUITED	141
FIGURE 51 INPUT IMAGES (LEFT COLUMN) WITH CORRESPONDING eCAMs (RIGHT COLUMN), ILLUSTRATING VISUAL FEATURES RECOGNISED BY THE CNN WHEN CLASSIFYING IMAGES. A) RECOGNITION OF ABNORMAL IPCLs PATTERNS. B) SPECULAR REFLECTIONS ARE IGNORED BY THE CNN C) HIGH SELECTIVITY BETWEEN NORMAL MUCOSA AND ABNORMAL IPCLs	143
FIGURE 52 SUMMARY OF CNN PERFORMANCE STATISTICS FOR DETECTION OF ABNORMAL IPCL PATTERNS	145
FIGURE 53 SUMMARY OF THE CNN SIDE OUTPUT INCORPORATED TO ALLOW CLINICAL INTERPRETABILITY OF CLASS ACTIVATION MAPS (CAMs), ADAPTED FROM ²⁴⁵	161
FIGURE 54 HD IMAGE ASSESSMENT PORTAL USED BY THE EXPERT PANELS IN THIS STUDY TO CLASSIFY IPCLs	162
FIGURE 55 SCHEMATIC REPRESENTATION OF THE CHAPTER 8 STUDY WORKFLOW	164
FIGURE 56 SUMMARY OF PATIENT NUMBERS RECRUITED TO THE STUDY BY HISTOLOGIC STAGE OF ESCN	165
FIGURE 57 THE DIAGNOSTIC PERFORMANCE MEASURES FOR THE DETECTION OF ABNORMAL IPCL PATTERNS FOR ALL FOLDS OF THE CNN DEVELOPED IN THIS STUDY	165
FIGURE 58 IMAGES CLASSIFIED WITH THE HIGHEST PROBABILITY BY OUR CNN WHICH WERE SUBSEQUENTLY FOUND TO BE A TRUE POSITIVE (TP) OR NEGATIVE (TN) OR A FALSE POSITIVE (FP) OR NEGATIVE (FN)	166
FIGURE 59 SUMMARY OF THE POOLED DIAGNOSTIC PERFORMANCE FOR BOTH EUROPEAN AND ASIAN EXPERT ENDOSCOPISTS WHEN ASSESSING AND CLASSIFYING ABNORMAL IPCL PATTERNS	167
FIGURE 60 RECEIVER OPERANT CHARACTERISTIC (ROC) CURVE FOR THE CNNs DIAGNOSTIC PERFORMANCE	168
FIGURE 61 INTEROBSERVER AGREEMENT FOR THE CLASSIFICATION OF NORMAL AND ABNORMAL IPCL PATTERNS IN A EUROPEAN AND ASIAN EXPERT ENDOSCOPIST COHORT	168
FIGURE 62 SUMMARY OF EXPERT ENDOSCOPIST PERFORMANCE STATISTICS FOR THE CLASSIFICATION OF TYPE A IPCL PATTERNS	169
FIGURE 63 : SUMMARY OF EXPERT ENDOSCOPIST PERFORMANCE STATISTICS FOR THE CLASSIFICATION OF TYPE B1 IPCL PATTERNS	169
FIGURE 64 SUMMARY OF EXPERT ENDOSCOPIST PERFORMANCE STATISTICS FOR THE CLASSIFICATION OF TYPE B2 IPCL PATTERNS	170

FIGURE 65 SUMMARY OF EXPERT ENDOSCOPIST PERFORMANCE STATISTICS FOR THE CLASSIFICATION OF TYPE B3 IPCL PATTERNS.	170
FIGURE 66 ABSENCE OF VISIBLE SUBMUCOSAL VESSELS MAY INDICATE THE PRESENCE OF ESCN.....	171
FIGURE 67 REPRESENTATIVE EXAMPLES OF CLASS ACTIVATION MAPS GENERATED BY THIS CNN	172

List of abbreviations used in this thesis

Abbreviation (in order of appearance)

SCC	Squamous cell carcinoma of the oesophagus
OAC	Oesophageal adenocarcinoma
BO	Barrett's oesophagus
IM	Intestinal metaplasia
GIST	Gastrointestinal stromal tumour
AJCC	American Joint Committee on Cancer
ONS	Office of National Statistics
ESCN	Early squamous cell neoplasia of the oesophagus
IARC	International Agency for Research on Cancer
NIT	Linxian Dysplasia Nutrition Intervention Trial
LGD	Low grade dysplasia
HGD	High grade dysplasia
USL	Unstained lesion
IMC	Intramucosal cancers
HD-WLE	High definition white light endoscopy
NBI	Narrow band imaging
ME-NBI	Magnification endoscopy with NBI
ME	Magnification endoscopy
IPCLs	Intrapapillary capillary loops
UH-ME	Ultra-high magnification endoscopy
WLE	White light endoscopy
EET	Endoscopic eradication therapy
SM2	Second submucosal layer of the oesophagus
M1	First mucosal layer of the oesophagus
M2	Second mucosal layer of the oesophagus
M3	Third mucosal layer of the oesophagus
AVA	Avascular areas
JES	Japanese Endoscopic Society
ER	Endoscopic resection
CT	Computerised tomography scan
MRI	Magnetic resonance imaging
PPV	Positive predictive value
NPV	Negative predictive value
EMR	Endoscopic mucosal resection
RFA	Radiofrequency ablation
ESD	Endoscopic submucosal dissection
SEER	US Surveillance, Epidemiology and End Results program
GORD	Gastro-oesophageal reflux disease
GOJ	Gastro-oesophageal junction
BSG	British Society for Gastroenterology
ACG	American College of Gastroenterology
BOBCAT	Benign Barrett's and CAncer Taskforce

GM	Gastric metaplasia
ASG	American Society of Gastroenterology
NDBO	Non-dysplastic Barrett's oesophagus
NICE	National Institute for Clinical Excellence
RCTs	Randomised control trials
QALY	Quality adjusted life year
SP	Seattle protocol
BING	Barrett's International NBI Group
OE	iScan optical enhancement
BLI	Blue Laser Imaging
PDT	Photodynamic therapy
MDT	Multidisciplinary team
AI	Artificial intelligence
NN	Neural networks
CNNS	Convolutional neural networks
RBG	Red-green-blue
MCM5	Minichromosomal maintenance complex 5
CAMs	Class activation maps

Chapter 1 – Introduction

1.1 An overview of oesophageal cancer

Oesophageal cancer is a potentially devastating condition, with an almost uniformly dire prognosis if detected late in its disease course. The eighth most common cause of cancer deaths worldwide¹, it is a condition with significant morbidity and mortality in both the developed and developing world.

The two main subtypes of oesophageal cancer are squamous cell carcinoma (SCC) and adenocarcinoma (OAC). The first typically arises from the normal squamous lined oesophageal epithelium and is more commonly seen in the developing world, particularly countries on the ‘Cancer Belt’ – sub-Saharan Africa, Iran, Afghanistan, China and Japan². The latter most commonly arises from Barrett’s oesophagus (BO) - glandular tissue containing intestinal-type metaplastic (IM) tissue in the distal half of the oesophagus³. There are other subtypes of oesophageal cancer (**Figure 1**) but owing to their relative rarity, they will not be discussed in great length in this thesis.

Subtypes of oesophageal cancer

Squamous cell carcinoma (SCC)
Adenocarcinoma (OAC)
Melanoma
Lymphoma
Gastrointestinal stromal tumour (GIST)
Leiomyosarcoma

Figure 1 The main histologic subtypes of oesophageal cancer

Oesophageal cancer is an aggressive form of cancer, with relative rapid growth both intraluminally and through the oesophageal layers (**figure 2⁴**) into the structures adjacent to the oesophagus (aorta, trachea, diaphragm). The subdivisions of the oesophageal layers are also relevant to the treatment options indicated in patients. This will be discussed further in chapters 2 and 3 but a schematic of these layers, along with relevant abbreviations used in this thesis are displayed in (**figure 3⁵**). Presenting symptoms often include dysphagia and pain, typically due to extensive local spread.

Oesophageal cancer is usually diagnosed based on endoscopic findings, with histologic analysis of sampled tissue or based on radiologic appearances in patients unable to undergo endoscopy.

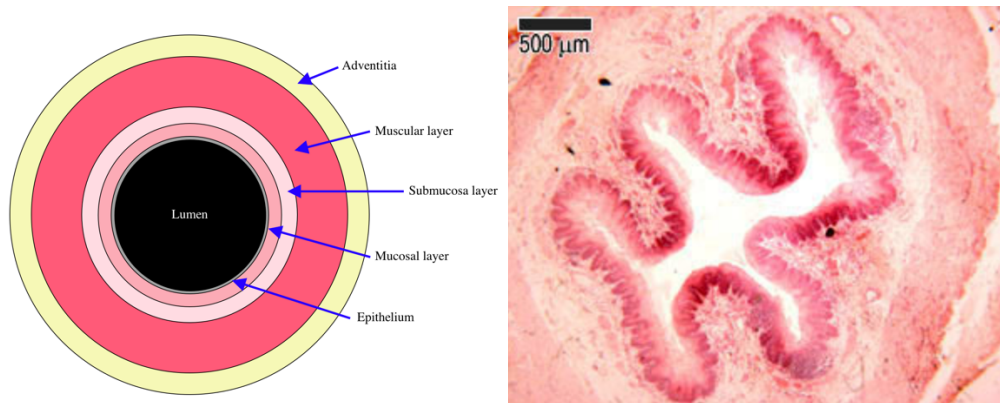


Figure 2 Schematic representation (left) and histologic specimen demonstrating the layers of the oesophagus seen macroscopically. Histologic slide (right) demonstrating these layers in section (from University of Leeds⁴).

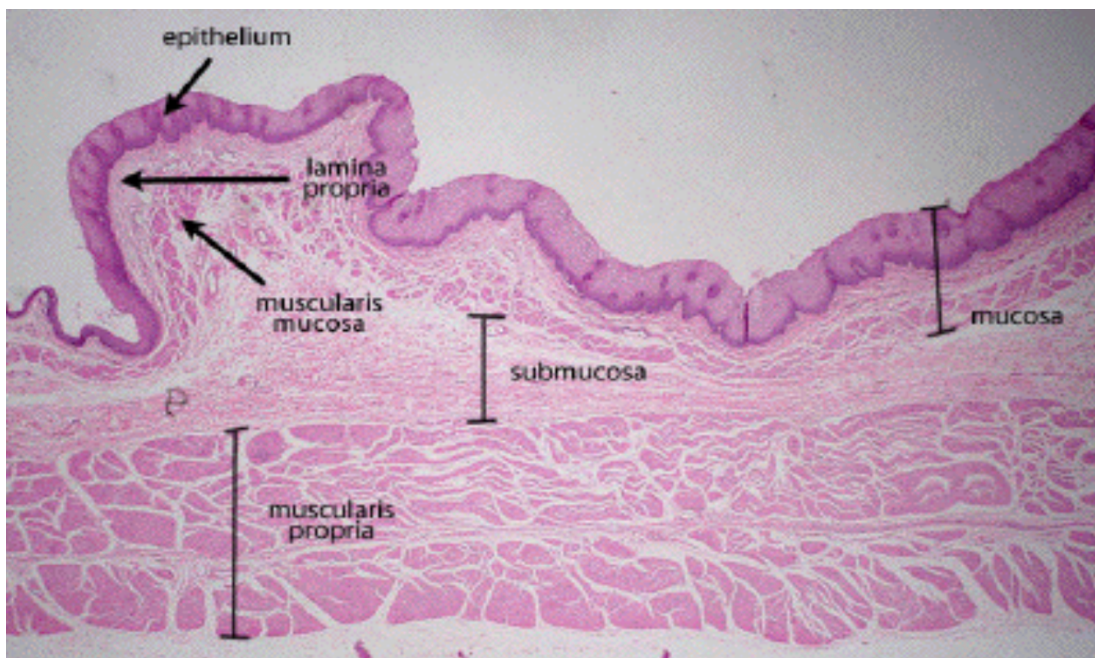


Figure 3 Histology slide demonstrating the microscopic appearance of the oesophageal layers. The luminal epithelium (EP) covers the mucosal layer comprising the lamina propria (LP) and muscularis mucosa (MM). Beneath these layers are the submucosa (SM) and muscularis propria (MP) (from Dartmouth College⁵)

In the UK, oesophageal cancer is staged conventionally according to the TNM staging system proposed by the American Joint Committee on Cancer (AJCC)⁶. As outlined in (figure 4) this staging system reflects the degree of local and distant spread and allows prognostication and selection of treatment modalities in patients diagnosed.

TNM Staging	
T (local staging)	<i>Invasion depth</i>
T _{is}	High grade dysplasia
T1	Lamina propria, muscularis mucosa, submucosa
T2	Muscularis propria
T3	Adventitia
T4a	Invades resectable local structures (pleura, diaphragm)
T4b	Invades unresectable local structures (trachea, aorta)
N (nodal staging)	<i>Extent of node involvement</i>
N0	None
N1	1 – 2 positive nodes
N2	3 – 6 positive nodes
N3	>7 positive nodes
M (metastatic staging)	<i>Extent of distal spread</i>
M0	No distant metastasis
M1	Distant metastasis

Figure 4 An overview of the TNM staging system for oesophageal cancer

1.2 Morbidity and mortality associated with oesophageal cancer

Around 70% of oesophageal cancers are detected late in the course of the disease⁷, largely due to the lack of symptoms, which typically only occur when lesions are large enough to cause pain or dysphagia,. In the UK in 2014, at least 59% of cancers were diagnosed at stage III or above^{8,9}. The attendant mortality associated with late stage oesophageal cancer makes it the fourteenth most common cause of cancer deaths in the UK with similar estimates reflecting the poor overall five-year survival of less than 15% worldwide¹⁰.

UK data compiled by the Office of National Statistics (ONS) between 2013 and 2017, indicates that as the cancer stage at diagnosis increases the estimated 1 and 5 year survival rates drop dramatically¹¹. For patients diagnosed with stage 1 oesophageal cancer the estimated 1 and 5 year survivals are 84.5% and 52.8% respectively. In contrast for patients diagnosed at stage 4, the 1 and 5 year survival is 20.8% and unrecordably low respectively. Patients with stage 3 cancer at diagnosis are reported

as having 5 year survivals of 16.3% (**figure 5**). There is therefore a clear need to ensure that patients with oesophageal cancer have early access to diagnostic tests and that physicians are capable of recognising the symptoms of early oesophageal cancer. One of the biggest issues facing clinicians is that this type of cancer rarely generates symptoms until it is advanced, so the detection of early lesions is often incidental.

Oesophageal cancer survival by stage		
Stage at diagnosis	1 year survival (%)	5 year survival (%)
T1	84.5	52.8
T2	68.3	29.9
T3	54.8	16.3
T4	20.8	not recorded

Figure 5 UK Office for National Statistics (ONS) 1 year and 5 year survival statistics for UK oesophageal cancer diagnosis by stage

1.3 The global epidemiology of oesophageal cancer

When considering the epidemiology of oesophageal cancer the two main subtypes, SCC and OAC have notable differences and so will be considered separately.

SCC is the predominant histologic type of oesophageal cancer worldwide, with a particular geographic distribution described as the 'oesophageal cancer belt'; extending from Eastern and Sub-Saharan Africa, through the Middle East and into China^{12,13}. Incidence rates of SCC in these countries are disproportionately high and can be up to 100 cases/100000 person years¹². Risk factors for SCC include smoking¹⁴, alcohol intake and possibly the consumption of preserved meat or vegetables, the incidence of which is more common in the developing world and eastern hemisphere and decreasing in the western world¹⁵. It should be noted however that the increased incidence in these areas may not be explained entirely by exposure to such external or environmental factors; for instance in areas of China where female smoking rates are comparably low¹⁶, the incidence of SCC remains high, or in predominantly Islamic countries such as Iran tobacco and alcohol use is low¹⁷, but SCC incidence remains high.

In the western world, OAC now represents a more emergent health burden than SCC¹⁵. In the UK alone there has been a three-fold increase in the incidence of adenocarcinoma since the early 1970s¹⁸, with similar trends seen across other developed nations^{3,19}. The underlying cause for this has not been fully characterised but may be related to changes in lifestyles. The most studied risk factor for the development of OAC is the concomitant presence of Barrett's oesophagus (BO), which is more commonly seen in patients with long term exposure of the oesophageal mucosa to refluxed gastric acid or bile¹⁵. Over 90% of patients diagnosed with OAC have BO²⁰, but estimates of the true incidence of OAC arising from BO vary^{20,21}, since in a large number of patients their BO may remain undetected until a diagnosis of cancer is made. Old age and male sex have also been identified as independent risk factors for the development of OAC. The epidemiology and risk associated with BO and oesophageal cancer will be discussed in more detail in **chapter 3**.

Chapter 2 - Squamous cell carcinoma of the oesophagus

2.1 Introduction

This chapter will outline current evidence in the natural history, epidemiology and pathogenesis of squamous cell carcinoma of the oesophagus. Following this it will discuss the endoscopic diagnosis and endoscopic treatments available for early squamous cell neoplasia (ESCN) of the oesophagus. This chapter aims to outline how current practice in the endoscopic imaging of ESCN may be improved by the use of advanced endoscopic imaging devices such as magnification endoscopy. In **chapters 7 and 8**, this thesis will discuss how in conjunction with these advanced imaging modalities, artificial intelligence may assist expert and non-expert endoscopists in the endoscopic diagnosis of ESCN.

2.2 Epidemiology

Squamous cell cancer (SCC) of the oesophagus is the most prevalent histological subtype of oesophageal cancer worldwide². According to the International Agency for Research on Cancer (IARC), 88% of the 450,000 reported cases of oesophageal cancer identified in 2012, were caused by squamous cell carcinoma².

There is a significant variation in the geographic distribution of SCC worldwide; a disproportionately high incidence of the disease affects populations along two 'oesophageal cancer belts' which extend from Central to Eastern Asia and from the Middle East into the west coast of sub-Saharan Africa^{12,13}. Incidence rates of SCC in these countries are disproportionately high and can be up to 100 cases/100000 person years¹². Of note there is significant variation even within countries with a high incidence of the disease, with apparent local discrepancies in the cases observed²². Owing to its large population and high incidence, it is believed that almost half of all global cases of SCC in the world are reported in China^{22,23}. Globally, SCC affects men more commonly than women; in low risk areas such as the UK and USA the ratio of men to women affected approaches 4:1, whereas in high incidence countries such as China or Iran this ratio is likely closer to 1:1².

Studying the global incidence of SCC is complex since most published studies examine only high-risk populations such as those seen in Northern China, Iran and Africa. What is clear however is that the incidence over the last 50 years has dropped significantly in the UK²⁴, USA²⁵ and other western nations, whilst remaining high in central Asia, Africa and the far East.

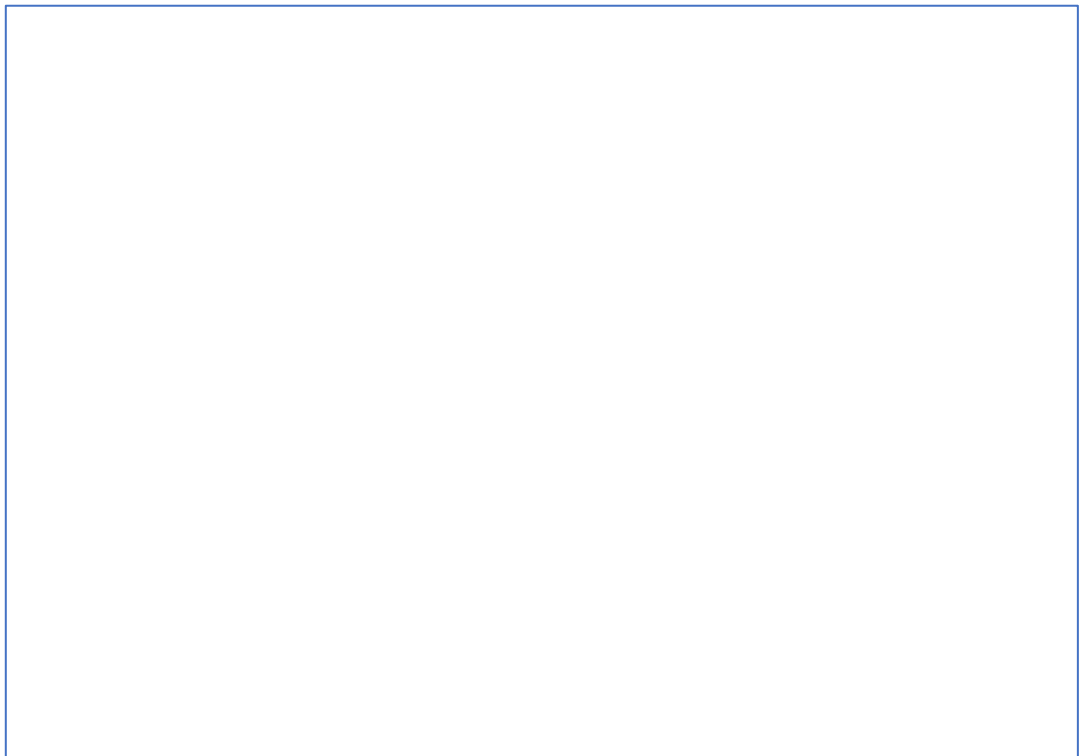
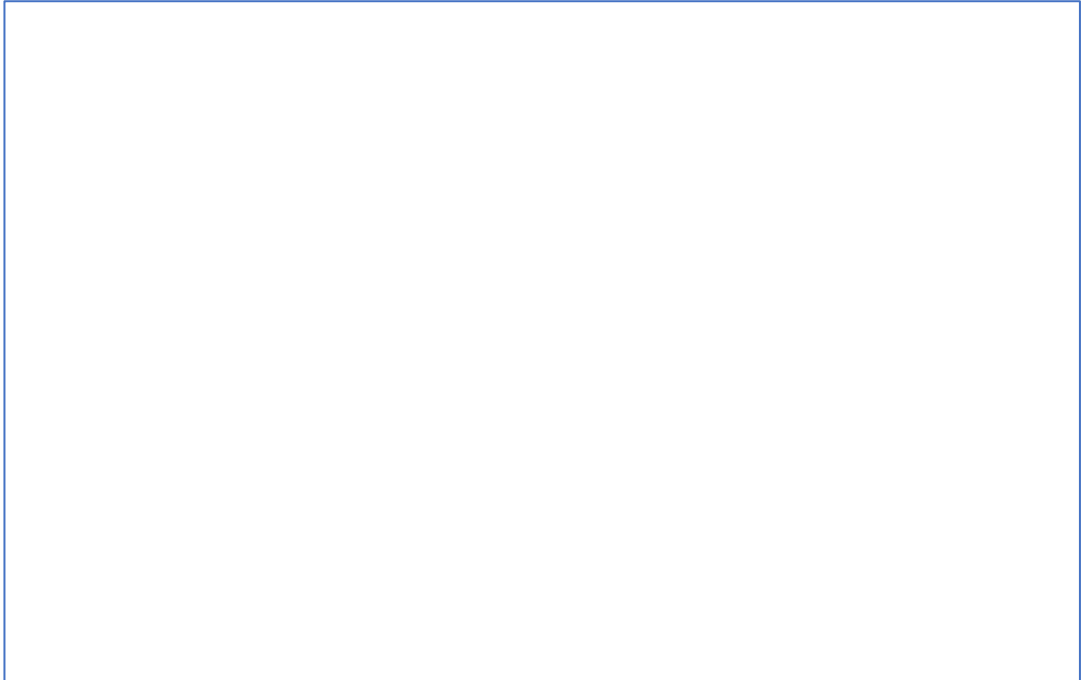
Numerous risk factors have been identified for the development of SCC, but they likely do not represent the whole causality for the disease. Smoking, increases the relative risk of SCC development by 3-9 fold, particularly in developed nations^{26,27}. The size of this effect is less pronounced in developing countries, particularly those with a high incidence of the disease^{16,17,23}. Excessive alcohol consumption has also been linked to an increased relative risk of SCC²⁸. As with tobacco smoking it increases the risk in studies of Chinese²⁹, Japanese²⁹ and Iranian¹⁷ patients by between 1.5 and 5 fold, but appears to increase the risk more significantly (up to 9 fold) in European populations³⁰⁻³². Other risk factors include exposure to polycyclic hydrocarbons, betel nut consumption, pickled foods and low dietary fruit and vegetable consumption³³⁻³⁶.

2.3 The dysplasia – carcinoma sequence in oesophageal squamous cell cancers

Squamous cell carcinoma of the oesophagus typically presents late in its disease course, hence the study of the early development of lesions was primarily confined to high-risk populations. Initially it was believed that SCC most likely arose from areas of oesophagitis; Qiu and Yang were the first to demonstrate that specific dysplastic changes in the oesophageal mucosa indicative of the subsequent development of SCC rather than oesophagitis³⁷.

The evolution of squamous dysplasia to SCC is stepwise and the diagnosis is histological. Dysplastic squamous epithelium exhibits a range of architectural abnormalities which typically includes nuclear enlargement, pleomorphism, mitotic figures and hyperchromasia (**figure 6**)^{38,39}. Squamous dysplasia can be classified as low grade dysplasia (LGD) or high grade dysplasia (HGD); with the above changes visible in the basal half of the epithelium in low grade and extending into the top half and

surface in high grade. If dysplastic squamous tissue remains undetected, there is a risk that it can develop into a squamous cell carcinoma, with progressive invasion into the mucosal layer and beyond.



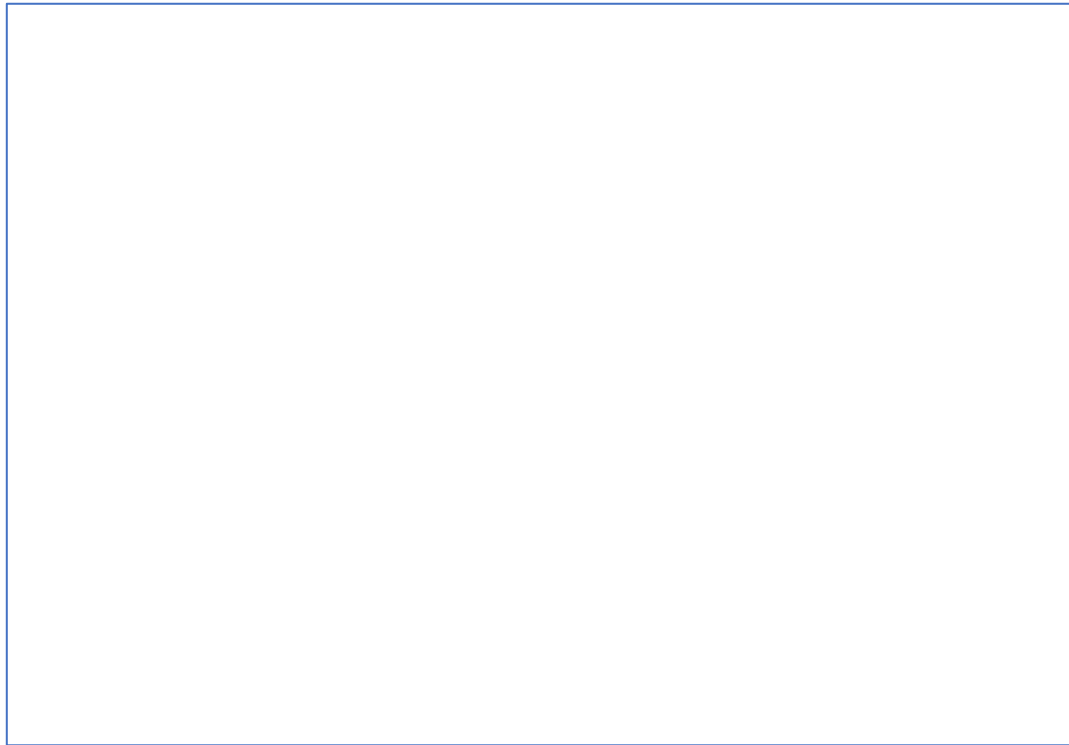


Figure 6 Histologic progression of LGD (top), HGD (middle), carcinoma-in-situ (bottom) for ESCN (from Pathology Outlines available at ³⁹)

Evidence from the large, prospective Linxian Dysplasia Nutrition Intervention Trial (NIT) demonstrated that the presence of dysplasia was a predictive factor for the development of SCC over the next 3.5 years⁴⁰. In this study, the presence of mild, moderate or severe dysplasia was associated with a worsening relative risk of developing SCC of 2.2, 15.8 and 72.6 respectively⁴⁰. These results were again confirmed in the same cohort, with even higher relative risks after 13.5 years of follow up⁴¹. The risks for those with severe dysplasia are even higher, with Dawsey et al. demonstrated higher rates of progression to SCC of 60%⁴². Estimates of the prevalence of squamous dysplasia vary widely in studies depending on the specific populations and endoscopic screening methods employed, but are estimated to be within the range of 2-38%^{43,44,45,46}. Since the incidence of SCC is lower in Western populations there is scarce literature to quantify the prevalence of squamous dysplasia in low risk populations.

The gold standard for the diagnosis of squamous dysplasia is gastroscopy with endoluminal biopsies. Identifying squamous dysplasia is key to reducing the incidence

of SCC, since prompt recognition can allow for the endoscopic therapies discussed later in this chapter to prevent lesion progression. On endoscopic assessment, dysplastic lesions may appear as a white plaque or patch, or as an eroded lesion with a friable surface⁴⁷. Since lesions are subtle, flat and variable in appearance, they are easily missed on endoscopy, particularly in low incidence areas where endoscopists may be unfamiliar with the appearance of dysplasia.

2.4 Improving squamous dysplasia detection – chromoendoscopy

Owing to the challenges of identifying early squamous lesions, since the early 1990s numerous studies have examined the utility of real-time chromoendoscopy using a solution of Lugol's iodine applied topically to the mucosa during endoscopy. Iodine stains glycogen molecules within the normal squamous mucosa a brown to orange colour. Dysplastic or neoplastic mucosal tissue is often glycogen deplete and so following the application of Lugol's they either remain unstained or quickly void their brown colour compared to the surrounding normal mucosa. These unstained lesions (USLs) can therefore be target biopsied (**figures 7, 8 and 9**).

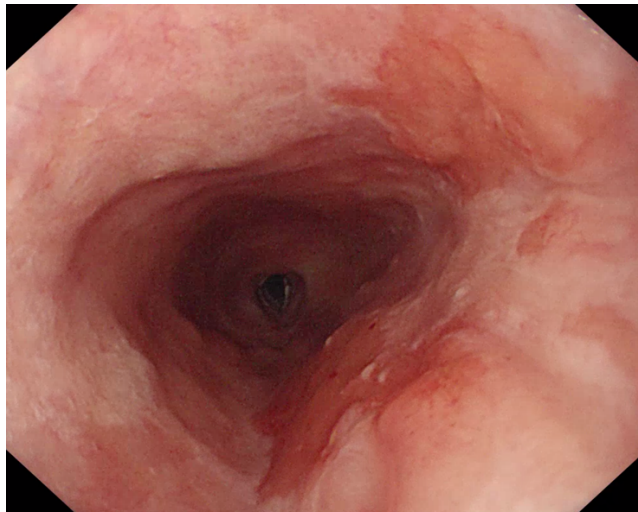


Figure 7 Macroscopically normal squamous mucosa prior to the application of Lugol's iodine

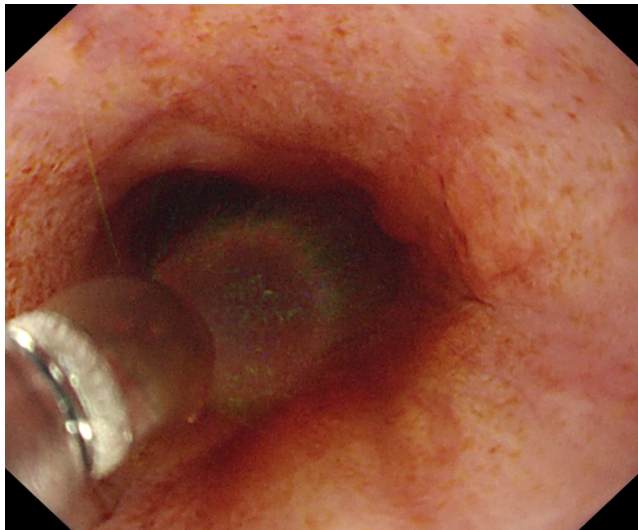


Figure 8 Lugol's iodine applied via a spray catheter

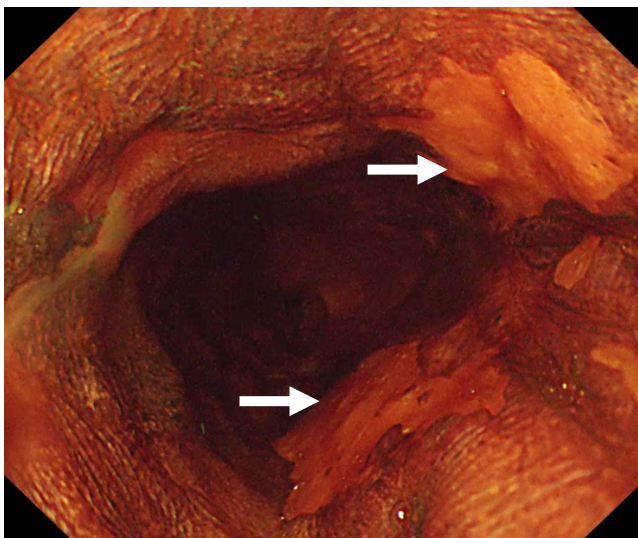


Figure 9 The appearance of unstained lesions following Lugol's application

Reported sensitivities and specificities using Lugol's for ESCN detection are high. Per patient analyses of sensitivity range from 80%-100%⁴⁸⁻⁵² and for specificity from 0.52-0.94⁴⁸⁻⁵². A large meta-analysis however suggests that per lesion specificity may be as low as 37% when using Lugol's due to its propensity to highlight benign lesions such as papillomas⁵¹.

A prospective study conducted in China demonstrated sensitivities of 63%, 93%, 96%, 100% for the detection of low or high grade dysplasia and intramucosal cancers (IMC) and invasive SCC respectively, following the application of Lugol's⁵³. Similarly metachronous oesophageal SCC detection rates are higher in patients with head and neck SCCs who undergo endoscopic screening with Lugol's^{54,55}.

In a low incidence region of China, squamous dysplasia was detected in 7.3% of patients of undergoing high definition white light (HD-WLE), this increased to 13.7% in patients in those who underwent HD-WLE with the addition of Lugol's⁵⁶. This suggests that use of Lugol's may improve the endoscopic detection of squamous dysplasia and SCC in low incidence settings such as the developed world.

The routine use of Lugol's is associated with a number of potential issues that may limit its use in endoscopic ESCN screening. The depletion of glycogen within the oesophageal mucosa is associated with a degree of oesophageal spasm and irritation, which may be uncomfortable for the patient. Iodine hypersensitivity reactions or aspiration pneumonias involving Lugol's can cause significant harm. Lastly, studies assessing the use of Lugol's have demonstrated that it is poorly specific for the detection of ESCN – which in addition to the resource and procedural time required to stain the mucosa, may also contribute to a high number of unnecessary biopsies being taken⁵³. Given the inherent issues with the use of Lugol's, endoscopic technologies to provide the visual enhancement of chromoendoscopy, without the need for topical agents have garnered much attention in recent years.

2.5 Improving squamous dysplasia detection – virtual chromoendoscopy

A number of virtual chromoendoscopy platforms are available to endoscopists. More detail on the technical aspects of the various virtual chromoendoscopic technologies is provided in **section 3.5**, since their use in a Western healthcare setting is better studied in the context of detecting dysplasia in Barrett's oesophagus. This section outlines the current evidence for the use of enhanced imaging technologies in early squamous neoplasia.

Narrow Band Imaging (NBI, Olympus, Japan) is the most widely researched virtual chromoendoscopy platform used in ESCN assessment (**figure 10**), although its value remains contested. Studies comparing NBI to HD-WLE are limited since the use of Lugol's iodine is now widespread and effective. From the reported literature NBI appears to offer favourable conditions for the detection of ESCN in the oropharynx and oesophagus over HD-WLE. A tandem trial involving 113 patients randomised to undergo either NBI or HD-WLE endoscopic assessments demonstrated a significant improvement in ESCN detection rates when NBI was used (70.2% vs 35.7%)⁵⁷.

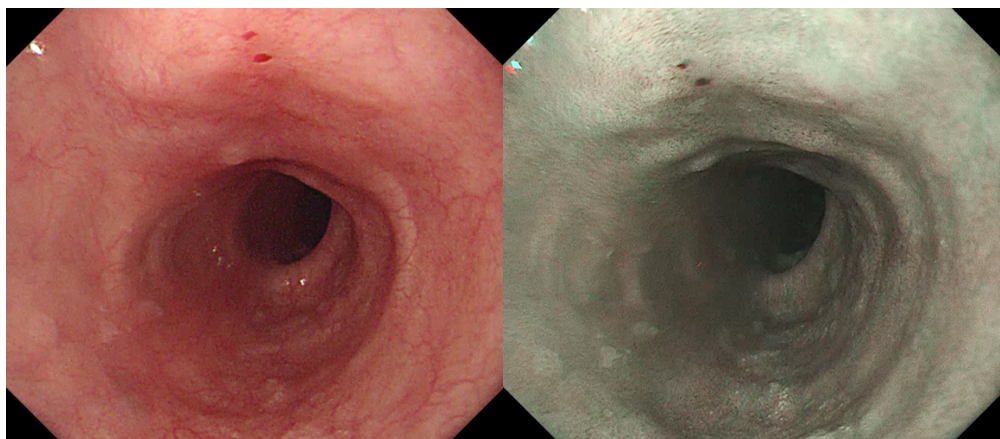


Figure 10 HD-WLE imaging of the oesophageal squamous mucosa (left) compared to NBI (right)

Muto et al, in a multicentre randomised control trial demonstrated high diagnostic performance for NBI in the detection of superficial ESCN in both the head, neck and oesophagus. Comparing HD-WLE and NBI showed a significantly higher diagnostic accuracy for NBI and sensitivities and specificities of 55% vs 97% and 63% and 42% respectively⁵⁸.

A meta-analysis comprising 12 studies and 1911 patients compared the use of Lugol's or NBI in the detection of HGD or ESCN. In a per lesion analysis using Lugol's and NBI, sensitivity was 98% and 94% respectively, specificity 37% and 65% respectively. This suggests that NBI is comparable to Lugol's in the detection of ESCN, and is likely better at differentiating it from other oesophageal pathology⁵¹.

NBI has been predominantly used in combination with magnification endoscopy (ME-NBI) discussed further in **section 2.6**. Several studies have demonstrated that in combination with ME, NBI confers diagnostic accuracies of between 76.8 – 85.2%^{59,60}.

iScan (Pentax, Japan), is a post-processing imaging technology. The use of iScan in the detection of ESCN is not well documented. A single centre, prospective, non-inferiority trial suggests that it offers a lower detection rate for ESCN than HD-WLE in combination with Lugol's; 10.4% compared to 12.9% respectively⁶¹. Further work is needed to assess whether it has a role in the detection of ESCN.

2.6 Improving squamous dysplasia detection – magnification endoscopy

Magnification endoscopy, used in conjunction with high-definition white light and virtual chromoendoscopy may provide an additional benefit to endoscopists in the identification of early lesions. The utility of magnification endoscopy in the oesophagus was first investigated by Inoue et al⁶². Ultra-high magnification endoscopes, capable of continuous magnification of the mucosa at resolutions of up to 150x, facilitated the interrogation of the microvascular structures seen within the mucosa. This study reported a previously unreported mucosal microvessel – these small looped vessels were defined as intrapapillary capillary loops (IPCLs). Although they were recognised as a feature of the normal oesophageal mucosa, previous magnification endoscopes were only capable of up to 35x resolution so the IPCL structures had not been visualised.

Normal IPCLs are fine calibre looped structures which arise from the deep submucosal vessels of the oesophagus. These vessels course into the basal layer of the

oesophageal epithelium and run adjacent to it. Within this layer IPCLs can be seen as red dots on conventional or low magnification endoscopy. On magnification however they are visualised as delicate looped structures as demonstrated in (**figure 11**).

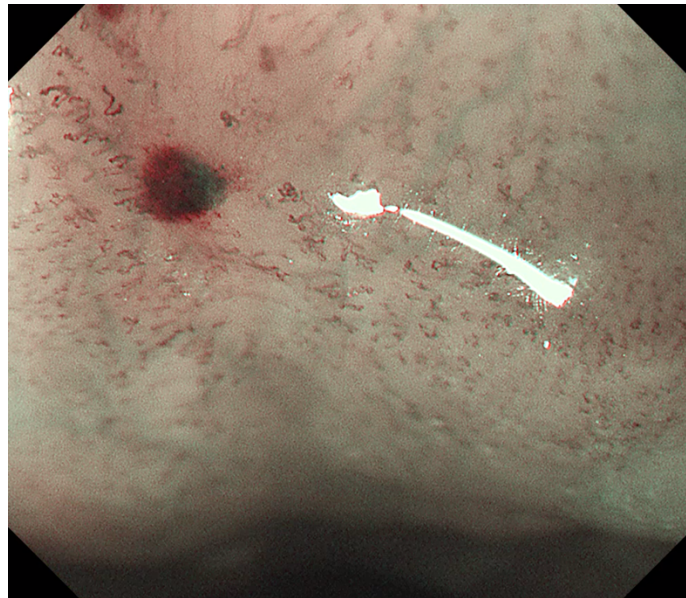


Figure 11 Representative example of IPCLs seen on ME-NBI endoscopy

Ultra-high magnification endoscopy (UH-ME) enables the interrogation of the oesophageal mucosa in both normal and diseased states. In 1997, Inoue et al. reported the use of UH-ME in combination with WLE to assess flat lesions associated with squamous dysplasia and early squamous cell neoplasia (ESCN). When interrogating these lesions with UH-ME they noted abnormal IPCL morphologies. They reported that the morphology of these IPCLs in ESCN lesions were dilated, irregular and more tortuous than those seen in the normal oesophageal mucosa or in oesophagitis⁶³. Lesions which exhibited these abnormal IPCL patterns were confirmed histologically to be either squamous dysplasia or ESCN; progressive changes in the IPCL morphology could be correlated with the progression of neoplasia.

2.7 IPCL classification systems for the recognition of ESCN

Morphological changes in IPCL patterns are linked to the presence of squamous dysplasia and ESCN. Various classification systems have been developed to identify

the various normal and abnormal morphologies and their correlation with histologic findings in resected samples.

2.7.1 Inoue classification (2001)

The initial ICPL classification system was proposed by Inoue et al⁶⁴. This classification subcategorised IPCLs into 5 groups and considered four distinct morphological features as well as their appearances after staining with topical Lugol's iodine. The four abnormal IPCL morphologies described included weaving, dilation, irregular calibre or inconsistent shapes in visualised IPCLs. Type I IPCLs were normal calibre, regular and looped vessels that stained positively with Lugol's and were found in normal oesophageal mucosa. Type II IPCLs could have up to two of the above described morphological changes, typically elongation and mild dilation of the ICPLs was observed. Type II IPCLs, stain positively with Lugol's and are characteristically seen in patients with oesophagitis. Abnormal IPCL patterns are described as type III-Vn. Type III IPCLs are morphologically similar to type II but negatively stained with iodine and were associated with low grade dysplasia. Type IV IPCLs are seen in areas of high grade dysplasia, demonstrating two or three abnormal morphological features and remain unstained with Lugol's. Type V^{1,2,3,N} IPCLs are the most abnormal patterns described and both exhibit all four morphological abnormalities while remaining completely unstained with Lugol's. Recognising IPCLs representing each of the subdivisions requires some degree of interpretation. Type V¹ correlates with an ESCN invading the first mucosal layer (M1) and is regarded as exhibiting all four of the above abnormalities. V² is described as an 'extension' of the V¹ pattern with more prominent morphologic change; these correlate with lesions invading the second mucosa layer (M2). Type V³ IPCLs demonstrate advanced destruction of the IPCL vessels with the formation of large and sometimes linear, dilated vessels. This pattern is seen within lesions that have invaded into the boundary of the mucosa (M3) and the first submucosal layer (SM1). In larger lesions that have invaded into the second submucosal layer (>SM2) or deeper, there is often a degree of neovascularisation, with the formation of new, grossly dilated vessels and almost complete destruction of

the normal IPCL morphology. The type V IPCL patterns are associated with advancing squamous carcinoma of the oesophagus and span the endoscopically recognisable changes seen in endoscopically resectable lesions and surgically resection or palliative lesions. The Inoue classification has been well validated by multiple studies, some of which are discussed below.

Sato et al (2015) prospectively assessed and classified endoscopic images of Inoue type V lesions⁶⁰. Two endoscopists assessed images of 446 type V lesions and classified them as type V^{1-N} on two occasions with a washout period in between. They demonstrated high accuracies of 91% and 86% for the classification of V¹ and V² lesions with substantial interobserver and intraobserver agreement. They demonstrated reduced accuracies for the classification of V³ lesions of 43%, which is clinically relevant as it meant that 28% of expert assessments underestimated lesions invading into the second submucosal layer (SM2) that were in fact not amenable to EET. It was not clearly reported how many mucosal lesions were over diagnosed as submucosal lesions; in clinical practice this could mean that a patient with a resectable lesion may be referred for surgical management rather than endoscopic therapy.

Goda et al demonstrated that the Inoue classification was less sensitive for the distinction of mucosal from submucosal lesions; which is of clinical relevance since the only mucosal lesions are an absolute indication for endoscopic eradication therapy (EET). ME-NBI in combination with the Inoue classification offered a sensitivity of 78%, which was not significantly better than normal high resolution WLE or EUS (sensitivity of 72 and 83% respectively) for the classification of these lesions⁵⁹. Further work by Goda et al demonstrated that in early squamous lesions, use of the Inoue classification allowed the correct identification of more than 90% of LGD and HGD lesions⁶⁵.

2.7.2 Arima classification (2005)

Arima et al proposed a relatively simplified IPCL classification system and compared the IPCL patterns seen on 405 UH-ME images of superficial oesophageal lesions, including 191 cancers⁶⁶. They classified IPCL patterns into four distinct types. Type 1

IPCLs were described as regular, looped and thin calibre capillaries that were observed within the subepithelial papillae of the normal oesophageal squamous mucosa. Type 2 vessels retained the looped structure seen in type 1 vessels, but were of a dilated calibre. These vessels were typically seen in oesophagitis. Type 3 vessels were defined as tortuous, dilated, spiral vessels with red spots. The arrangement of the vessels was irregular and appeared 'flattened' compared to the more linear type 1 and 2 vessels. Type 3 vessels were typically seen in cancers invading the first or second mucosal layer (M1 or M2). The final IPCL subgroup described was type 4 vessels, all of which correlated with invasion to at least the third mucosal layer (M3) or beyond. The type 4 vessels were subdivided as multi layered, irregular branched or reticular; all of which in their various forms were dilated and irregularly arranged compared to the normal vessels. Arima et al also introduced the concept of avascular areas (AVAs), regions of mucosa free of any vasculature that were observed between abnormal IPCLs. The presence and increasing size of these AVAs were associated with more advanced lesions that invaded the submucosal layers.

The Arima classification was internally validated and demonstrated accuracies of 94% for the prediction of lesions containing HGD-M2 ESCN tissue. Their simplified system also afforded accurate identification of >M3 lesions in up to 90% of cases, although 10% of lesions were overestimated⁶⁶. This system, with fewer subdivisions of ICPL patterns within the >M3/SM lesions, enabled better differentiation of lesions by clinicians of mucosal lesions compared to submucosal lesions. A limitation of the Arima classification is its grouping of IPCLs seen in mucosal and submucosal lesions as type 4 vessels. As noted previously it is important to differentiate mucosal lesions accurately, since unlike some submucosal lesions, they can be treated endoscopically.

2.7.3 The Japanese Endoscopic Society IPCL classification system (2011)

A 2011 simplified classification system for magnified endoscopy images of IPCL patterns was proposed by Oyama et al⁶⁷, which was subsequently adopted and validated by the Japanese Endoscopic Society (JES) in 2012⁶⁸. The JES criteria use

morphological abnormalities seen in IPCL patterns to determine if ESCN was present. A secondary diagnostic criteria utilised AVA size to assess the likely invasion depth of these lesions.

Using the JES classification system, IPCL patterns were defined as normal (type A) or abnormal (type B), with abnormal vessels subclassified as type B1, B2 or B3. Type A vessels are small, narrow calibre looped vessels that are typically found within the normal oesophageal mucosa. More elongated or dilated type A vessels may also be seen in oesophagitis.

Type B vessels are associated with progressively more advanced squamous neoplasia and are all indicated by the presence of morphologically abnormal IPCL patterns. Type B1 vessels retain their looped structure but are more tortuous and dilated and may be associated with a red/brown discoloration of the mucosa. Type B1 vessels often appear more densely arranged than type A vessels due to their dilation. Type B2 vessels are severely irregular, dilated and are associated with complete loss of their normal looped arrangement. Type B2 vessels are often seen to follow a tortuous horizontal running pattern across the mucosa. Type B3 vessels are a more aberrant formation of severely dilated and tortuous B2 vessels. B3 vessels are described as being at least three times the diameter of the type B2 vessels which often surround them. The endoscopic appearance of the various IPCL patterns described in the JES classification are shown in (*figure 12*).

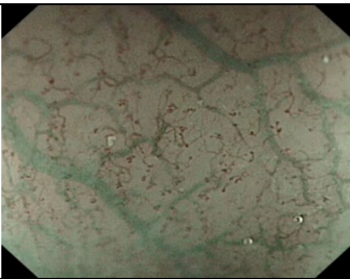
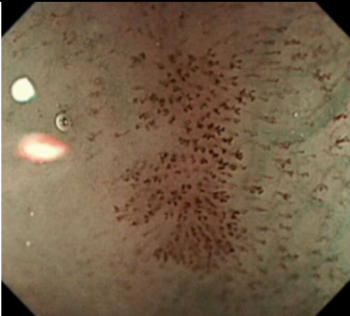
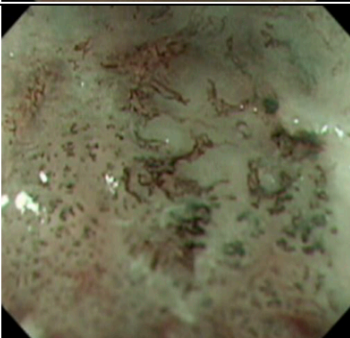
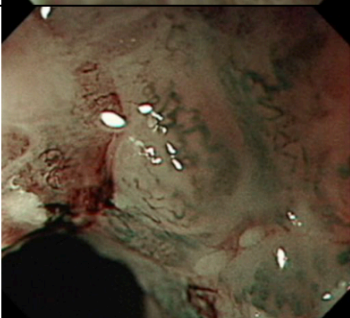
JES type		Typical morphology	Typical histology
A		Small, non-dilated, fine calibre looped vessels with no gross abnormality. Some vessels may take a more elongated form with inflammation or LGIN. Submucosal vessels may be visible with the background mucosa of a uniform colour	Normal/LGIN
B1		Subtle abnormalities in IPCLs observed. Increased tortuosity, increased vessel calibre and density but looped structure is retained. There may be a brown hue to the mucosa	HGIN/LP
B2		Grossly abnormal IPCLs. Linear vessel formation with loss of the normal loop structure. There is also gross dilation and tortuosity of vessels. Formation of avascular areas between vessels observed	MM/SM1
B3		Dilated non-looping vessels typically 3x calibre of B2 vessels seen. IPCLs are tortuous and highly irregular in appearance. Evidence of extensive neovascularisation and avascular areas seen. Mucosal contour may be distorted. Other adjacent IPCLs are typically abnormal	SM2 or deeper

Figure 12 An overview of the Japanese Endoscopic Society classification for intrapapillary capillary loop patterns.

2.8 Predicting ESCN invasion depth using the JES classification system

A prospective multicentre study by Oyama et al assessed the predictive value of the JES magnification classification for determining the likely invasion depth of ESCN⁶⁹. 211 consecutive patients with SCC were enrolled across 5 referral centres. Expert endoscopists were required to classify the IPCL patterns in these ESCN lesions according to the JES classification. They then predicted the invasion depth of the ESCN lesions which were imaged using an ME-NBI endoscopy system. Histological analysis of endoscopic resection (ER) specimens was used to determine invasion depth and CT/MRI imaging was used to confirm or exclude lymph node or distant metastatic spread.

The accuracy of the JES classification as reported in this study was high – with the overall accuracy for histology prediction 90.5% across type B1-3. Using the JES classification clinicians correctly predicted the invasion depth of lesions containing type B1, B2 and B3 IPCL patterns with 97.5%, 75% and 55% sensitivity respectively. Overall accuracy for their predictions were 91.9%, 93.4% and 95.9% for type B1, B2 and B3 IPCL patterns respectively. The diagnostic value of the JES IPCL classification system are outlined in **(figure 13)**:

IPCL classification	Sensitivity (%)	Specificity (%)	PPV (%)	Accuracy (%)
B1	97.5	72.9	92.4	91.9
B2	75.0	96.2	75.0	93.4
B3	55.0	100	100	95.9

Figure 13 Diagnostic performance of endoscopists predicting ESCN invasion depth using the JES classification, adapted from⁶⁹

The overall accuracy of the JES IPCL classification is sufficiently high for it to have utility in clinical use; it is now widely used by endoscopists. It is also arguable that the JES classification relies less on subjective interpretations of the IPCL patterns, and due to few subdivisions may be more accessible to endoscopists with less experience assessing ESCN. The sensitivity of type B1 vessels for endoscopically resectable ESCNs that extend only as far as the lamina propria was high at 97.5%, with the PPV of B1

vessels for the presence of these lesions similarly high at 92.4%. Although less high, the sensitivity and PPV of type B2 IPCLs for the presence of ESCN extending as far as SM1 was 75% for both measures. This suggests that type B1 and to a lesser extent type B2 vessels offer a highly sensitive endoscopic marker for the presence of endoscopically resectable ESCN. In the Oyama study interobserver agreement was not assessed and these results reflect the diagnostic ability of expert endoscopists, familiar with the assessment and treatment of ESCN lesions. It should be noted that they may not reflect the accuracies of less experienced endoscopists.

Kim et al assessed the accuracy and interobserver agreement using the JES IPCL classification⁷⁰. ME-NBI images of 70 lesions, which were confirmed histologically as SCC, were reviewed by two experienced endoscopists. This study reported an overall accuracy for prediction of ESCN invasion depth of 78.6%. Importantly they demonstrated an accuracy of 88.6%, 78.6% and 90% for the correct prediction of invasion in lesions classified as having type B1, B2 or B3 IPCL patterns respectively. Although this study only compared the classification by two endoscopists, the accuracies achieved using the JES system compared favourably to the Inoue classification. Interobserver agreement for classifying lesions as type B1 was excellent at 0.75.

A recent study by Wang et al (2017), assessed the utility of ME-NBI in addition to HD-WLE using the JES IPCL classification to accurately predict cancer invasion depth, demonstrating mixed results. In an experienced group of endoscopists accuracy using HD-WLE alone was 37%, this increased to 51% when using HD-WLE in combination with ME-NBI. In an inexperienced group of endoscopists the accuracy with HD-WLE alone was 34% and HD-WLE with ME-NBI was 31%. These figures may be representative of the diagnostic accuracy of the average endoscopist, raising the suggestion that miss rates of ESCN on endoscopic assessment may be high. These results may also suggest that in inexperienced hands, the use of ME-NBI may confer no advantage in ESCN detection. Additional training in the use of ME-NBI or adjuncts such as artificial intelligence systems which provide a 'red flag' for endoscopist to help identify ESCN may ameliorate this lack of experience.

2.9 Treatment modalities for oesophageal squamous cell cancers

2.9.1 Endoscopic mucosal resection

Endoscopic mucosal resection (EMR) is a minimally invasive and safe method for the resection of early, localised squamous lesions of the oesophagus. Using a distal cap attachment or multiband mucosectomy device the suspicious area of mucosa can be ligated and resected. EMR was pioneered in Japan in the early 1990s⁷¹ and has since been used extensively in the management of ESCN with a long established record of good outcomes. Early studies on the use of EMR suggested excellent durability and 5 year survival rates of up to 95%⁷².

Debate exists as to whether EMR is suitable for the management of lesions invading into the muscularis mucosa, since the rate of recurrence and metastasis is not negligible compared to epithelial or lamina propria lesions. A large multi-centre study suggests however that of 104 patients with ESCN invading the MM, only 1.9% went on to develop lymph node metastasis after EMR (median follow up 43 months)⁷³.

The durability of EMR is largely dependent on the ESCN invasion depth at the time of index treatment Yamishina et al. report a large cohort study of 570 patients with ESCN treated by ER. Patients were followed up over 15 years, with a mean follow up of 50 months. The 5 year overall survival rates for epithelial, mucosal and submucosal ESCN was 90.5%, 71.1% and 70.8% respectively. The cumulative 5 year incidence of metastasis for epithelial, mucosal and submucosal ESCN was 0.4%, 7.7% and 36.2% respectively⁷⁴.

2.9.2 Radiofrequency ablation (RFA)

The use of radiofrequency ablation (RFA) is well established in patients with Barrett's oesophagus associated neoplasia. ESCN confined to the mucosa has also been investigated as a target for RFA therapy, provided that the lesion is flat, is unstained on Lugol's and does not invade beyond the muscularis mucosae. RFA comprises the delivery of high frequency energy via a bipolar electrode, either via a

balloon catheter inflated within the oesophageal lumen or via a catheter placed on the endoscope tip. The application of RFA induces intense heating within the diseased tissue, which in turn induces tissue damage and cell death within the treated area. This local destruction of tissue can be utilised to clear dysplastic tissue.

Numerous studies of RFA in ESCN demonstrate that it may be effective for superficial ESCN. A retrospective study undertaken in a high volume Chinese centre demonstrated that 84% of 96 enrolled patients had complete remission of their squamous LGD/HGD at 12 months; with reported stricture rates of 21%⁷⁵. The same cohort demonstrated 86% of patients had sustained remission of their dysplasia over a 5 year follow up period (median follow up 48 months)⁷⁶.

Wang et al. demonstrated in 30 patients, followed up over a mean follow up period of 40.1 months, that 71.4% achieved complete remission of dysplasia. Over the follow up period no patients developed lymph node metastasis or cancer progression, and of the 20% who developed locally recurrent neoplasia, all were treated completely with EMR or endoscopic submucosal dissection (ESD)⁷⁷.

Outcomes from the UK RFA registry suggest a more mixed efficacy for RFA in ESCN. Of 20 patients who completed an RFA treatment protocol, only 50% remained dysplasia free at 12 months of follow up but 30% progressed to invasive cancer at 1 year⁷⁸.

2.9.3 Endoscopic submucosal dissection (ESD)

Endoscopic submucosal dissection is another established treatment for ESCN. The use of ESD had two primary advantages in the case of ESCN; firstly it allows for en bloc resections, which reduces recurrence and aids histopathologic delineations of resection margins; secondly it facilitates the resection of lesions which extend deep to the muscularis mucosa and into the submucosa.

Oyama et al report en bloc resection rates of 95% in a cohort of 102 patients undergoing ESD for ESCN over a four year period. Overall a mean follow up period of

21 months, there was a 0% recurrence rate. 7 patients out of 102 developed a benign stricture but none experienced a perforation⁷⁹.

Nagami et al. further assessed 5yr survivals for patients undergoing ESD for superficial ESCN. 107 lesions were resected from 84 patients, one cohort had lesions superficial to the lamina propria and the other were deep to it. Over a 5 year follow up period cause specific survival was 100% and 85% respectively. 18% of patients in this study developed endoscopically treatable strictures, 4% experienced perforation and none had major bleeding⁸⁰.

A similar single centre retrospective study demonstrated favourable 5 year survivals in 94 patients undergoing ESD. 5 year survivals in patients with ESCN superficial to the lamina propria or deep to it were 100% and 88% respectively, again indicating that ESD is a highly effective treatment for superficial ESCN⁸¹.

A multi-centre study, undertaken in a Western healthcare setting by Berger et al. suggests that for lesions invading up the oesophageal wall up to SM1, ESD appeared to be preferable to EMR⁸². The ESCN recurrence rate in patients treated with EMR was 23.7%, compared with just 2.9% in the ESD subgroup. Similarly 5yr recurrence free survival rates favoured treatment with ESD – 95.2% treated with ESD survived five years without a recurrence, compared to 73.4% in the EMR group. This underscores once more that the accurate recognition of likely invasion depth in ESCN lesions and the appropriate triage to endoscopic, or surgical therapies is vital to improve patient survival.

Chapter 3 - Adenocarcinoma and Barrett's oesophagus

3.1 Introduction

This chapter will outline current evidence in the natural history, epidemiology and pathogenesis of adenocarcinoma of the oesophagus, with a specific focus on the diagnostic challenges presented by Barrett's oesophagus (BO). Advanced endoscopic imaging modalities present an opportunity to improve the recognition of early Barrett's associated neoplasia, and so improve prompt access to the endoscopic treatments discussed in this chapter. After presenting the limitations of endoscopic surveillance methods for Barrett's oesophagus, **chapter 5 and 6** will present two studies; the first on a novel biomarker for the detection of oesophageal cancer and dysplastic Barrett's, the second will validate a novel advanced endoscopic imaging system for the same purpose.

3.2 Epidemiology

Prior to the 1980s oesophageal adenocarcinoma was regarded as an uncommon cause of cancer. Representing between 0.8-3.7% of all reported oesophageal cancers, it was far outweighed in incidence by squamous carcinoma¹⁰. Since that time numerous population based studies from the USA and western Europe have demonstrated a sharp increase in its incidence. Data from the US Surveillance, Epidemiology and End Results (SEER) program suggested that the incidence of OAC in white American males nearly doubled between the 1970s to 1980s¹⁰. Similar trends were observed in the UK, Denmark and other developed nations⁸³⁻⁸⁵. Current data suggests that the incidence continues to rise into the last decade such that the incidence of OAC now exceeds that of SCC in most of the western world⁸⁶. At present, OAC is the most common histological subtype of oesophageal cancer seen in the western world; accounting for around 53% of cases and with an incidence of 0.7 cases per 100,000 person years². The UK currently has one of the highest country specific incidences of OAC in the world at 7.2 cases per 100,000 person years in men².

OAC is predominantly a disease of males, with an estimated incidence ratio of 9:1 for men compared to women^{87,88}. Although it typically affects Caucasians to a greater

extent than non-Caucasians, the reasons for this are unclear^{88,89}. The incidence of OAC is greatest in patients over 50 and tends to peak in males between 55-65 years. Of concern over the last few decades up to the early 2000s is that the incidence of OAC in males under 65 has doubled, whereas the incidence in patients over 65 has increased up to 3-fold in the same time^{89,90}.

Numerous risk factors for the development of OAC have been proposed. Gastro-oesophageal reflux disease (GORD) has also been identified as a risk factor for OAC in several large meta-analyses of population studies; posing an estimated 6.2 fold risk in affected patients⁹¹⁻⁹³. The role of obesity and central abdominal adiposity has also been implicated as increasing the risk of OAC, conversely there is some evidence that weight loss may reduce the risk⁹⁴⁻⁹⁷. In common with SCC, tobacco smoking has been demonstrated to raise the risk of OAC up to 3-fold and yet in contrast alcohol consumption appears in some studies to have little or an unclear effect⁹⁸⁻¹⁰¹. Interestingly the presence of *Helicobacter pylori* infection appears to have a protective effect on colonised individuals; some studies show a reduction in risk of between 40-60%¹⁰²⁻¹⁰⁴. The tendency to eradicate *H. pylori* infection in western healthcare systems in recent decades, along with the increasing incidence of OAC would seem to support this. Another well documented risk factor, and one relevant to the contents of this thesis is the presence of Barrett's oesophagus⁹⁸. This entity is therefore discussed in more detail in the following sections.

3.3 Barrett's oesophagus

Barrett's oesophagus (BO) was first described in 1950 by Norman Barrett, a British surgeon who noted a columnar-lined oesophagus above the GOJ in patients with oesophagitis¹⁰⁵ (**figure 14**). Barrett's initial hypothesis was that BO was in fact gastric mucosa that was identified within an intrathoracic stomach due to the shortened oesophagus of affected patients. In 1957 Barrett reported that these macroscopic findings were due to the replacement of the squamous epithelium of the oesophagus with columnar epithelium¹⁰⁶. It was not until 1975 that Naef and colleagues linked the presence of BO with the development of oesophageal adenocarcinoma (OAC)¹⁰⁷.

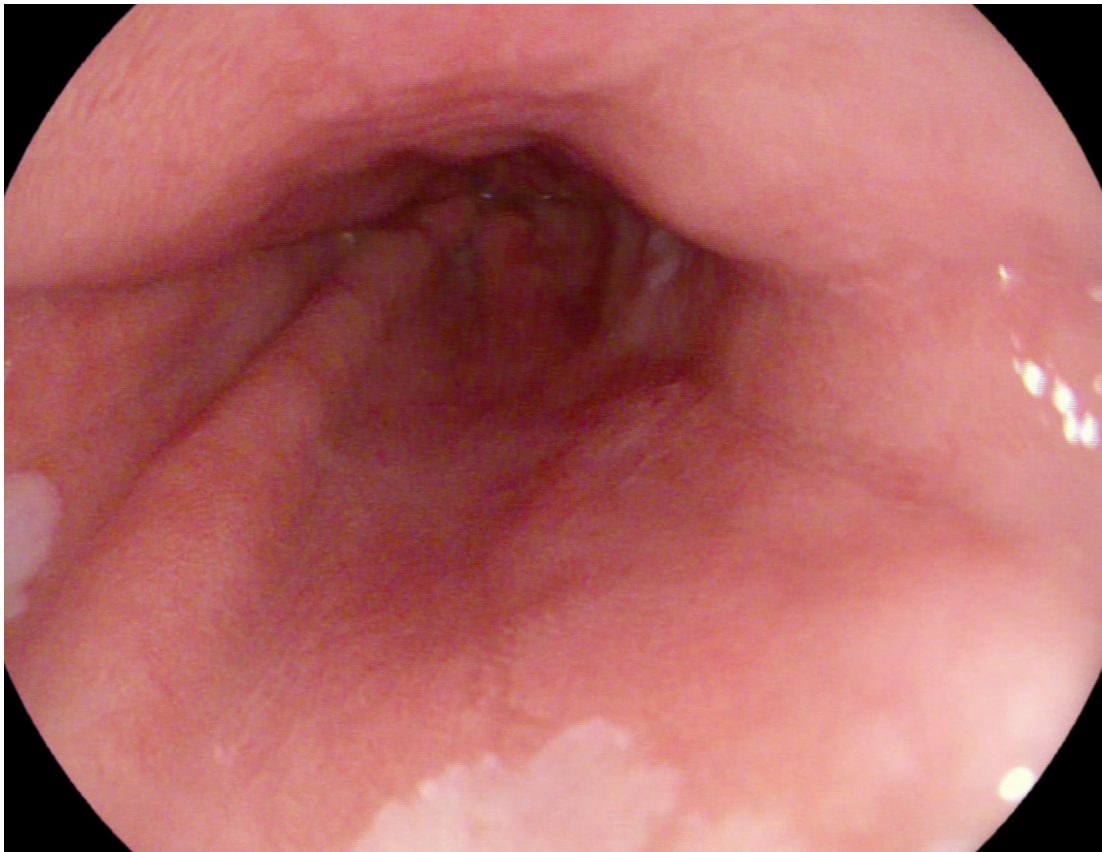


Figure 14 Endoscopic appearance of non-dysplastic Barrett's oesophagus

3.3.1 Diagnosing Barrett's oesophagus

In current practice the diagnosis of Barrett's oesophagus relies on a typical endoscopic appearance, which is confirmed by histologic analysis. Endoscopically, the BO mucosa is readily visible under high definition white light (HD-WLE), as a salmon pink villous mucosa extending above the gastrointestinal junction (GOJ). There is international consensus on the endoscopic landmarks associated with a formal diagnosis of BO. The British Society of Gastroenterology (BSG), American College of Gastroenterology (ACG) and the Benign Barrett's and CAncer Taskforce (BOBCAT) 2015 international consensus all state that columnar epithelium must be visible more than 1cm above the GOJ^{108–110}. The endoscopic landmark for the GOJ is defined by all three guidelines as the top of the gastric folds.

There remains some international variation in the histologic features required for a formal BO diagnosis. The BSG recommend that there must be biopsy confirmation of columnar metaplasia, which includes gastric metaplasia (GM) in addition to intestinal

metaplasia (IM)¹⁰⁸. American Society of Gastroenterology (ASG) guidelines stipulate that intestinal metaplasia is a prerequisite for a diagnosis of BO¹⁰⁹. The BOBCAT consensus recommends that any biopsy confirmed columnar metaplasia is sufficient but that special note by the pathologist of IM above the GOJ should be noted by the pathologist¹¹⁰.

The contention over the inclusion of histologically confirmed IM is due to conflicting evidence over its role in the premalignant potential of BO, which is discussed below. Several studies have demonstrated that the presence of IM in BO confers an increased risk of adenocarcinoma, but that in BO with only GM or cardia type metaplasia there was a low incidence of adenocarcinoma. A 2012 study by Chandrosoma et al demonstrated no proven risk of adenocarcinoma in patients with BO without IM¹¹¹. However a smaller retrospective study showed that there was in fact no significant difference in the incidence of adenocarcinoma in patients with BO that either did or did not contain IM¹¹². With this conflicting evidence base UK guidelines elect to include all metaplasia in the diagnostic criteria; regardless it is generally accepted that the presence of BO disposes individuals to a higher risk of OAC.

3.3.2 Barrett's oesophagus as a risk factor for oesophageal adenocarcinoma

The true prevalence of Barrett's oesophagus is difficult to quantify precisely, as it is often asymptomatic; as such the relative risk of OAC it confers is also difficult to quantify. The most accurate estimates of BO prevalence in a large patient populations come from a Swedish sample of 3000 adults undergoing gastroscopy and an Italian study of 1033 patients. In the first study the incidence of Barrett's oesophagus 1.6%¹¹³ and in the second it was 1.3%¹¹⁴; suggesting the presence of Barrett's in the general population is significantly more common than the incidence of OAC.

Two recent analyses have demonstrated an increased incidence of OAC in patients with BO, compared to the general population. A meta-analysis by Sikemma et al estimated the rate at 5.3 cases/1000 patient years; another metanalysis estimates this figure to be 6.5 cases/1000 patient years^{21,115}. A further study of 11,028 patients,

followed up for a median period of 5.2yrs estimates the incidence at 1.2 cases/1000 patient years¹¹⁶. Despite this lower incidence, patients in this study with BO demonstrated a relative risk of OAC of 11.3 compared to the general population.

While it should be noted that the majority of patients diagnosed with BO do not go on to develop OAC, it is generally accepted that the conversion rate is in the range of 0.1-0.3% per year^{117,118,119}, which is significantly higher than the generally population risk per year. As such significant research to understand the evolution of BO into OAC has been undertaken to establish the clinical and histopathologic progression through BO to adenocarcinoma.

3.3.3 The dysplasia – adenocarcinoma sequence in Barrett’s oesophagus

Barrett’s oesophagus is of clinical significance as it represents the only endoscopically detectable precursor lesion to OAC. As with squamous dysplasia, there is a well-established linear progression from non-dysplastic BO, through low and high grade dysplasia to adenocarcinoma. The progression of dysplastic changes within BO tissue is defined according to the Vienna classification¹²⁰. The detection of dysplasia within histologic samples taken from BO will determine future management strategies as discussed below. All three major guidelines discussed in this chapter require the determination of dysplasia to be made by two expert gastrointestinal pathologists.

In non-dysplastic BO (NDBO) tissue the squamous epithelium of the oesophagus is replaced by intestinal type columnar epithelium which contains goblet cells. The epithelial surface cells should appear more mature than the deeper cells. All cytologic features should appear normal; with smooth nuclear membranes and nucleoli, nuclear polarity is maintained^{121,122} (**figure 15**). Histologic specimens should have normal architecture and an abundance of lamina propria, there should also be the absence of any features concerning for dysplasia.

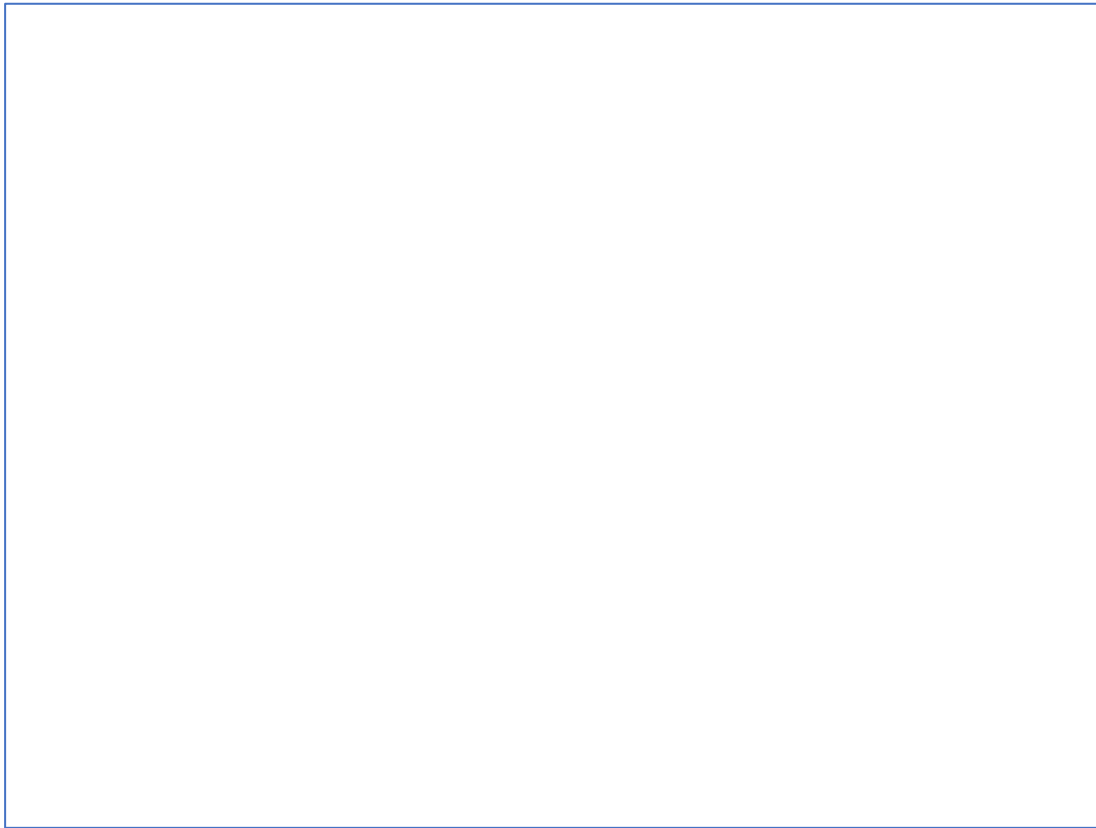


Figure 15 Histologic appearance of non-dysplastic Barrett's oesophagus (from Pathology Outlines available at ¹²²)

Low grade dysplastic (LGD) changes include nuclear enlargement and crowding along the basal surface of cells. There may be mild to moderate cytological atypia that extends to the luminal surface, including up to moderate luminal hyperchromasia and enlargement (**figure 16**). The diagnostic accuracy and high interobserver variability of pathologists describing LGD remains high and is a source of contention. A 2015 study found that in 293 BO samples with 'confirmed' LGD, only 27% had this diagnosis supported after review by two expert GI pathologists – the rest were downstaged to either non-dysplastic BO or were labelled as indefinite for dysplasia¹²³. In any case, the presence of confirmed LGD equates to between a 3-23% risk of conversion to OAC per year^{123,124}.

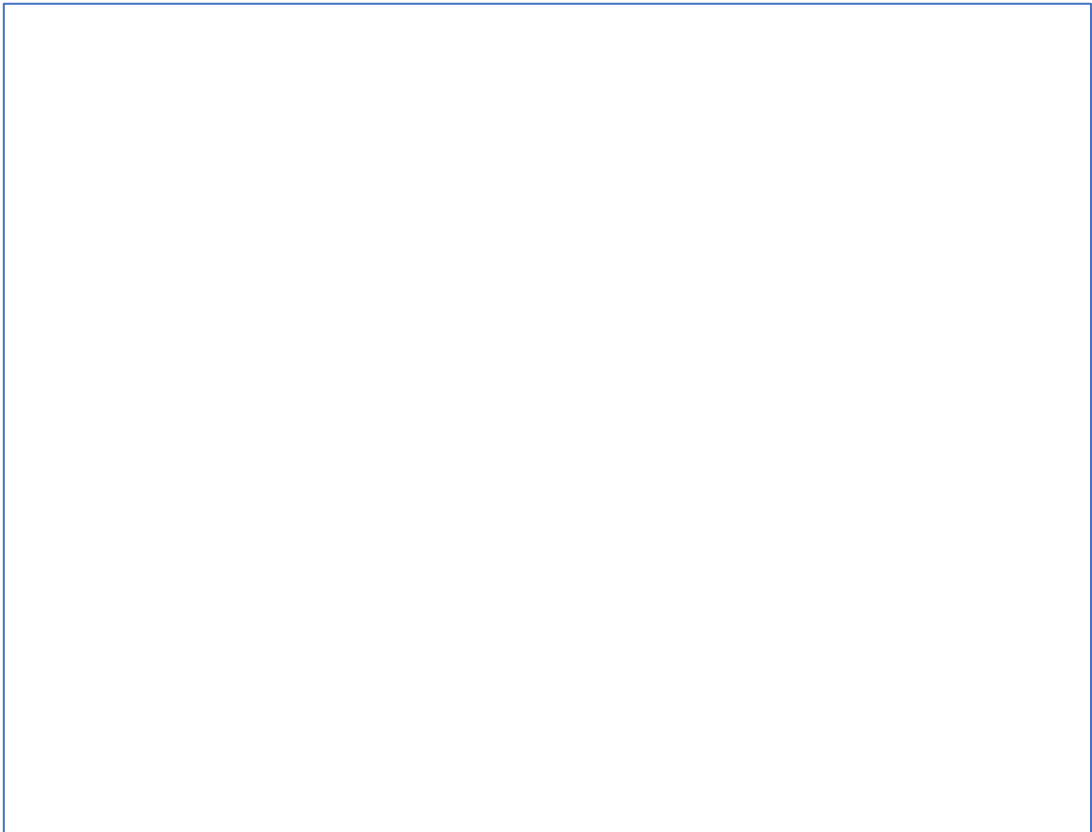


Figure 16 Histologic appearance of Barrett's oesophagus with low grade dysplasia (from Pathology Outlines available at ¹²²)

The histologic changes associated with high grade dysplasia (HGD) include more marked cytologic atypia, there may be a high degree of nuclear stratification and irregularity (**figure 17**). The nucleus to cytoplasmic ratio is often markedly increased. The degree of abnormal tissue architecture is also more prominent, with lateral budding, villus formation and the presence of dilated glands filled with necrotic material. The risk of progression from HGD to adenocarcinoma is markedly higher, with estimates ranging from 4-55% over a five year follow up period^{125, 126}. It is generally accepted that the minimum incidence of cancer per year in patients with confirmed HGD is at least 6%¹²⁷. As such it is recommended that patients with HGD should undergo prompt endoscopic eradication therapy, in order to prevent the progression to OAC (**figure 18**).

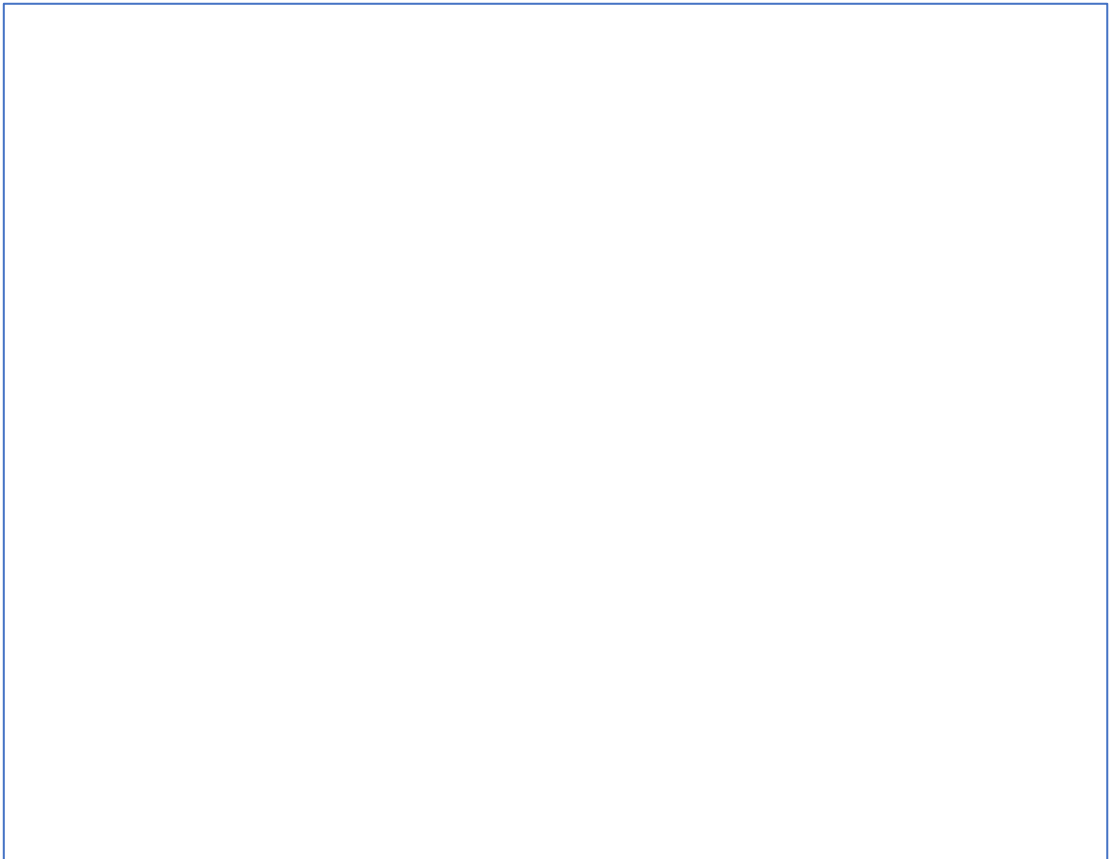


Figure 17 Histology appearance of Barrett's oesophagus with high grade dysplasia (from Pathology Outlines available at ¹²²)

Given the heightened risk of OAC in patients with known BO, they are often entered into endoscopic surveillance programs. The underlying principle of these programs is to identify early neoplasia at a stage where it can be treated with the endoscopically eradication therapies (EET) discussed in **chapter 3.6**. This spares patients from requiring more invasive surgical management, which is the only curative therapy for more advanced disease, but its efficacy on a population level remains disputed¹²⁸.

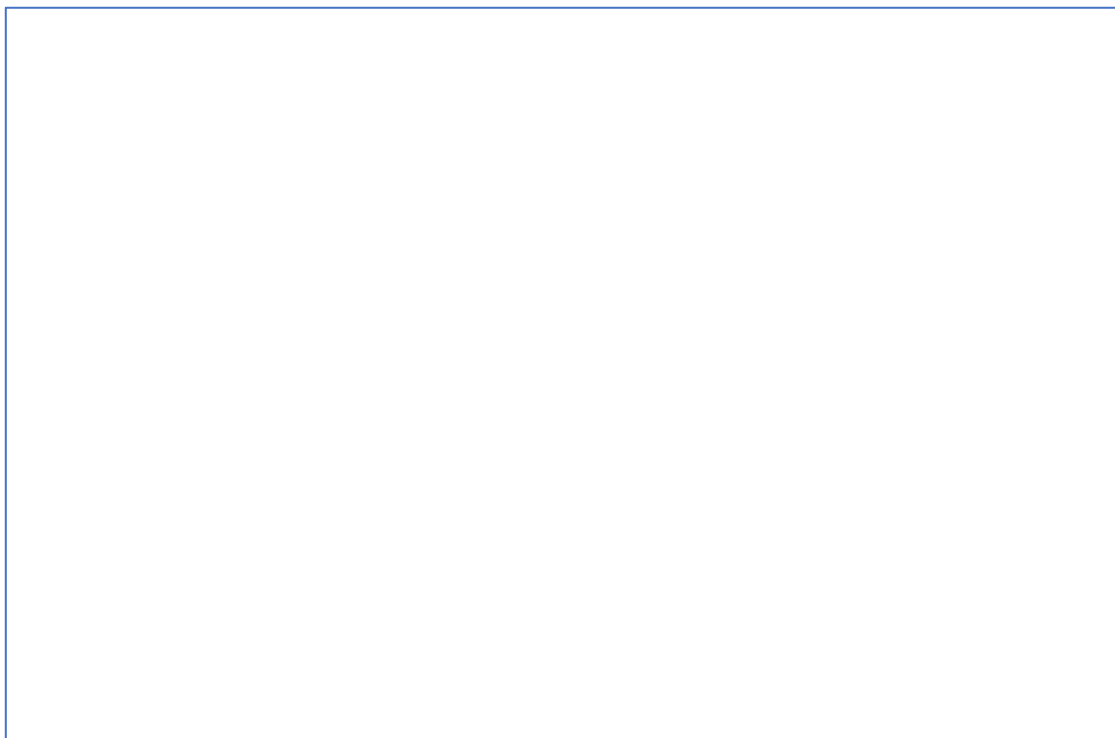


Figure 18 Histology appearance of Barrett's associated adenocarcinoma (from Pathology Outlines available at ¹²²)

3.3.4 Endoscopic surveillance of Barrett's oesophagus – the Seattle Protocol

As recommended by the National Institute for Clinical Excellence (NICE) and the BSG, the current paradigm in the UK for the surveillance of BO patients involves the interval endoscopic assessment of their BO segment in order to identify treatable early neoplastic lesions. There is currently no good evidence that screening for asymptomatic patients or those with GORD for BO is cost effective^{129,130}.

It is generally accepted that patients enrolled in BO surveillance programs, who progress to OAC are likely to have their cancer detected at an earlier stage than those outside of surveillance programs – although no randomised control trials (RCTs) exist to demonstrate this. Similarly patients within BO surveillance programs have generally been shown to have improved survival^{131–134}. There is however a lack of randomised control trial evidence for this, with most studies consisting of small sample retrospective analysis.

In the UK, treatments and interventions are generally approved by NICE at a cost of between £20-30,000 per quality adjusted life year (QALY) gained¹³⁵. By this measure,

the cost effectiveness of BO surveillance is debatable. A Dutch study suggested that 5 yearly surveillance of patients with NDBO was cost effective when paired with EMR and RFA for HGD and OAC assuming a threshold of €35,000/QALY¹³⁶. A review of 7 publications noted that in more than half of the included studies Barrett's surveillance was not cost effective and in some cases could cause harm¹³⁷. A study by the same authors suggests that in patients with NDBO the cost per QALY ratio is as high as \$60,858, but does note that this cost may be reduced in high risk patients such as those with known dysplasia¹³⁸. Inadomi et al. demonstrate a cost per QALY of \$10,440 when patients with known dysplastic BE undergo surveillance and treatment for OAC, suggesting that this likely remains a cost effective option¹³⁹.

The current gold standard for endoscopic surveillance is the Seattle protocol (SP)¹²⁷. The endoscopist begins by identifying the anatomical landmarks from which the BO segment is measured. The top of the gastric folds are identified and the length of any associated hiatus hernia is noted. The BO mucosa is washed thoroughly with a solution of water and a mucolytic, to remove adherent mucous and to permit a thorough assessment of abnormal surface patterns that may represent dysplasia.

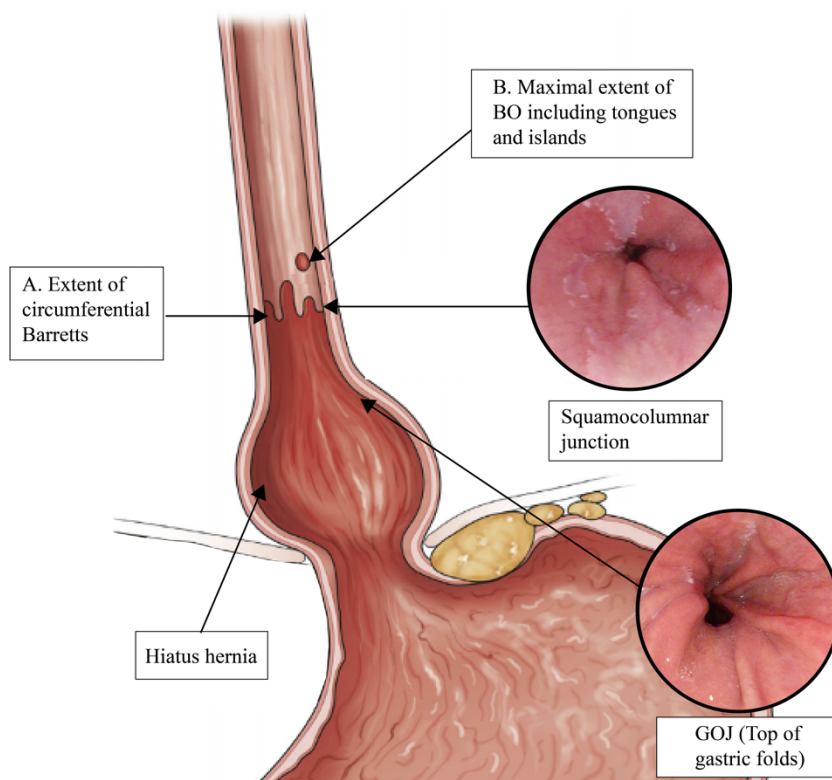


Figure 19 Schematic of how the CM Prague classification is recorded. Circumferential length (A) and maximal extent (B). Adapted from ¹⁴⁰

In order to standardise and improve communication between physicians managing patients with BO, the segment should be measured and described in accordance with the Prague CM criteria¹⁴¹. Endoscopists record the circumferential extent of the BO segment from the top of the folds in centimetres (**A in figure 19**), as well as the maximal extent of visible BO mucosa including tongues and islands (**B in figure 19**).

Current guidelines state that the endoscope should be slowly withdrawn proximally throughout the BO segment, with the endoscopist observing for visible lesions suggestive of early neoplasia. If lesions suspicious for early neoplasia are detected, clinicians should record their findings based on the Paris classification^{142,143}. This simple classification divides lesions into three broad categories; flat, depressed and protruding, with further subcategorization based on the mucosal colour and appearance of vessels (**figure 20**). The implications of this will also be discussed in a later section. Once targeted biopsies have been taken from the suspicious area – in order to secure a histological diagnosis and to plan therapy, the endoscopist proceeds

with quadrantic biopsies. These are taken from every quadrant at 2cm intervals of the BO segment, starting at the GOJ and ending at the proximal squamo-columnar junction.

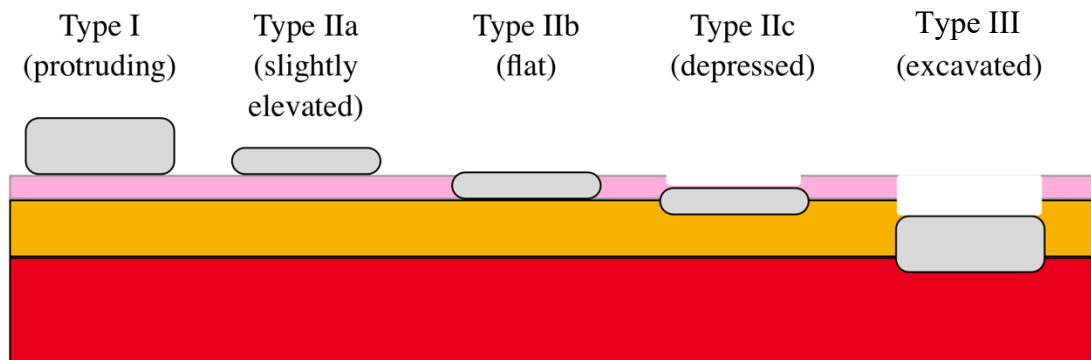


Figure 20 Schematic of the Paris classification used to describe oesophageal lesions (indicated in grey)

3.3.5 Limitations of the Seattle Protocol

Early neoplastic lesions with BO are often subtle, flat, focal and hence may be difficult to distinguish from normal tissue during endoscopic assessment (**figure 21**). There are a number of limitations that may mean the Seattle protocol, while the current gold standard, could be improved. A major theme of this thesis is to propose methods that may mitigate these deficiencies and improve cancer detection



Figure 21 A typical subtle, flat lesion seen in early Barrett's associated neoplasia

Large population based studies have demonstrated between 12.7-33.5% of OAC is diagnosed within a year of a normal index endoscopy^{116,144}, suggesting a potentially high miss-rate of early lesions. A meta-analysis by Visrodia et al, showed that in 24 cohorts of adults with NDBO at baseline, followed up for at least three years endoscopically, 25.3% of OACs were missed¹⁴⁵. Although adherence to the SP in all studies was questionable, there are clearly inherent deficiencies in OAC detection under the current methods of surveillance.

Tschanz et al. highlight the susceptibility of the protocol to sampling error. The average BO mucosal surface area is estimated at 14cm², with random forceps biopsies sampling only around 0.5cm² of this area - representing just 3.5% of the total surface area¹⁴⁶. If lesions are missed on visual inspection of the mucosal surface, there is a high chance that they may also be missed using this sampling method.

Of particular relevance in the modern endoscopy suite, where time is pressured, is that the time spent assessing the oesophagus has an impact on the detection rate of early neoplasia during BO surveillance endoscopies. Gupta et al. demonstrated that the detection rate of HGD/OAC improved with assessments where endoscopists inspected each centimetre of the BO segment for more than one minute¹⁴⁷. Studies have also demonstrated that compliance with the protocol is often poor and worsens with longer segments – with sensitivity for dysplasia/neoplasia detection ranging widely from 33->90%¹⁴⁸.

Given the aforementioned limitations of the SP, there is an unmet need to develop a system that improves the visual recognition of early BO associated neoplasia – to enable endoscopists to take more targeted biopsies and possibly improve their dysplasia detection rates. Endoscopic technology has advanced significantly in recent years, with large numbers of innovations focused on improving the quality of endoscopic imaging that is available to clinicians.

3.4 Adjuncts to improve Barrett's dysplasia detection - chromoendoscopy

In a similar manner to that seen in the endoscopic assessment for ESCN, clinicians have sought to improve visualization of early BO associated neoplasia through the use of chromoendoscopy. Subtle changes in the mucosal surface features may be associated with the presence of dysplasia in BO segments, but they are often hard to detect with HD-WLE alone. The application of topical dyes aims to improve lesion recognition; several such preparations have been studied in the oesophagus with the aim of improving surface enhancement and subsequent dysplasia detection in Barrett's oesophagus, but with conflicting results.

Both methylene blue and indigo carmine have been investigated to see if they improve the dysplasia detection rate of targeted biopsies¹⁴⁹. Recent studies have shown that there is little additional benefit to be derived from their use¹⁵⁰. Another such agent which has been extensively investigated for its use in the detection of BO associated neoplasia in high risk populations is acetic acid. Following application of a dilute solution of acetic acid, the reversible acetylation of cellular proteins occurs, giving the

mucosal tissue a characteristic white appearance¹⁵¹. The loss of this ‘aceto-whitening’ effect, where sections of the mucosa more rapidly return to their normal colour, has been identified as a sensitive and specific marker for the presence of dysplasia^{151,152,153}. While some early studies showed little additional benefit in its use¹⁵⁴, there is now more experience and its use is commonplace in centres undertaking frequent assessment of high risk Barrett’s populations. The PREDICT study provided evidence for a validated classification system for use in the identification of Barrett’s neoplasia using acetic acid. This study demonstrated sensitivities for the identification of neoplasia of 98.1% and 95.5% in endoscopists and non-endoscopists respectively¹⁵⁵. Despite the evidence in favour of using acetic acid, there are no randomised control trials comparing its use against the current gold standard of the Seattle protocol, as well as the additional procedure time and demand on resources, the routine use of chromoendoscopy is neither widespread nor recommended for routine practice by any major gastroenterological associations.

3.5 Adjuncts to improve Barrett’s dysplasia detection – virtual chromoendoscopy

The use of real-time chromoendoscopy may lengthen procedure times and consumes resources, with questionable levels of benefit derived from their use. Concurrently, imaging technologies available to endoscopists have developed rapidly over the last two decades. Significant commercial and clinical interest has seen the development of several virtual chromoendoscopy platforms, designed to increase the visibility and discrimination of early neoplastic lesions by endoscopists in a way similar to chromoendoscopic methods. These systems aim to provide high definition, digital surface enhancement of the oesophageal mucosa, as well as enhanced visualisation of mucosal microvessels. The main imaging platforms used in current practice are outlined below:

3.5.1 Narrow band imaging

The oldest and most widely studied of the advanced imaging modalities is Narrow Band Imaging (NBI, Olympus). NBI transmits predominantly in two wavelengths of light, with a lower baseline transmitted for the rest of the visible light spectrum; blue light at around 415nm and green light at 540nm¹⁵⁶. The blue spectrum penetrates the superficial layer of the mucosa and highlights the microvascular patterns. The transmission of green light penetrates deeper into the mucosa and allows visualization of deeper vessels which may be disordered or enlarged in neoplastic lesions. NBI utilizes a special filter installed within the light source in order to deliver these specific wavelengths and so is considered a pre-processing technology. A criticism of NBI technology is that it delivers less intense light across the visible light spectrum, as a result the illumination of the oesophagus may be reduced compared to HD-WLE.

Most studies into the use of advanced endoscopic imaging have assessed NBI. Taking advantage of the improved visualization of mucosal and vascular patterns, most studies attempted to classify dysplastic and non-dysplastic areas of BO based on these macroscopic features – with varying degrees of success^{157,158,159}. A meta-analysis by Mannath et al. demonstrated the clinical utility of NBI in the detection of high grade dysplasia within BO, with pooled sensitivities and specificities of 96% and 94% respectively¹⁶⁰. The Barrett's International NBI Group (BING), presented a prospective study which defined and validated a simple criteria that could predict dysplasia with up to 85% accuracy, based on mucosal and vascular patterns¹⁶¹.

3.5.2 iScan Optical Enhancement

The iScan system (Pentax) has emerged as a potential alternative to NBI. iScan employs both novel pre and post-processing technologies to provide surface enhancement of the superficial structures of the mucosa, as well as improving the visibility of the mucosal microvasculature¹⁶². Using a new optical filter, iScan optical enhancement mode (OE) delivers specific wavelengths of light which correspond with the main absorption spectrum of human haemoglobin (415nm, 540nm and 570nm)

and so, like NBI, highlight the microvasculature within the most superficial layers of mucosa. However, OE uses additional post-processing technology in order to digitally enhance mucosal surface features. The resulting high definition image allows detailed visualization of the minute vessels, but unlike NBI, also enhances mucosal microstructures¹⁶³. Because OE also raises the intensity of light transmission across the visible spectrum, it is argued that the images obtained are brighter than NBI which allows better visualization in areas of the gastrointestinal tract with larger lumens.

Currently no studies have been published on the use of OE in the detection of early Barrett's associated neoplasia. Lipman et al, using early iterations of the iScan system demonstrated accuracies of 69% without and 79% with the use of topically applied acetic acid in a validation dataset¹⁶⁴. This study also validated a simple classification system based on mucosal and vascular patterns for the detection of neoplasia on zoom magnified endoscopy. A small study by Verna et al. demonstrated a low dysplasia detection rate of between 5.5-9.2% on target biopsies, using a combination of magnification endoscopy, acetic acid and iScan enhancement¹⁶⁵. It is likely that the lower reported accuracies for iScan compared to NBI are a reflection of the technology being at a relative early stage of its development, as well as a 'learning curve' effect due to limited clinician experience.

This body of work seeks to add to the literature assessing the utility of OE for use in BO associated neoplasia detection.

3.5.3 Blue laser imaging

Another new technology is Blue Laser Imaging (BLI, Fujifilm). BLI works using two monochromatic lasers installed within the endoscope light source. The first laser transmits blue light at 410nm in order to highlight the microvascular patterns of the mucosa. A second laser, transmitting at 450nm provides high intensity white light through the excitation of white phosphor in the tip of the endoscope¹⁶⁶. Very recent work by Subramaniam et al proposed the BLINC classification. This demonstrated that BLI shows promise as an adjunct for dysplasia detection. In the assessment of 50 BLI

images, the sensitivity of dysplasia detection in expert endoscopists was 96% and in non-expert endoscopists it was 85.3% (improving to 95.7% following a training program)¹⁶⁷.

3.6 Endoscopic eradication therapy (EET) of early oesophageal neoplasia

The overarching aim of improving the endoscopic detection of HGD or early OAC, is to allow prompt access to EET, thereby sparing patients the morbidity and mortality of more invasive surgical options such as oesophagectomy. The selection of endoscopic eradication therapy depends on the histological stage and appearance of the early neoplastic lesions. This section will outline the various therapeutic options available and discuss the evidence behind their use.

3.6.1 Photodynamic therapy (PDT)

PDT was an early ablative therapy trialled for use in the eradication of BO, it involves the oral or intravenous administration of photosensitizing agents to the patient before endoscopy. Typically, 5-aminolevulinic acid administered orally or porfimer sodium intravenously would distribute throughout the body tissues and sensitize tissues to light exposure at certain wavelengths. During endoscopy, a catheter which emits the specific wavelength of light would be targeted at the BO tissue. Upon exposure of the PDT agent to the light energy, an oxygen free radical is generated which causes targeted tissue damage and subsequent necrosis of the BO tissue.

Patients treated with PDT were demonstrated in one randomised control trial to have a reduced risk of developing adenocarcinoma compared to untreated individuals with BO (13% vs 28%)¹⁶⁸. In another study patients with HGD treated with PDT alone had similar 5 year mortality to those treated with oesophagectomy (9% vs 8.5%) – but without the considerable morbidity associated with such a major procedure¹⁶⁹. Despite early promise shown by PDT, its expense, high stricture rate, as well as the development of other high-quality treatments such as radiofrequency ablation limited its widespread use.

3.6.2 Radiofrequency ablation (RFA)

RFA is currently the most widely used and recommended ablative therapy for dysplastic BE. There is also evidence that it is effective for the treatment of dysplastic squamous lesions. Numerous devices are commercially available for RFA therapy and include circumferential balloon electrodes, focal devices that attach to the endoscope tip or electrode paddles that can be passed via a working channel.

In preparation for RFA, the oesophageal mucosa is washed with a dilute solution of n-acetylcysteine to remove any adherent mucus. The delivery device is then positioned adjacent to the oesophageal mucosa and ablations deliver heat energy into the superficial layers. The heat energy delivered to the epithelium causes the thermal necrosis, thereby destroying early neoplastic lesions within the superficial mucosal. By avoiding thermal injury to the muscularis mucosae and deeper structures the risk of fibrotic stricture formation is thought to be reduced. Multiple treatment doses can be delivered at 2-3 month intervals until histologic or endoscopic remission is achieved.

Several small studies have demonstrated that RFA is an effective therapy for the eradication of low and high grade dysplasia. Bergman et al (2010) reported a 97% histologic resolution of LGIN/HGIN lesions in a cohort of 29 patients treated with RFA. A larger prospective study by He et al (2015) showed that complete histologic remission was achieved in 84% of patients treated with RFA at 3 monthly intervals⁷⁵.

The use of RFA in the treatment of BE associated dysplasia has a more robust evidence base. In a randomised control trial undertaken by Shaheen et al (2009), patients with confirmed dysplastic BE were assigned to either receive RFA, or a sham control procedure. 81% of patients with HGD achieved remission of dysplasia compared to only 19% in the control group (which in itself suggests sampling error rather than true histologic remission). Furthermore, higher progression rates of dysplasia were observed in the control group compared to the treatment group (19% vs 2.4%)¹⁷⁰. Other studies have also confirmed favourable outcomes and durability of remission of intestinal metaplasia and dysplasia following RFA treatment^{126,171}.

3.6.3 Cryoablation

Cryoablation is a promising new ablative therapy that involves the topical application of liquid nitrogen spray to the oesophageal mucosa. The deep-freezing effect, followed by thawing is thought to cause disruption of the cellular architecture, as well as superficial mucosal necrosis in response to the induced vascular ischaemia. Early studies into the use of cryoablation as a BE eradication therapy, as well as new devices to improve its uniform delivery to the oesophageal mucosa, demonstrate promising results^{172,173}.

3.6.4 Endoscopic mucosal resection

Endoscopic mucosal resection (EMR) is the most studied EET for use in BO associated neoplasia. Visible lesions containing either HGD or suspected intramucosal OAC are amenable to EMR. The choice of method for EMR varies between operators and centres, but techniques include ligation assisted, cap assisted or cap-snare assisted resections. The aim of EMR is to selectively resect dysplasia confined to the mucosal layer of the oesophagus, whilst leaving healthy tissue intact.

EMR has been widely demonstrated as a safe technique for the resection of early neoplasia; demonstrating high cure rates of between 87-96% for neoplasia^{174,175–177}. While a safe modality of treatment the complications of EMR are serious and include stricture formation, incomplete resections and perforation. One study indicates that within a structured training program perforation rates of endoscopists performing their first 20 EMRs were as high as 5%¹⁷⁸. In experienced hands the perforation rate of EMR is significantly reduced, with most studies reporting rates of around 0.5%^{72,179,180}.

Recurrence rates after EMR monotherapy are estimated at around 6-7%¹⁸¹. Risk factors for recurrence include piecemeal resections, long segment BO and where clinicians do not follow EMR with ablative therapies¹⁸².

A systematic review of 8 cohort studies identified promising results using EMR. Complete eradication of early neoplasia was observed in 97% of patients, with complete resolution in metaplasia seen in 85%. The stricture rates of 37% in these cohorts as well as significant heterogeneity between included studies meant that the authors could not definitively recommend EMR as a sole treatment for BO neoplasia¹⁸³.

3.6.5 Endoscopic submucosal dissection

One of the main drawbacks of EMR is that for larger mucosal lesions, piecemeal resection is usually required, thereby increasing the recurrence risk. Endoscopic submucosal dissection (ESD) was pioneered in Japan in the late 1980s for the endoscopic resection of early gastric cancers. An advantage of ESD over EMR is that it permits en bloc resection of lesions extending into the submucosa and so allows histologic assessment of the lateral and deep margins. Much debate on the utility of ESD compared to EMR has continued in recent years.

A meta-analysis of 11 studies on the safety and efficacy of ESD, suggests it is potentially an effective form of EET for early BO associated neoplasia¹⁸⁴. This analysis reported overall en bloc, R⁰ and curative resection rates of 93%, 74.5% and 64.9% respectively. The risk of bleeding and perforation were relatively low at 1.7% and 1.5%. The estimated recurrence rate after curative ESD was 0.7% at 22.9 months; longer term follow up on the durability of ESD are needed.

Studies comparing ESD and EMR in the resection of BO associated neoplasia are limited largely because Asian centres, where experience with ESD is greater, typically see more SCC than OAC. A meta-analysis¹⁸⁵ including 1080 patients suggests that ESD is more likely to allow en-bloc resections than EMR (97.1% vs 49.3%). Notably ESD achieved higher cure rates (92.3 vs 52.7%), a lower incidence of local recurrence (0.3% vs 11.5%) and similar rates of complications; except for perforations where ESD was riskier. Another meta-analysis demonstrated a non-significant trend towards higher

recurrence in EMR (2.6%) compared to ESD (0.7%)¹⁸⁶. Again similar complications rates were recorded in both groups, but EMR was associated with faster mean procedure times. More work on the comparative efficacy of EMR and ESD is required.

3.6.6 UK guidelines for the assessment and endoscopic eradication therapy of Barrett's neoplasia

The current guidelines on Barrett's surveillance and progression to therapy are defined by the BSG. Patients found to have uncomplicated, NDBO should undergo clinical review and consultation to assess their fitness to enter a surveillance program. Patient with a segment length <3cm with IM, who are deemed suitable, undergo surveillance every 3-5 years. Patients with a segment length >3cm undergo OGD every 2-3 years¹⁸⁷.

Patients with biopsies demonstrating BO that are indefinite for dysplasia or LGD for the first time require six months of maximal acid suppression followed by a repeat OGD. Those who are upstaged to dysplasia or who have a second finding of LGD, require discussion in a multidisciplinary team (MDT) meeting regarding endoscopic eradication therapy.

Patients with confirmed HGD or confirmed intramucosal cancer (IMC) should undergo EET promptly, delivered at a centre with experience in these techniques. NICE recommends that for patients with a localised visible lesion consistent with HGD or IMC, EMR should be offered as a first line alternative to oesophagectomy. Residual or recurrent disease can be retreated with further EMR, or if it is non-visible, ablative therapy such as RFA or PDT¹⁸⁸.

Chapter 4 - Artificial intelligence in the diagnosis of oesophageal cancer

4.1 Defining artificial intelligence

The use of artificial intelligence (AI) in medicine is a burgeoning field. Across numerous disciplines the role of AI in diagnosis and clinical decision making is being demonstrated. AI is a broad and highly technical field, hence understanding it, unfamiliar terminology and applications can be confusing and prone to error when described by clinicians. This chapter serves to provide an overview of relevant concepts in artificial intelligence, as well as to appraise its role to date in the endoscopic diagnosis of early neoplasia.

AI is an umbrella term used to define the development of computer systems that require or mimic human-like forms of intelligence: natural language processing, spatial reasoning and computer vision are examples. Machine learning is often used interchangeably with AI but actually relates more specifically to the construction of mathematical models designed to make predictions regarding unseen data, based on observed patterns in known data.

Machine learning can be described as supervised and unsupervised. In this body of work we employ a form of supervised machine learning to enable a neural network to identify unseen images as cancerous based on previously seen images of cancerous tissue. In this case supervised indicates that the data used to train a network is labelled by a human against a gold standard; normal or neoplastic in the case of this thesis. Based on the model developed when trained with this input, the network then attempts to predict the labels on an unseen selection of data. Unsupervised learning does not require labelled data, instead a network attempts to independently cluster or identify patterns within a dataset, in order to identify the same patterns within an unseen dataset.

Deep learning is a sub-field of machine learning that utilises neural networks (NNs) and is the field of AI to which this body of work pertains most closely. NNs bear some resemblance to the mammalian brain, in that 'neurones' within the network are

connected to downstream neurones in layers. Within the field of computer vision and image processing, convolutional neural networks are a frequently used machine learning algorithm.

4.2 Convolutional neural networks (CNNs)

Convolutional neural networks are a relatively new addition to the world of computer vision, rising to prominence over the last decade. CNNs were inspired by the seminal work of Hubel and Weisel who demonstrated that discrete clusters of neurones within the feline visual cortex were specialised to recognise specific features present within the visual field¹⁸⁹ – edges of objects arranged at a specific orientation will be seen by certain clusters of neurones for example. A commonly used application of CNNs is for image processing and classification. Humans from an early age are able to look at images and subconsciously label what they are seeing, a cat compared to a dog for example. CNNs when appropriately trained are potentially able to achieve the same thing.

CNNs are complex and the technical aspects of their construction vary depending on their application; a full overview of their architecture and application is beyond the scope of this thesis. The following sections aim to provide an overview to better understand the key components of CNN design, training and testing as it relates to the area of computer aided endoscopic diagnosis. The core concepts relating to CNN design are framed within the following sections using the later studies discussed in this thesis. The aim is to provide clinicians undertaking similar work in the future to gain enough of an understanding of the principles of machine learning to be able to work more closely with those of a computer science background

4.2.1 Key terminology relating to CNN design

This section aims to define some of the key terms that relate to CNN development and this thesis, such that the descriptions outlined in subsequent sections are better understood.

Machine learning

Machine learning is a branch of artificial intelligence and computer science that aims to build algorithms, which imitate the way in which humans learn, in order to improve accuracy for a given task over time.

Convolutional neural networks

A CNN is a type of machine learning algorithm which is designed to process arrays of data. Such arrays include images hence the utility of CNNs in image classification and computer vision. The typical structure of a CNN involves an input layer, a series of hidden or convolutional layers and an output layer

Supervised learning

Supervised learning refers to machine learning where labelled data is used against which a neural network trained to classify. The CNN outputs are then compared to the ground truth of this labelled data and the weights of the network are adjusted accordingly until the predictions become more accurate. The studies presented in this thesis all relate to supervised learning.

Input layer

The input layer refers to the data that is being provided to the CNN. In the case of our studies it refers to the layer in which the numerical pixel matrix that corresponds with an endoscopic image is provided to the CNN.

Hidden layers

Hidden layers are found between the input and output layers of the CNN. It is in these layers that input data is sequentially convolved using kernels and directs them through an activation function prior to the output layer.

Output layer

The output layer of the CNN is the last layer of the network in which the data generated from the preceding hidden layers is converted into a class prediction.

Convolutions

A convolution is a linear operation in which an array or matrix of weights are multiplied by another array of input values. In the context of CNN design this array of weights is a kernel or filter and the convolutions occur in a stepwise manner with the filter moving across the image.

Weights

A weight is a parameter or number which transforms input data within a networks hidden layers. Weights arranged within a matrix are known as kernels.

Kernels

A kernel is also known as a filter. It is a matrix of weight values which can be applied to input data in order to multiplied it. Kernels are applied to input images in a sequential manner during convolutions.

Folds

A fold is a set of data taken from the overall dataset. For example in a study including data from one hundred patients, ten folds of ten patients or four folds of twenty five patients could be generated. Folds can then be used in k-fold cross validation where each fold is further subdivided into a training set of data, validation set and testing set of data.

Loss

Loss functions can be used to compute the margin of error between the networks output predictions and the known 'ground truth' value. The loss refers to how incorrect the network predictions are compared to the correct value. The aim of a CNN is to reduce the loss over a number of iterations; meaning that it's predictions become more accurate.

Backpropagation

Backpropagation is an algorithm that enables the calculation of the loss function gradient for each weight in the network, working backwards from the output layer through the convolutional layers. In this way individual weights can be automatically adjusted to reduce the loss function and improve the network's accuracy over many iterations.

4.2.2 Using images as input data for CNNs and basic CNN architecture

Images, however complex, can be thought of as a series of pixels. For colour images each pixel is represented by three numbers, with each number denoting the intensity of the red, green and blue (RGB) components of the overall pixel. The intensity values of the RGB components of 8 bit images ranges from 0 (no intensity) to 255 (maximum intensity). For example, a pure black pixel's RGB components could be written as 0,0,0; a pure white pixel would be 255, 255, 255. These pixels, when considered as numerical expressions of RGB components and combined with potentially large numbers of other pixels that make up an image, it becomes clear that an image can be thought of as a long list of numbers. As this string of numbers, images can therefore be used as input data for training a CNN.

The basic structure of a CNN as related to the work in this thesis is to take an input image and produce an output classification that through an iterative process of 'learning' becomes gradually more accurate at making correct classifications. An unknown function is used to take the numerical input data and to associate it correctly

with an output classification. A function is a mathematical operation that associates a set of data X , with a single domain within a set of data Y . In the case of this body of work, the function is required to associate the numerical values contained within the pixels of magnification endoscopy images, with an output classification (normal vs dysplastic).

Clearly a function capable of associating the huge volume of numerical data within a high-definition image, with a binary outcome would likely be highly complex and beyond the calculation of human operators. Using the rapid processing capabilities of computers, the 'trial and error' process of identifying the correct function can be achieved in a shorter time - the function of a CNN within this context is to operate as a function approximator.

A CNN can operate as a 'function approximator'. It is assumed that the unknown function associating the image data with a class output is a series of convolutions applied one after the other. This function forms the basis of the CNN structure. To achieve this, the image, represented by a string of numbers arranged in a 3D matrix to account for the RGB intensities for each pixel, are provided to the CNN. At the initiation of the network's training, a number of weights are randomly selected. These weights are numerical values arranged within a matrix, these smaller subgroups of weights are known as convolutional kernels.

Convolutional kernels are then applied to the image – in the case of our studies the majority of the kernels were a 3x3 matrix of numbers. This matrix is applied sequentially to an area of the image, the numbers within the image pixels are multiplied by the numbers within the kernel to generate another series of numbers. The kernel then moves, or convolves, to the right until the end of the image is reached. The kernel convolves left to right along the next line of the image down. Once the kernel has convolved across the whole input image, the output is another matrix of numbers or 'image' which correspond with the products of the input image pixel intensities, multiplied by the weights within the convolutional kernel.

The CNN is arranged such that there are a series of convolutional layers which work in the same way – with a randomly initialised set of weights/kernels applied to the product image of each layer. Since the weights within the kernels are randomly initiated some of the output numbers can be negative values. An activation function is therefore applied between convolutional layers. The activation function for the purpose of our study was the Rectified Linear Unit (ReLU). ReLU functions by activating only neurons where the linear transformation is a positive value, any which generate a negative value are converted to 0. The repetition of these stepwise convolutions per image, application of the ReLU activation function, and generation of a new image to input into the next convolutional layer is known as forward propagation.

Kernels are therefore able to capture correlations. Weights within the kernels (analogous to image features) when passed over similar looking regions in the image (also made up of numbers) produce an output number or sum that will be a high number. When kernels pass over images areas that are not similar to it, the numbers which make up the output images will likely be lower.

During forward propagation, between each 'layer', images are downsampled (made smaller) by a combination of pooling, stride and kernel size. In our network for example, input images underwent the following convolutions with the following kernel sizes to produce gradually smaller resolution images.

- 32 convolutions/kernels (3x3x3) + ReLU – produces 32 images (128x128)
- 64 convolutions/kernels (3x3x32) + ReLU – produces 64 images (64x64)
- 128 convolutions/kernels (3x3x64) + ReLU – produces 128 images (32x32)
- 256 convolutions/kernels (3x3x128) + ReLU – produces 256 images (16x16)
- 512 convolutions/kernels (3x3x256) + ReLU – produces 512 images (8x8)

At the output layer of the neural network another activation function is applied to the output of the final convolution. This generates a single output number that is effectively a probability that the input image belongs to a class – in the case of our studies normal or dysplastic. Since our neural network is supervised, the output class

prediction of the network can be compared to the ground truth data (the known histology results) to establish whether the network prediction was correct.

This introduces the concept of loss. When the network makes a prediction on the probability that an input image belongs to a certain class, how 'correct' that prediction was can be determined against the known ground truth. Loss refers to how close to the correct prediction the network was; where the networks prediction is close to the true value the loss is low, where it is far from the true value the loss is high. High loss indicates that the weights used during the convolutional layers of the network are poor at capturing correlations with the image or features that are associated with a particular class output. The process of training a network so that it 'learns' the correct weights that can accurately predict a class output, are strongly related to reducing this loss.

4.2.3 How a neural network learns

Since a CNN operates as a 'function approximator', identifying the weights that allow a correct class prediction at the output layer is essential. To do this the weights which are randomly initialised at the start of network training are adjusted in an iterative process that is aimed at gradually reducing loss. Since the network training is supervised, there is an expected output class for each piece of data provided to the CNN. In the case of this work the input images are labelled as either dysplastic or non-dysplastic, hence the networks predictions can be identified as correct or incorrect, and more importantly *how* incorrect they were. When the correct weights are identified to correctly associate input data with output classifications, the loss will be low and the neural network should be able to accurately classify.

Adjusting the various weights which enable the CNN to produce accurate class predictions would be far too laborious to do by hand, since there are hundreds of thousands if not millions of weights to be adjusted through each of the many iterations of the network. Various methods of automating this weight adjustment process exist – for the purposes of this body of work backpropagation using gradient

descent will be discussed. Through backpropagation the weights within kernels applied at each layer are adjusted, working backwards through the network, with the overall aim of reducing the loss using the weights selected for the next iteration.

In theory if the loss could be calculated for all possible values of all the weights used in a CNN, the resulting graph (or loss curve) would be a parabola with very low loss values at a given weight and higher loss values at values above and below this particular weight value. The weight at which the loss is lowest is known as the local minimum – therefore the aim of training the neural network is to get all of the weights as close to the local minimum for loss. Gradient descent allows the network to do this.

Take the randomly initialised weights as a starting point on the loss curve, the gradient descent algorithm then calculates the gradient of the loss curve at the starting point. In calculating this gradient it is possible for it to compute in which direction weights must be adjusted in order to descend the gradient towards values which allow the network to reach its local minimum loss value. These weights will be adjusted after each iteration in a way which gradually decreases the loss value. Backpropagation refers to the manner in which the network works backwards during this process, taking the loss value for a given set of weights and adjusting them at each layer moving backwards to reduce the loss on the next iteration.

Since the process is iterative, the adjustments to the weight values are done in a stepwise manner. The magnitude of these weight adjustments is influenced by the learning rate - a hyperparameter which can be adjusted during the training of the CNN. At learning rates which are too small the adjustments to the weights are very small – this can lead to inefficient training of the network as it takes too many iterations to reach the local minimum for loss. In other words the network descends the loss curve gradient too slowly. Conversely learning rates that are too high will result in learning that is too rapid and the values for weights can overshoot the local minimum and result in increasing loss values again. A full discussion of how the ideal learning rate and other hyperparameters are selected are beyond the scope of this work – however

the above description should provide a basic overview for clinicians to understand the process if working alongside engineers or computer scientists.

The iterative process of CNN training ends when the weights are adjusted until they reach the lowest possible loss values; a process known as convergence. Once the network has converged it can be tested against unseen data in order for its accuracy to be characterised.

This body of work proposes that the HD images acquired at endoscopy, could provide a potential source of input data for the development of a CNN. We speculate that the additional contrast and clear border demarcation afforded to IPCLs by ME-NBI may also facilitate the feature recognition required of a CNN that can successfully classify these images as neoplastic or non-neoplastic. Additionally the simple and intuitive JES IPCL classification, with its correlation with histology provides a useful label for input images used to train a potential CNN. In **chapter 7**, we aim to provide proof of concept, then develop and validate an algorithm that can fulfil this purpose.

4.3 Previous work on the role of neural networks in early squamous neoplastic lesions

The body of literature focusing specifically on the role of machine learning in the detection or classification of oesophageal squamous neoplasia is sparse. What data there is demonstrates promising but limited results¹⁹⁰.

Zhang et al¹⁹¹ employed a convolutional neural network to classify images of ESCN based on their IPCL patterns. A single expert endoscopist delineated lesions on a mixed pool of 218 images of dysplastic and non-dysplastic squamous mucosa – pixels within the delineation margin were termed ‘regions of interest’ (ROI). Each image was then randomly sampled to generate 1000 patches of 50x50 pixels with patches from within the ROI deemed dysplastic and those outside non-dysplastic. Patches were used to train a CNN, which was then tested against a panel of unseen patches. Zhang et al reported reasonable sensitivity, specificity, NPV and PPV of 73.4%, 83.5%, 72.1% and 84.4% respectively, indicating the use of CNNs to classify ESCNs is feasible. A

potential criticism of this method includes that the network again used a limited number of images of ESCN lesions and therefore may be unable to generalise to classify images obtained across the spectrum of disease. The CNN reported in this study was also unable to classify in real time and so its utility for in vivo classification is limited. Again, only a single endoscopist was used to delineate ROIs, meaning that the network is biased towards this clinician's interpretation of the images. Areas outside of the ROI that did in fact contain ESCN tissue or vice versa may also have impaired the way in which the CNN classifies, particularly given the small numbers of images used where any incorrect classification would have a major effect.

More recent work by Horie et al¹⁹² developed a CNN for the macroscopic detection of SCC on endoscopy, using 8428 still images acquired from 384 patients assessed at a referral centre in Japan. Training images included 397 SCC lesions and 32 OAC lesions. The CNN was then assessed for its diagnostic performance using 1118 testing images from 47 independent patients. Within the testing set 50 patients with no neoplasia were including, increasing the prevalence of neoplasia in the testing set to 50%, compared to the training cohort where it represented nearer 5%. This study reported an accuracy and sensitivity of 98%¹⁹². This promising work recorded a box around areas suspicious for neoplasia and so further work is required to make it clinically interpretable or suitable to guide resection therapy.

Zhao et al¹⁹³ assessed the role of a CNN for the prediction of lesion histology based on IPCL patterns, they then compared the CNN with the diagnostic performance of a group of expert and non-expert endoscopists based in a single high volume Chinese centre. 1350 images were acquired from 219 patients for this study, however there was a significant imbalance in the what lesions were included, which would have biased the network performance (206 type A, 945 type B1 and 199 as type B2). Lesions containing B3 IPCLs were excluded which is another significant limitation of the study. For type A, B1 and B2 IPCL patterns their CNN demonstrated accuracies of 71.5%, 91.1% and 83.0% respectively; equating to a pooled diagnostic accuracy of 87%. This was lower than the pooled accuracy of the expert group (92%), but exceeded that of the non-expert group (73.3%)¹⁹³. This work indicates that CNNs for this purpose may

have a role in improving the diagnostic accuracy of non-expert endoscopists, particularly in a low volume western setting.

Study overview and aims

The following chapters aim to evaluate adjuncts to improve the endoscopic diagnosis of early oesophageal neoplasia. The first study will assess the role of a novel biomarker in the detection of early oesophageal neoplasia and Barrett's dysplasia. The following chapter will outline studies assessing the utility of image enhanced and magnification endoscopy in the detection of early oesophageal neoplasia and how this may be paired with artificial intelligence systems to further improve diagnosis

Study 1 - Minichromosomal maintenance component complex 5 (MCM5) as a marker of Barrett's oesophagus related neoplasia – a feasibility study

1. To describe a method for quantification of MCM5 expression in exfoliated oesophageal epithelial cells obtained from aspirated gastric fluid during endoscopic assessment of the upper gastrointestinal tract.
2. To quantify MCM5 expression in patient cohorts with histologically distinct oesophageal mucosae. To assess whether there is an incremental increase in expression between patients with a normal oesophageal squamous mucosa, those with non-dysplastic Barrett's oesophagus, dysplastic Barrett's oesophagus and those with oesophageal adenocarcinoma.
3. To establish the feasibility of quantifying MCM5 expression, using a proprietary assay (Arquer), as a means of identifying the presence of Barrett's oesophagus associated dysplasia or adenocarcinoma.

Study 2 – Assessing whether virtual chromoendoscopy using iScan Optical Enhancement improves the detection of Barrett's oesophagus associated neoplasia in expert and non expert endoscopists.

1. To assess the diagnostic performance of iScan OE compared to HD-WLE in the endoscopic detection of Barrett's oesophagus associated neoplasia in a cohort of expert endoscopists.

2. To assess the diagnostic performance of iScan OE compared to HD-WLE in the endoscopic detection of Barrett's oesophagus associated neoplasia in a cohort of trainee endoscopists.
3. To validate a previously published¹⁶⁴ consensus driven magnification endoscopy classification system for use with OE and to compare the diagnostic performance for the recognition of Barrett's associated neoplasia in a cohort of expert endoscopists using both HD-WLE and OE.

Study 3– Developing a clinically interpretable convolutional neural network to aid in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof of concept study.

1. To develop a novel AI system that can classify IPCL patterns as normal or abnormal in **endoscopically resectable lesions** (<SM1 invasion) in real time on videos acquired during magnification endoscopy
2. To develop a methodology that can be used to develop further AI systems, for use in vivo, that accurately predict the grade of a lesions histology and invasion depth based on IPCL patterns

Study 4 – Validating a clinically interpretable convolutional neural network for the prediction of early squamous cell neoplasia of the oesophagus; comparing diagnostic performance with a panel of expert European and Asian endoscopists

1. To train a convolutional neural network capable of classifying intrapapillary capillary loop patterns as normal or neoplastic based on the JES IPCL classification¹⁹⁴.
2. To expand the patient numbers and image count in our dataset significantly in order to produce a neural network better able to

generalise its predictions on a novel patient population. In addition to images of normal oesophageal mucosa (type A IPCLs), this study's dataset was now to include images of lesions which were both endoscopically resectable (type B1 and B2 IPCLs) and non-endoscopically resectable (type B3 IPCLs).

3. To quantify the diagnostic performance of a cohort of expert Asian endoscopists against that of a cohort of expert European endoscopists using the JES classification; a comparison currently unreported in the literature. The diagnostic performance of these two cohorts could then be compared to the diagnostic performance of our neural network.

Chapter 5– Minichromosomal maintenance component complex 5 (MCM5) as a marker of Barrett’s oesophagus related neoplasia – a feasibility study

The work presented in this chapter formed the basis of a peer reviewed publication. Text and figures were adapted for publication. Citation:

Everson, M., Magee, C., Alzoubaidi, D., Brogden, S., Graham, D., Lovat, L.B., Novelli, M. and Haidry, R., 2019. Minichromosomal Maintenance Component Complex 5 (MCM5) as a Marker of Barrett’s Esophagus-Related Neoplasia: A Feasibility Study. *Digestive diseases and sciences*, 64(10), pp.2815-2822.

5.1 Introduction

Access to upper gastrointestinal endoscopy is now commonplace in the developed world; patients are typically referred to services with either symptoms suggestive of upper GI pathology, or as part of a surveillance program for Barrett’s oesophagus (BE). In the UK set criteria are mandated by the National Institute for Clinical Excellence that should prompt a referral for endoscopic assessment¹⁹⁵. One of the major pathologies to be excluded at upper GI endoscopy in such these patient cohorts is oesophageal cancer.

As outlined above, the first cohort of patients relevant to this chapter are those reporting dysphagia, anaemia, new or refractory dyspeptic symptoms who are referred for screening gastroscopy to rule out oesophageal pathology. In the UK the demand on endoscopy services is increasing and access to urgent endoscopy is subject to ever expanding waiting lists; only 55% of patients referred for urgent endoscopy in 2017 were seen within a two week target¹⁹⁶. Although national data for individual cancers are not formally reported, the endoscopic oesophageal cancer detection rate for patients with alarm symptoms is low at around 3.8%¹⁹⁷. Similarly the detection of serious benign pathology was only 12.8% in the same series; suggesting that a large number of endoscopies requested under fast-track rules are not necessary¹⁹⁷. Furthermore, evidence suggests that screening gastroscopies are often falsely reassuring and that cancers are frequently missed. Menon et al. suggested in a meta-analysis that up to 11.3% of upper GI cancer may go undetected despite the patient undergoing an upper gastrointestinal endoscopy within the 3 years preceding

diagnosis^{198,199}. The constituent studies of this analysis demonstrate that in 75% of the cases of missed cancers the root cause was endoscopist error; ranging from lesions going undetected, lesions detected but not biopsied or false negatives due to insufficient biopsies being taken from lesions.

In light of the above, there is evidence that the detection rate of oesophageal cancer in patients presenting with alarm features is low and that even in patients with a cancer present there is a propensity for the lesion to be missed. There is therefore a need to find alternatives to endoscopy in patients with alarm symptoms that may facilitate an earlier, more accurate and less invasive way to diagnose early cancers.

The second group of patients who routinely present to upper GI endoscopy services are those with Barrett's oesophagus undergoing interval surveillance endoscopy. Since early Barrett's oesophagus associated neoplasia has low rates of distant spread prompt detection is vital to allow access to endoscopic eradication therapy (EET). Endoscopic mucosal resection (EMR) and radiofrequency ablation (RFA) now afford patients with early BE associated neoplastic lesions high rates of cure, particularly if lesions are confined to the mucosa^{127,170,171,200}. As previously discussed, the Seattle protocol employed for Barrett's surveillance is the gold standard of care, but may be suboptimal, with the sensitivity of neoplasia detection estimated between 33-90²⁰¹. Likely as a result of its inherent deficiencies, including a low percentage sampling of the Barrett's mucosa with quadrantic biopsies, it is estimated that up to 25.3% (95%CI:16.4-36.8%) of BE associated adenocarcinoma may be missed during an endoscopic assessment in the preceding year¹⁴⁵.

Barrett's surveillance is both labour and cost intensive; there is mixed evidence regarding the cost effectiveness of Barrett's surveillance programs²⁰², or indeed whether enrolment in such a program improves overall survival^{203,204}. Given the number of biopsies required to be taken at intervals of between two and five years, there is significant time and financial expenditure. With an incidence rate of only 0.2-2% in one study¹²⁵ and one cancer diagnosis for every 1/52²⁰⁵ to 1/285²⁰² patient years under surveillance reported in others, the random biopsies mandated in the

Seattle protocol may place an excessive strain on resources, particularly in health systems with constraints.

Given the inherent deficiencies in endoscopy for the detection of cancer outlined above, there is a need to explore adjuncts that may improve the detection of neoplasia by reducing the reliance prolonged endoscopies and biopsies. Biomarkers obtained from patients under investigation for oesophageal cancer may provide an adjunct to the endoscopist, by offering objective evidence on whether or not neoplasia is present. We envisage that a validated biomarker could serve two purposes to mitigate the shortcomings outline above; firstly if acquired through tissue samples intra-endoscopy, a positive value could alert physicians to the presence of a cancer that has been missed, thereby prompting a second interval endoscopy or a reassessment of any lesions identified. Secondly a validated biomarker could be used to 'rule out' the presence of neoplasia; thereby preventing patients undergoing further unnecessary endoscopic assessment, biopsies and so reducing the cost and time burden on endoscopy or pathology services. One such biomarker which we have selected to investigate in this study is minichromosomal maintenance complex component 5 (MCM5).

Minichromosomal maintenance complex component 5 (MCM5) is a cell cycle protein which forms part of the DNA replicative helicase²⁰⁶. MCM proteins are upregulated during the transition from G₀ to G₁ of the eukaryotic cell cycle and so are believed to be implicated in DNA replication and cell cycle regulation^{206–208}. The dysregulation of MCM5 expression has been previously associated with the development of cancer in epithelial tissues – including cervical²⁰⁹ and urothelial cancers^{210,211}.

Epithelial cells surrounding the lumen of hollow viscus shed cells intraluminally, allowing them to be harvested and used for laboratory analysis of MCM5 expression. The oesophagus, in common with the rest of the alimentary tract is lined with epithelial cells, which shed into the lumen and are found mixed with the normal transitory contents of the gut. Going et al have previously investigated the

expression of MCM5 proteins in immunostained, formalin-fixed histological specimens taken from Barrett's oesophagus and squamous mucosa. They demonstrated that the failure of MCM5 expression to downregulate, a feature seen in non-dysplastic tissue, was observed in cells classified as dysplastic²¹². This phenomenon was observed in both glandular Barrett's associated dysplasia, and in squamous dysplasia²¹². Importantly this study also demonstrated that cells with raised MCM5 expression were present up to the luminal surface of the oesophageal mucosa²¹², suggesting that they could be exfoliated into the alimentary tract and collected endoscopically through fluid aspiration.

The epithelial cells exfoliated into the oesophagus and gastric fluid may provide a useful target for quantifying the expression of MCM5 in the Barrett's oesophageal mucosa. Williams et al. previously reported a method for quantifying the expression of MCM5 in shed oesophageal cells harvested from gastric fluid, as measured by an immunofluorometric assay. This study demonstrated MCM5 expression was significantly raised in individuals with adenocarcinoma or squamous cell carcinoma of the oesophagus compared to those with a normal oesophagus²¹³. Interestingly, almost half of the non-cancer cohort in this study comprised patients with non-dysplastic Barrett's oesophagus, but the relationship between the presence of Barrett's oesophagus, with or without dysplasia and how this relates to epithelial MCM5 expression was not characterised. We propose that there may be an incremental increase in MCM5 expression between patients with a normal oesophageal mucosa, non-dysplastic Barrett's oesophagus, dysplastic Barrett's oesophagus and oesophageal cancer. During endoscopy, a common finding is gastric fluid which contains exfoliated oesophageal epithelial cells; this is often aspirated and discarded during the examination to allow better visualisation of the gastric mucosa. We proposed that MCM5 expression could be quantified by direct analysis of its presence in these shed oesophageal epithelial cells, to interrogate whether there is differential MCM5 expression in the patient cohorts outlined above.

5.2 Aims of this study

There are two aims to this chapter, which presents a feasibility study. Firstly we assessed whether MCM5 expression can be quantified in oesophageal epithelial cells acquired endoscopically from gastric fluid samples. Secondly, we aimed to characterise the different expression levels of MCM5 in cohorts of patients stratified by their oesophageal histology. Our specific aims were:

1. To describe a method for quantification of MCM5 expression in exfoliated oesophageal epithelial cells obtained from aspirated gastric fluid during endoscopic assessment of the upper gastrointestinal tract.
2. To quantify MCM5 expression in patient cohorts with histologically distinct oesophageal mucosae. To assess whether there is an incremental increase in expression between patients with a normal oesophageal squamous mucosa, those with non-dysplastic Barrett's oesophagus, dysplastic Barrett's oesophagus and those with oesophageal adenocarcinoma.
3. To establish the feasibility of quantifying MCM5 expression, using a proprietary assay (Arquer), as a means of identifying the presence of Barrett's oesophagus associated dysplasia or adenocarcinoma.

5.3 Methods

Patient recruitment, inclusion and exclusion criteria

Patients with a normal oesophagus, Barrett's oesophagus (non-dysplastic, high-grade dysplasia) and known adenocarcinoma were recruited from a single tertiary referral centre in the UK (University College Hospital) between August 2017 and April 2018. Patients were recruited into one of four groups, depending on the histology results taken at the index endoscopy from which the gastric aspirates were obtained. The four subgroups were; macroscopically normal squamous oesophagus (NS), non-dysplastic Barrett's oesophagus (NDBE), high grade dysplastic (HGD) Barrett's oesophagus, oesophageal adenocarcinoma (OAC). Epithelial histology was verified by two expert gastrointestinal pathologists for all recruited patients.

Patients were excluded if they had concomitant systemic inflammatory conditions, active sepsis or infection. Since MCM5 expression is raised in cells obtained from other epithelial cancers and it is not known whether this effect is seen systemically, patients with other solid organ or epithelial malignancy were excluded. Similarly patients with severe oesophagitis or active oesophageal ulceration at endoscopy were also excluded. Patients were also excluded if they had previously been found to have low grade dysplasia or had biopsies indeterminate for dysplasia during a previous endoscopy. This decision was taken given the high inter-observer variability in the diagnosis of LGD and the current practice that these patients are often recalled for a follow up endoscopy as a result^{214,215}. Patients were also excluded if they had previously received chemo-radiotherapy or ablative endoscopic therapies for dysplasia (radiofrequency ablation, argon plasma coagulation, photodynamic therapy) since the effect of these interventions on MCM5 expression is unquantified. Patients with food contamination, with no gastric fluid to aspirate or those in who histologic samples could not be taken at the time of the index endoscopy were also excluded.

Our study had full ethical approval (IRAS 214612) and participants were required to give full written consent prior to enrolment.

Endoscopic procedures

All patients underwent upper gastrointestinal endoscopy under conscious sedation with Midazolam and Fentanyl. The use of a topical anaesthetic agent (Xylocaine) was permitted as per routine practice in our unit. All endoscopies were performed using a Pentax EG-2990Zi MagniView endoscope with *i*-Scan EPK-i7010 high-definition video processor. Prior to the procedure the working channel of the endoscope was flushed with air to prevent water contamination of the gastric fluid. The endoscope was carefully passed down the oesophagus to avoid mucosal trauma and no suction was applied during the passage into the stomach. A sterile plastic catheter was then passed down the working channel and used to aspirate up to 10ml of gastric fluid. The oesophagus was then cleaned with a solution of 2% simethicone and carefully inspected. This procedure is outlined in **(figure 22)**. After gastric fluid samples had been acquired, patients underwent a full upper endoscopy with or without biopsies depending on the suspected or confirmed histology seen. For patients with a normal oesophageal mucosa and no suspicious lesions they did not undergo histologic sampling, in line with routine practice. Patients with likely non-dysplastic Barrett's oesophagus underwent a gold standard upper gastrointestinal endoscopy with Seattle protocol biopsies^{140,141,187}. Patients with suspected high-grade dysplastic Barrett's oesophagus or with a lesion consistent with adenocarcinoma underwent histologic sampling either by forceps biopsy or endoscopic mucosal resection (EMR) at the endoscopists discretion.

Histologic classification of samples

Patients enrolled in this study with evidence of either non-dysplastic, dysplastic Barrett's oesophagus or adenocarcinoma underwent histologic sampling by either forceps biopsy or endoscopic mucosal resection. Patients with low grade dysplasia or those with samples indeterminate for dysplasia at the index or any previous endoscopy were excluded as discussed above. Patients found to have adenocarcinoma were stratified, as is routine practice according to invasion depth; those confined to the epithelium (M1); those invading into the lamina propria (M2); those extending into the muscularis mucosa (M3) and those invading into or beyond the first submucosal layer (SM1).

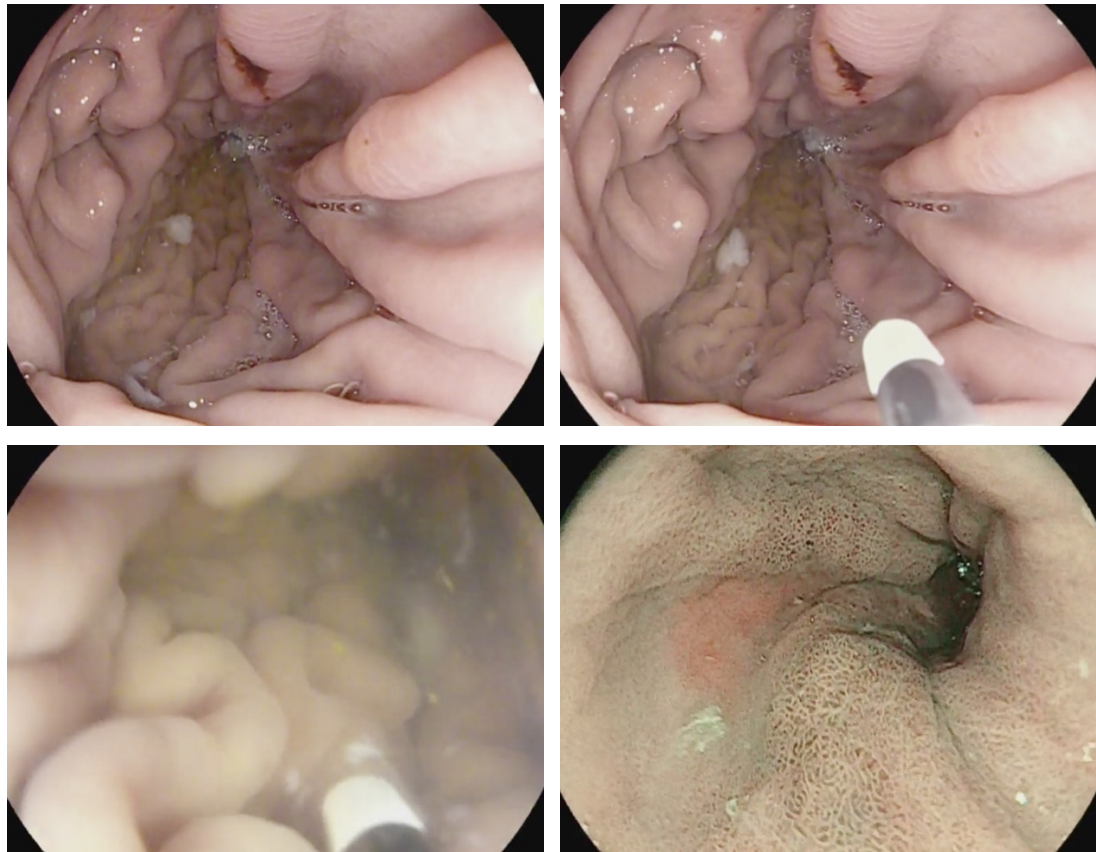


Figure 22: Procedure for collecting gastric aspirates intra-endoscopy. Clockwise from top left. 1) gastric fluid free of food or blood identified within the stomach 2) a sterile plastic suction catheter is passed down clean, dry working channel of endoscope into the gastric fluid 3) 5-10ml of gastric fluid aspirated from stomach and immediately refrigerated before analysis. 4) Histologic confirmation of neoplasia if present by biopsy or EMR

Sample preparation and storage

Samples of aspirated gastric fluid were stored in sterile vials and refrigerated at 4°C immediately after collection. All samples were processed within 4 hours of collection to prevent degradation of the oesophageal cells and MCM5 protein. The fluid was centrifuged at 1500g for 5 minutes to form a cell pellet and the supernatant was then carefully aspirated by pipette and discarded. The cell pellet was then fully resuspended using 500µl of cell lysis buffer. The lysed cell suspension was then stored for up to 3 weeks frozen at -80°C prior to analysis. To prevent degradation of MCM5 cells were transferred to an external lab for processing using dry ice to prevent thawing. The lysis buffer used in this method was a proprietary formulation provided for this study by Arquer Diagnostics Ltd.

Determining MCM5 expression

MCM5 protein expression levels were calculated using a proprietary immunoassay (Arquer Diagnostics Ltd) and reported in pg/ml. Epithelial cell numbers could not be quantified prior to cell lysis, in order to avoid degradation or alteration of the MCM5 protein. Therefore MCM5 expression levels for each subject were normalised according to the volume of gastric aspirate acquired at the index endoscopy of that patient. The level of MCM5 protein expression could then be calculated per mL of gastric fluid acquired.

Statistical analysis

Patients and their associated MCM5 expression level were stratified according to their histological results obtained from samples from the index endoscopy. Differences in the expression of MCM5 based on histological group were assessed using Kruskal-Wallis testing. This statistical test was chosen as it is suitable to assess for whether there is a significant difference between medians within a ranked, non-parametric dataset. To assess for differences in the medians between individual subgroups Dunn's tests of multiple comparisons was used. The diagnostic performance of the immunoassay was assessed by calculating its sensitivity and specificity. ROC curves were used to calculate the AUC to assess the accuracy of the assay. The AUC was categorised to assess how good a discriminator MCM5 expression was to detect patients with neoplastic compared to non-neoplastic oesophageal mucosae. The AUC was interpreted according to the following scale; 1-0.9 (very good), 0.9-0.8 (good), 0.8-0.7 (fair), 0.7-0.6 (poor), 0.6-0.5 (fail). Since this was a feasibility study and there are no previously reported similar investigations, formal power calculations were not undertaken. In this study we aimed to recruit 60 patients across all four subgroups to assess whether there was value in our group extending these results into a larger study at a later date.

5.4 Results

Patient demographics

In total 61 patients were included in this study. The mean segment length of Barrett's oesophagus is stratified by histologic subgroup and biopsy results taken during the index endoscopy in **(figure 23)**. Patients were recruited into one of four histological subgroups; normal oesophagus/acid reflux only; non-dysplastic BE; high-grade dysplastic BE; and adenocarcinoma. The mean age patients participating in this study was 67 years (range 26 – 89years). We observed no significant difference in the mean segment length of Barrett's oesophagus in patients of each of the histologic subgroups (NDBE vs Cancer $p > 1$; NDBE vs HGD $p = 0.4$; HGD vs Cancer $p = 0.28$). Of the cancer subgroup all patients had adenocarcinoma confirmed after EMR or ESD with variable invasion depths; those confined to the epithelium (M1; 3); those invading into the lamina propria (M2; 4); those extending into the muscularis mucosa (M3; 5) and those invading into or beyond the first submucosal layer (SM1; 3). The histologic characteristics of patients with histologically confirmed OAC is summarised in **(figure 24)**.

Patient demographics		
Age (years)	67 (range 26-89)	
Mean Barrett's length (total C+M in cm)	Normal	0
	NDBE	6.4
	HGD	5.8
	Cancer	9
Histology (by subgroup)	Normal	14
	NDBE	14
	HGD	18
	Cancer	15

Figure 23 Summary table showing patient demographics and histological subgroups recruited in the MCM5 study

Characteristics of OAC subgroup		
Invasion depth	M1	3
	M2	4
	M3	5
	SM1	3
Differentiation	Good	3
	Moderate	8
	Poor	4
Lymphovascular invasion	Present	1
	Absent	14
Metastatic disease	Present	1
	Absent	14

Figure 24 Summary table showing histological characteristics of patients with confirmed adenocarcinoma.

(M1 – intraepithelial invasion, M2 – lamina propria invasion (M1 – intraepithelial invasion, M2 – lamina propria invasion, M3 – muscularis mucosa invasion, SM1 – upper third of submucosal layer

MCM5 expression is raised in patients with high grade dysplastic Barrett's or adenocarcinoma compared to a macroscopically normal oesophagus

As shown in (figure 25), we demonstrate that levels of MCM5 expression are significantly raised in patients with oesophageal adenocarcinoma compared to patients with a macroscopically normal oesophagus. (mean expression 146.0 pg/ml vs 14.4 pg/ml, $p = 0.03$). When all four subgroups were analysed individually we only observed a significant difference in MCM5 expression between patients with adenocarcinoma and a normal oesophagus.

We did not demonstrate a significant difference in MCM5 expression between patients with a macroscopically normal oesophagus and those with non-dysplastic Barrett's or high-grade dysplasia. Similarly, there was no significant difference in MCM5 expression between patients with a non-dysplastic Barrett's and those with high grade dysplasia or cancer. The mean expression and differences between MCM5

expression levels we observed across the histological subgroups are summarised below in **(figure 25)** and **(figure 26)**.

Interestingly we did note a stepwise trend of increased mean MCM5 expression between patients with a normal oesophagus, non-dysplastic Barrett's oesophagus compared to those with dysplastic Barrett's or adenocarcinoma. While these results are not statistically significant, we note that this is a feasibility study with relatively low patient numbers. We postulate that there may some association between dysplasia or neoplasia and raised MCM5 expression, but clearly only a sufficiently powered study more patients in each subgroup will be able to confirm or refute this association.

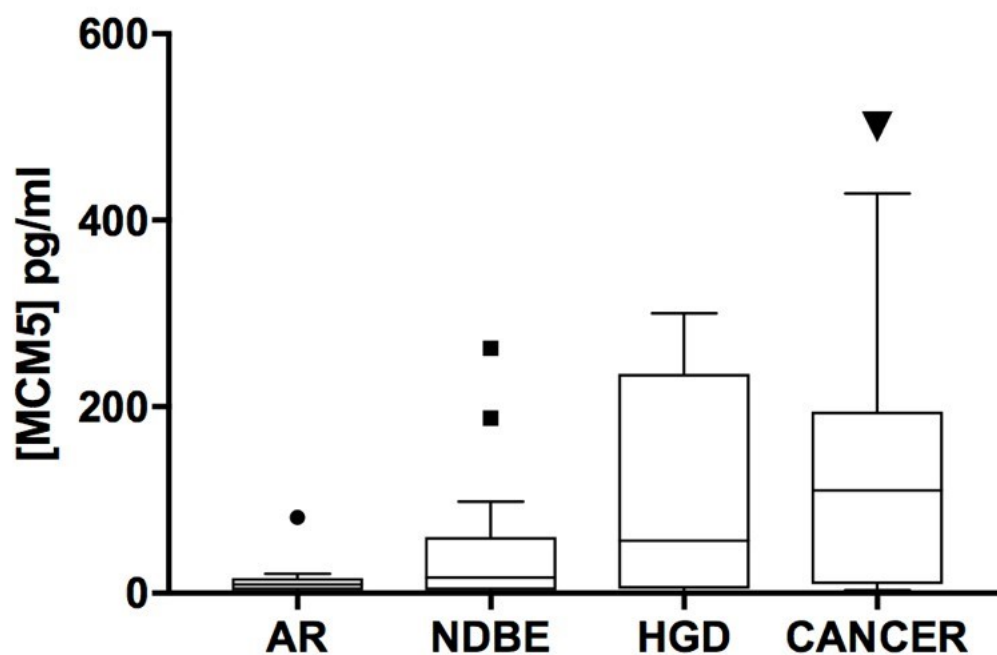


Figure 25 Comparison of MCM5 expression levels in gastric aspirate samples (pg/ml) between patients of each histological subgroup (AR: macroscopically normal/acid reflux without visible oesophagitis only, NDBE: non-dysplastic Barrett's oesophagus, HGD: high grade dysplasia and cancer).

Histologic subgroup	Mean MCM5 expression (pg/ml) [range]	Median MCM5 expression (pg/ml)
Normal	14.4 [0 – 81]	9.3
NDBE	49.8 [0.4 – 262.9]	16.7
HGD	112.3 [0 – 300]	56.0
Cancer	146.0 [3.3 – 500]	110.3

Difference in MCM5 expression	Significance (p value)
Normal vs NDBE	>0.99
Normal vs HGD	0.33
Normal vs Cancer	0.04*
NDBE vs HGD	>0.99
NDBE vs Cancer	0.34
NDBE vs Normal	>0.99

Figure 26 Comparison of differences between mean MCM5 expression levels and histological subtypes

We also assessed the diagnostic performance of this immunoassay for MCM5 expression as an indicative marker for the presence of oesophageal adenocarcinoma using a ROC curve (**figure 27**). Quantifying MCM5 expression was able to discriminate with fair accuracy the presence of oesophageal adenocarcinoma, compared to a macroscopically normal oesophagus (AUC 0.73 [95% CI: 0.62 – 0.96]). Using this immunoassay, 73% of patients presenting with an oesophageal adenocarcinoma would have a higher MCM5 expression compared to a patient with a macroscopically normal oesophagus ($p = 0.007$).

A feature that has high clinical utility for a biomarker for use in Barrett's surveillance is its ability to distinguish between non-dysplastic Barrett's, high grade dysplasia and adenocarcinoma. We observed a non-significant difference in MCM5 expression between these three histologic subgroups. Our assay is therefore unable to

differentiate patients with neoplastic histology (defined as high grade dysplasia or adenocarcinoma) from those with non-dysplastic Barrett's. There is however some promise that it may be able to differentiate between patients with a macroscopically normal mucosa and those with adenocarcinoma. There was no evidence that more invasive adenocarcinoma correlates with incrementally higher MCM5 expression levels. Analysis using a ROC curve showed that MCM5 expression is a fair differentiator between patients with non-neoplastic and neoplastic histology (AUC 0.70 [95% CI: 0.57 – 0.83]). Our results show that in 70% of cases a patient with either HGD or adenocarcinoma would have a higher MCM5 expression than a patient with either NDBE or normal histology ($p = 0.008$).

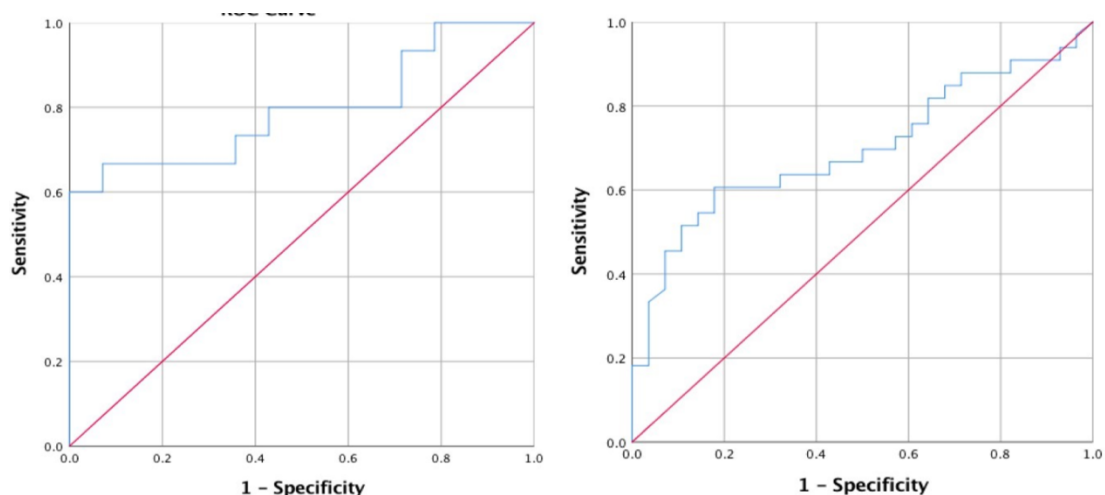


Figure 27 (Left) ROC curve (blue) for the diagnostic performance of our MCM5 expression assay for the characterisation of either a macroscopically normal oesophagus or adenocarcinoma compared to the null hypothesis reference line (red). (Right) ROC curve (blue) for the diagnostic performance of our MCM5 expression assay for the characterisation of patients as having neoplastic histology (adenocarcinoma or HGD) compared to non-neoplastic histology (NDBE or normal histology). The null hypothesis reference line is shown in red.

5.5 Discussion

MCM5 is a cell cycle protein that has been demonstrated to play a putative role in cell cycle regulation through its involvement in forming the DNA replicative helicase. Subsequent work observed that aberrant MCM5 expression and the resultant dysregulation of the cell cycle is implicated in numerous epithelial cancers – including bladder, cervical and oesophageal cancer^{209,211,216}. Williams et al have previously shown that exfoliated oesophageal epithelial cells can be acquired from gastric fluid aspirates, when paired with immunofluorometric analysis, is a feasible method of quantifying MCM5 expression²¹³. Furthermore, they demonstrated that MCM5 expression was raised in patients with oesophageal squamous cell carcinoma or adenocarcinoma, compared with those with a macroscopically normal oesophagus. In this study, almost half of the patients with a normal oesophagus in fact had Barrett's oesophagus without dysplasia.

This study arose from the finding that a significant number of patients in the Williams et al. study, with non-dysplastic Barrett's oesophagus appeared to have a normal level of MCM5 expression whereas patients with cancer had raised expression levels. We postulated that there may be a step-wise increase in MCM5 expression between patient with a normal oesophagus, non-dysplastic Barrett's through to adenocarcinoma. Given the inherent limitations of the Seattle protocol for Barrett's surveillance, as well as the high miss rate of endoscopy for the detection of adenocarcinoma¹⁴⁵, quantification of MCM5 could raise the suspicion of clinicians that a cancer has been missed.

This prospective feasibility study assessed whether a proprietary assay developed by Arquer Diagnostics Ltd could quantify MCM5 levels in oesophageal cells obtained from gastric aspirates. Furthermore we have reported a methodology to assess whether raised MCM5 expression in gastric fluid is associated with the presence of dysplastic Barrett's oesophagus or oesophageal adenocarcinoma in patients undergoing endoscopy. This study demonstrates that MCM5 levels are significantly raised in patients with oesophageal adenocarcinoma compared to

patients with a macroscopically normal oesophagus (154.1 vs 14.4 [$p = 0.04$]). We observed a stepwise association in the mean MCM5 expression levels between patients with a normal oesophagus, NDBE, HGD and oesophageal adenocarcinoma (14.4, 49.8, 112.3, 146.0 respectively), but note that there was no statistically significant difference in MCM5 expression levels between patients with NDBE, HGD and adenocarcinoma. Using ROC curves to assess the diagnostic performance of our assay, we demonstrate that MCM5 expression is a fair discriminator between patients with a macroscopically normal oesophagus and those with adenocarcinoma (AUC 0.73, $p = 0.007$). The MCM5 expression level was also a fair discriminator (AUC 0.70, $p = 0.008$) between patients with neoplasia (HGD or cancer) compared to those without neoplasia (NDBE or a normal oesophagus).

These results show promise that MCM5 may be a candidate biomarker for the detection of oesophageal adenocarcinoma in patients. While we aimed in this study to investigate how MCM5 expression corresponds to the metaplasia-dysplasia sequence seen in Barrett's oesophagus our results are not statistically significant. As a feasibility study this may be that the study was underpowered to detect this effect. As outlined in the introduction, there is still a significant resource burden produced by patients being referred with 'red flag' upper gastrointestinal symptoms to secondary care for endoscopy, who are found to have no serious underlying pathology. We demonstrate in this chapter that MCM5 expression in patients with a macroscopically normal oesophagus is significantly lower than those with histologically confirmed adenocarcinoma. A promising role for the quantification of MCM5 expression that should be investigated is whether it can accurately distinguish patients referred with red flag symptoms with cancer, from those who have symptoms but a benign pathology. We report a method by which MCM5 can be quantified from gastric aspirate samples; this fluid need not necessarily be obtained by endoscopy, less invasive methods such as nasogastric tubes could also be used to obtain these samples. We propose that further work should therefore investigate less invasive methods by which gastric aspirates could be obtained, to allow for the use of MCM5 to triage patients with alarm symptoms who may not need to undergo endoscopy since their risk of neoplasia is very low.

Limitations of this study and potential further work

The first limitation of this body of work is, as would be expected of a feasibility study, this study used a relatively small number of patients. Therefore, the observed lack of significant associations between MCM5 expression and the presence of neoplastic lesions in the oesophagus may be because the study was underpowered to detect more subtle associations. Larger, multi-centre studies should interrogate MCM5 expression between these histological subgroups to better assess for variations in expression levels; this study will provide a benchmark methodology and data for power calculations to enable this.

Future studies should also assess for dysregulated MCM5 expression in NDBE and investigate whether increased segment length affects expression levels. We suggest that if MCM5 expression is raised in NDBE prior to the development of neoplasia, longer Barrett's segments may yield higher MCM5 expression and so could give false positive results for dysplasia. Similarly, larger studies could be powered to interrogate whether there is a link between tumour size or invasion depth and MCM5 expression levels. There was no clear association between this in this study, but given the small sample size this may be due to underpowering.

Secondly, we observed a wide range in the level of MCM5 expression in patients with non-dysplastic Barrett's oesophagus, with the difference in mean MCM5 expression in the group not statistically significant compared to the mean expression in patients with high grade dysplasia or cancer. The gold standard for diagnosis used in this study, against which patients MCM5 expression levels was compared, was histological sampling of the oesophageal mucosa taken at the time of endoscopy. As discussed above the accuracy of this gold standard may be questionable. The sensitivity for dysplasia/neoplasia detection on random biopsies taken through a Barrett's oesophagus segment varies widely in reported studies (33->90%), largely because such a sampling technique samples less than 5% of the total mucosal surface area²¹⁷. It should therefore be noted that patients in this study categorised as having only non-dysplastic Barrett's based on their histology results, but who were found to

be high MCM5 expressors, may have had undetected neoplasia missed by random biopsies. Similarly, it may be that the evolution of dysplasia with BE tissue is a stepwise event, aberrant MCM5 may be one of several cellular changes that precede the development of dysplasia – hence raised MCM5 expression in isolation may not be consistently demonstrative of dysplasia. Future studies should consider whether patients with high MCM5 expression, but with no histologic evidence of dysplastic Barrett's, go on to develop neoplasia at a later date with a higher frequency than patients without raised expression.

A third limitation of this study relates to the methodology. Due to logistical implications, cells acquired from gastric aspirates needed to be lysed within four hours of collection at endoscopy. Expression was therefore quantified using a proprietary assay and the expression levels normalised compared to the volume of gastric aspirate acquired from each patient, since by the time of analysis it could not be normalised according to cell counts. This method was introduced to allow a fair comparison of expression levels between patients where variable quantities of gastric fluid were acquired. It was therefore assumed that the number of cells per unit volume of gastric fluid, would be roughly consistent between patients. An inherent limitation of this study is that gastric fluid volumes may not correlate with the number of cells present, for instance a patient with a large volume of gastric aspirate may have a higher proportion of that fluid made up of gastric juice than cells, compared to another patient who may have had a higher concentrations of sloughed oesophageal cells despite lower aspirated fluid volume. We suggest that future studies should aim to quantify the cell concentration in aspirated gastric fluid samples prior to any cell lysis step; using this figure to normalise MCM5 expression levels between patients. We also note that a small number of patients in this study recorded an MCM5 expression level of 0pg/ml. Studies have shown that MCM5 is expressed at a basal level throughout all stages of the cell cycle so it is improbable that patients should have undetectable levels. Given the supernatant aspiration step outlined in the methods section, some of the cell pellet may have been aspirated despite careful attention to leave it intact. Quantification of cell numbers prior to processing would allow more accurate identification of patients in whom a low total cell count would

give a falsely low MCM5 expression level; these patients could then be excluded from analysis.

Another potential limitation of this study is that in patients with dysplastic or neoplastic lesions it is unknown whether there is a higher degree of cell shedding, or whether the potentially friable mucosa is more easily disrupted by the passing endoscope in these patients. We attempted to mitigate this by ensuring that our protocol detailed a standardised procedure for intubating and passing through the oesophagus with minimal disruption to the mucosa. This potential source of bias could also be overcome as described above through normalising expression volumes against cell counts rather than by aspirate volume.

This study was industry sponsored and utilised a proprietary assay. This was a logistical choice by the study team. Since the quantification of MCM5 expression levels required the facilities to undertake an ELISA with specific labelled antibodies we did not have the resource to undertake this.

5.6 Summary of this chapter

This chapter presents a feasibility study to assess whether increased MCM5 expression in epithelial cells acquired from gastric fluid samples, correlates with the presence of progressively more dysplastic Barrett's oesophagus or adenocarcinoma.

1. We describe a method for the quantification of MCM5 expression by immunoassay in exfoliated oesophageal epithelial cells obtained from endoscopically aspirated gastric fluid.
2. We demonstrate that oesophageal epithelial cell expression of MCM5 protein is significantly lower in patients with macroscopically normal oesophageal mucosa, compared to those with adenocarcinoma

Chapter 6– Assessing whether virtual chromoendoscopy using iScan Optical Enhancement improves the detection of Barrett’s oesophagus associated neoplasia in expert and non-expert endoscopists.

The work presented in this chapter formed the basis of a peer reviewed publication. Text and figures were adapted for publication. Citation:

Everson, M.A., Lovat, L.B., Graham, D.G., Bassett, P., Magee, C., Alzoubaidi, D., Fernández-Sordo, J.O., Sweis, R., Banks, M.R., Wani, S. and Esteban, J.M., 2019. Virtual chromoendoscopy by using optical enhancement improves the detection of Barrett’s esophagus–associated neoplasia. *Gastrointestinal endoscopy*, 89(2), pp.247-256.

6.1 Introduction

As previously discussed in **chapter 3**, Barrett’s oesophagus is a known precursor to oesophageal adenocarcinoma. Undetected oesophageal adenocarcinoma typically presents late in the course of the disease, with an attendant overall 5 year survival in the UK of around 15% for stage 3 cancers¹¹. As outlined above, early neoplastic lesions arising in BO which are confined to the mucosa are amenable to endoscopic eradication therapy; including but not limited to endoscopic mucosal resection or endoscopic submucosal dissection. The use of endoscopic therapies for the eradication of Barrett’s associated neoplasia is associated with high cure rates^{126,170,171,200,218}. Of equal importance is that these therapies may potentially avoid the need for oesophagectomy, with its associated morbidity and mortality.

The gold standard for the detection and diagnosis of early Barrett’s associated adenocarcinoma, is gastroscopy. To facilitate the early detection of these lesions patients with histologically confirmed BE, should be enrolled into an endoscopic surveillance program¹⁸⁷. In the UK patients with a confirmed diagnosis of Barrett’s oesophagus are offered between 2 and 5 yearly endoscopic surveillance with Seattle protocol biopsies. The Seattle protocol (SP)²¹⁹ requires that the Barrett’s segment is examined carefully by high definition-white light (HD-WLE) endoscopy. Following a visual assessment any visible abnormalities are target biopsied, then random quadrant biopsies are taken at one to two centimetres through the remaining segment with the aim of detecting dysplasia not visualised on the initial assessment.

Since early lesions are often subtle, focal and therefore easily missed on endoscopic surveillance examinations, there are inherent limitations of the Seattle protocol. There is some evidence that the current sensitivity of HD-WLE for endoscopic Barrett's surveillance may not be sufficient for the detection of early neoplasia. Visrodia et al. estimate that up to 25.3% (95%CI:16.4-36.8%) of new adenocarcinoma diagnoses are made following a normal surveillance endoscopy in the preceding year²²⁰, and can therefore be regarded as 'missed lesions'. Furthermore, the sampling procedure outlined in the Seattle protocol is also prone to sampling error since less than 5% of the Barrett's epithelium is sampled during a typical endoscopy¹⁴⁶ – even in instances where the protocol is adhered to as recommended. It has also been shown that SP biopsies and adherence to the protocol worsens with increased segment length, which likely further compounds the miss rate of adenocarcinoma during routine HD-WLE Barrett's surveillance¹⁴⁶. Even where lesions that are endoscopically suspicious for dysplasia are identified, SP biopsies generate a large number of samples with a low reported sensitivity for dysplasia/neoplasia detection, ranging from 33->90%¹⁴⁸.

There are clear limitations to the sensitivity of HD-WLE and the Seattle protocol for the detection of Barrett's associated neoplasia. Prior to this study, an emerging body of work has studied the role of advanced endoscopic imaging to try to mitigate some of these limitations in endoscopic surveillance. Assessment of the oesophageal mucosa using advanced imaging systems may improve the detection of dysplasia. To date work has largely focused on identifying mucosal and vascular abnormalities that are associated with dysplasia^{164,221–223}.

In recent years our group has developed experience working with a new advanced imaging system. The iScan Optical Enhancement (OE) system (Pentax Hoya, Japan) is a novel advanced imaging technology demonstrating a range of clinical applications in upper and lower gastrointestinal endoscopy¹⁶². To facilitate improved visualisation of the mucosal surface of the upper gastrointestinal tract, the OE system employs both pre and post-processing technologies to provide surface enhancement of the superficial mucosal structures, as well as accentuating the visibility of the mucosal microvasculature.

A novel optical filter delivers specific wavelengths of light which correspond with the main absorption spectrum of human haemoglobin (415nm, 540nm and 570nm) at high light intensities. This increased illumination across the whole haemoglobin absorption spectrum allows the accentuation of the vascular structures within the most superficial layers of the oesophageal mucosa (**figure 28**). The endoscopes equipped with OE imaging technology are also often equipped with magnification endoscopy of up to 136x. This feature can be used in conjunction with either HD-WLE or OE and facilitates much closer interrogation of the microvasculature of the oesophageal mucosa. In line with other studies examining other advanced imaging modalities, we propose that these microvascular and mucosal abnormalities seen on magnification endoscopy may provide a visible endoscopic marker of dysplasia.

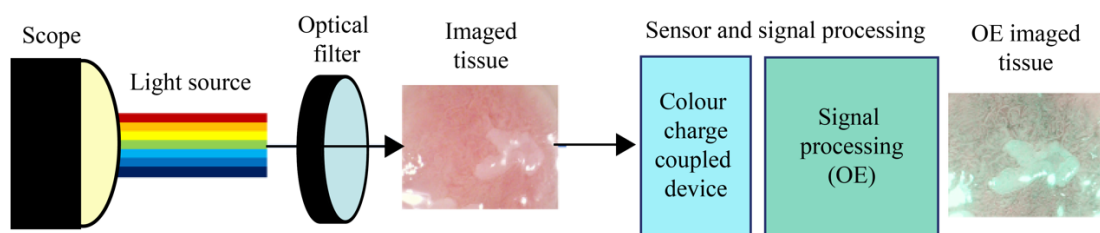


Figure 28 Schematic diagram of the image pre and post processing technology incorporated within iScan OE endoscopic imaging technology.

Our research group has previously validated the use of an early iteration of iScan technology (contrast, surface and tone enhancement or iScan 1/2/3) for use as an adjunct in the detection of BE dysplasia. Lipman et al proposed a simple classification system, based on mucosal and vascular patterns¹⁶⁴, to identify early neoplastic lesions. An additional improvement in the detection of dysplasia was achieved through the application of the chromo-endoscopic agent acetic acid to the mucosa; although clinically this may lengthen procedure times and may not always be available. One limitation of this study and others using advanced imaging technologies is that the improvement in diagnostic performance they confer is often assessed amongst expert endoscopists in a high volume setting. Whilst the majority of the endoscopic therapy for Barrett's associated dysplasia is undertaken in tertiary centres, the majority of surveillance is still undertaken outside of these settings. There is little reported in the

literature regarding the additional benefit for the detection of dysplasia these systems may offer non-expert endoscopists.

6.2 Aims of this study

The aim of this chapter is broadly to assess whether the detection of Barrett's oesophagus associated neoplasia can be improved through the use of the iScan OE advanced imaging platform, alongside the current gold standard Seattle protocol. Our specific aims are:

1. To assess the diagnostic performance of iScan OE compared to HD-WLE in the endoscopic detection of Barrett's oesophagus associated neoplasia in a cohort of expert endoscopists.
2. To assess the diagnostic performance of iScan OE compared to HD-WLE in the endoscopic detection of Barrett's oesophagus associated neoplasia in a cohort of trainee endoscopists.
3. To assess whether OE can improve the diagnostic performance for the recognition of Barrett's associated neoplasia in a cohort of expert endoscopists using both HD-WLE and OE. This analysis was performed using a previously published¹⁶⁴ consensus driven magnification endoscopy classification system for use with OE and HD-WLE.

6.3 Methods

Patient recruitment, inclusion and exclusion criteria

Patients attending one of three European referral centres (University College London Hospital, Hospital Clinico San Carlos, Madrid and UZ Leuven Hospital, Belgium) for either endoscopic surveillance or therapy of histologically confirmed Barrett's oesophagus were enrolled between Feb 2016 and Oct 2017. All included patients had a Barrett's segment of at least C1M2 length as classified by the Prague classification. Patients were excluded if they had received previous endoscopic eradication therapy for BE neoplasia, including EMR, ESD, RFA or other ablative therapies. Patients who had undergone surgical procedures to the oesophagus were not included. Patients with active oesophageal ulceration or varices were also excluded. The study had ethical approval and was registered with an International Standard Randomised Control Trial Number (ISRCTN) (Registration: 58235785).

Endoscopic procedures and image acquisition

All endoscopic examinations recorded for this study were undertaken by endoscopists with extensive expertise in the assessment and management of dysplastic BE (RJH, RB). After providing written consent, patients underwent either local anaesthesia or deep conscious sedation prior to the procedure. Mucous and food debris was removed from the oesophageal mucosa using a solution of simeticone and water, applied via the endoscopes working channel. A protocol was devised to ensure that all endoscopists followed the same procedure during their endoscopic assessments; in order to ensure comparability between videos recorded at disparate sites (**appendix 2 and 3**). By adhering to the protocol the endoscopist first undertook a 'pullthrough' manoeuvre through the oesophagus. Having reached the gastro-oesophageal junction (GOJ), the endoscopist began to slowly withdraw the endoscope from the GOJ to the proximal extent of the BE segment (**figure 29**). As would be encountered in clinical practice the endoscopist was instructed to stop if the view was obscured by passing saliva, or peristalsis of the oesophagus and to continue from the same position once the view was unobstructed.

Prior to any biopsies being taken the endoscopist was required to identify an area that they felt was dysplastic and from which in normal clinical practice they would take a biopsy. The endoscopist then used magnification endoscopy to record this area of mucosa at up to 136x magnification in both HD-WLE and in iScan OE. Following this recording a forceps biopsy was taken from this area to allow matched histology to be recorded. In patients with no areas suspicious for dysplasia the endoscopist was required to undertake this same recording procedure on an area they felt contained only macroscopically non-dysplastic Barrett's oesophagus.

All examinations were recorded in HD-WLE and iScan OE. All videos were recorded using a Pentax EG-2990Zi MagniView endoscope with *i*-Scan EPK-i7010 high-definition video processor. Once recording was complete patients either underwent routine Barrett's surveillance and Seattle protocol biopsies, or received endoscopic resection therapy if indicated.

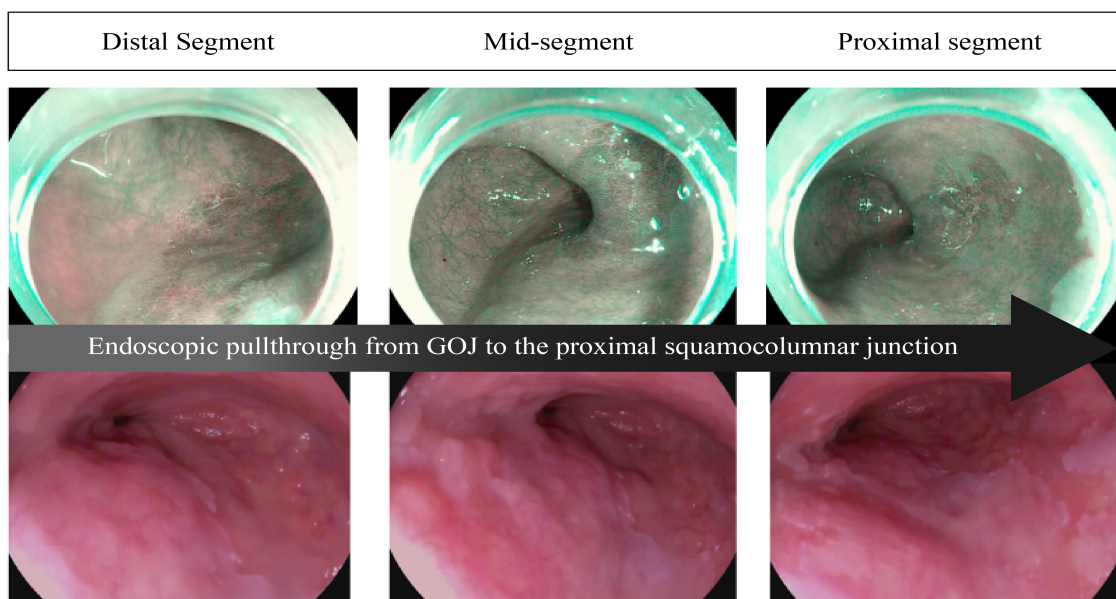


Figure 29 Representative example of how images were generated throughout the BE segment by carrying out a steady "pull through" sequence to simulate the normal endoscope withdrawal manoeuvre performed during BE surveillance endoscopy. From left to right: At the distal oesophagus/gastro-oesophageal junction, mid-section of the Barrett's mucosa, at the proximal squamocolumnar junction of the Barrett's segment. Top row: iScan OE, bottom row HD-WLE.

Tissue processing and histologic analysis

Where suspected neoplastic lesions were identified, the borders were demarcated using an electrocautery snare. Tissue was then resected by EMR. Biopsies of suspected normal or non-dysplastic areas were undertaken using forceps biopsy in accordance with the Seattle Protocol. Histology samples were retrieved and affixed with pins to cork board, prior to immediate fixation in formalin. Samples were processed within the individual hospital laboratory; embedded in paraffin and sectioned at multiple levels. All samples with confirmed dysplasia or cancer were assessed and reported by two expert GI pathologists.

Image pre-processing and analysis

Images were extracted as single frames from high definition video recordings and saved in the high quality, lossless .png format. As described above still images were generated throughout the distal, middle and proximal BE segment to simulate the normal 'pullthrough' manoeuvre performed during BE surveillance and outlined above in (**FIGURE 21**). All videos were assessed by a study member (MAE) for blurring, clarity and to ensure they were representative of 'real-life' endoscopic images, from which it was deemed that a meaningful diagnosis could be made. Videos and their images were also excluded if the location of biopsies taken at the time of endoscopy could not be established on the recorded video. A range of lesions were selected including LGD, HGD and OAC as well as videos of normal BE segments, in order to replicate the early lesions typically encountered in clinical practice. Images matched by position, lighting and endoscope angle using HD-WLE and iScan OE were selected where possible. A total of 262 images were included for analysis (130 HD-WLE and 132 OE).

Establishing an expert consensus on the delineation of neoplastic lesions

Both of the expert endoscopists, who recorded the endoscopic assessments described above and who had performed the biopsies or endoscopic resections of lesions suspicious for dysplastic lesions (RJH,RB), were unblinded to patients histology results. Both endoscopists were therefore excluded from the final analysis of pullthrough and magnification images that they had recorded. Instead, their

consensus was used to generate the gold standard; the area that was deemed to contain dysplastic Barrett's mucosa based on its endoscopic appearance.

To generate this gold standard, for both HD-WLE and OE images, each expert assessed each image and delineated areas that they felt represented dysplastic tissue based on the macroscopic appearance of the mucosa; their knowledge of the lesion from the viewed endoscopy video and their knowledge of the histology results for that patient. Delineations on each image were made using the GNU image manipulation program (GIMP V2.8.22). Both expert delineations were applied to the images in overlay on the to define the area deemed positive for dysplasia (**figure 30**). The area defined as positive for dysplasia was the area included within both experts delineations.

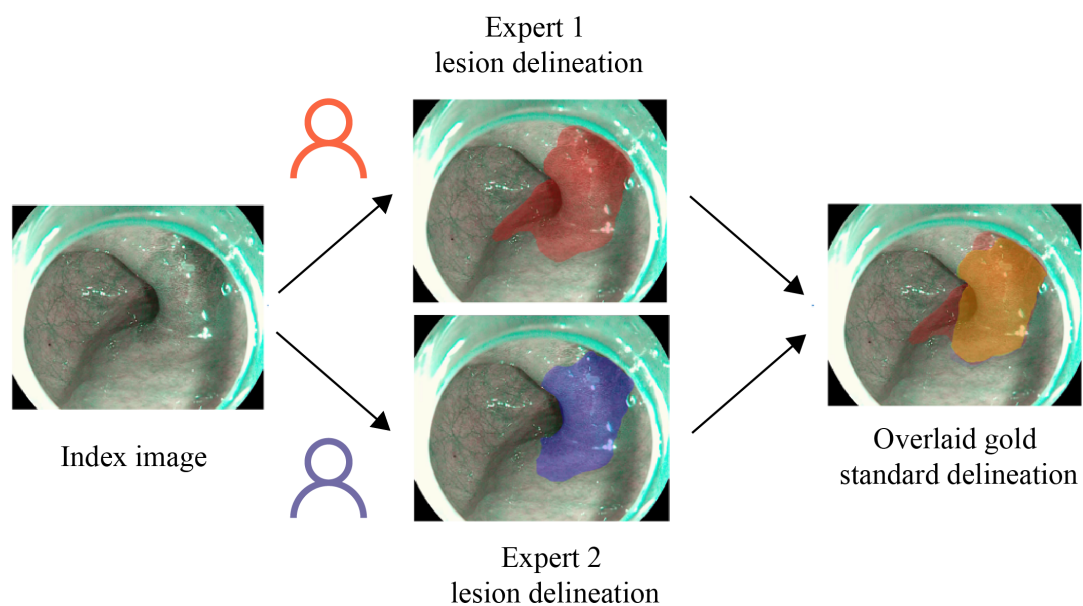


Figure 30 Representative example showing how the gold standard delineation (yellow, right) was generated from the two expert delineations shown in the middle column (red and blue) of a suspicious area seen here in iScan OE (left). This process was repeated for each individual image used in the study. The area marked in yellow represents the area deemed positive for dysplasia when the blinded endoscopists made their assessments of images for dysplasia.

Assessment of pullthrough images by trainees and expert endoscopists for dysplasia

A second group of 7 expert and 7 trainee endoscopists, all of whom were blinded to the initial endoscopic assessment video, biopsy location or resection margins and patient histology, were asked to assess each image and make a decision whether a dysplastic lesion was present in the image.

Experts were defined as clinicians who had completed their formal training and undertook >50 cases assessing and treating early BE neoplasia a year. Trainees were defined as those who had not yet completed formal training but had at least 3 years of endoscopy experience, previous exposure to BE surveillance endoscopy but with no formal training using OE.

Study participants reviewed the images of lesions alone, using high definition (HD) screens. To mimic how lesions would likely be assessed in clinical practice, participants reviewed HD-WLE images from all of the patients initially. They then assessed OE images from all of the patients – this delay between HD-WLE and OE assessments served as a short washout period between assessments for each patient. When assessing images that the endoscopist felt contained dysplasia, they were instructed to electronically place a single marker on each image over an area that they felt would be most likely to yield a biopsy with BE neoplasia if they were undertaking an endoscopy. A positive result was recorded when an endoscopists target biopsy fell within the consensus area delineated by the expert endoscopists (**figure 31**). If the endoscopists felt that the image did not display a lesion suggestive of dysplasia, they were instructed to make no mark on the image.

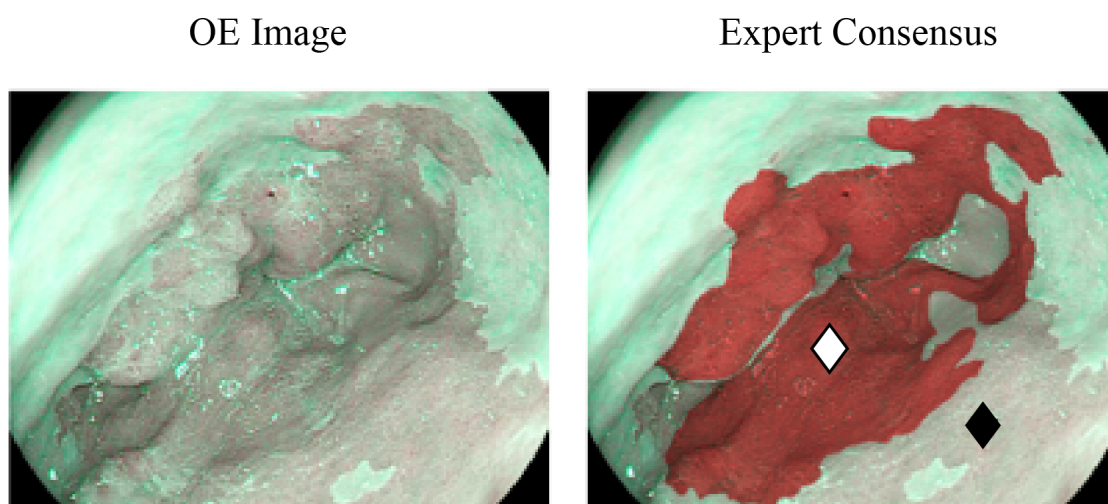


Figure 31 Illustrative example of the expert delineated consensus area considered positive for dysplasia (red). Assessor biopsy sites considered a true positive (white) and false negative (black). Images where the assessor made no mark were deemed to be classified as normal

Assessing the role magnification endoscopy using HD-WLE and OE for the identification of neoplastic tissue within Barrett's oesophagus

In phase two of this study, we assessed the role of OE compared to HD-WLE when used in conjunction with magnification endoscopy to interrogate microvascular and mucosal patterns. Magnification endoscopy was used to produce matched images of the oesophageal mucosal surface at up to 136x zoom in both HD-WLE and OE of normal and abnormal areas of BO (**figures 32, 33**). All images were captured at between 80-136x zoom. Any imaged area was either sampled by forceps biopsy from the area focused on, or was within the borders of a macroscopically dysplastic lesion that was histologically confirmed to be dysplastic by either EMR or ESD. Images depicting non-dysplastic tissue were only acquired from patients within the normal/non-dysplastic histology cohort. Conversely images depicting dysplasia were only obtained from patients with proven dysplasia, no non-dysplastic images were obtained from these patients.

For this analysis only the experts were asked to classify images as dysplastic or non-dysplastic based on the MV classification previously validated for use with the iScan system (**figures 32, 33 and 34**)²²⁴. This decision was made as currently magnification

endoscopy is not widely available in all endoscopy suites, nor is it currently recommended as part of a gold standard assessment for Barrett's surveillance. Within most UK healthcare settings, and by inference most developed healthcare settings, magnification endoscopy is not widely available. This technology is typically only used in high volume referral centres or in academic settings and may be used to assist lesion recognition and lesion demarcation for endoscopic resection planning. It was therefore decided that due to the lack of availability of this technology to non-expert endoscopists and the fact that it is not recommended for use in routine Barrett's surveillance at present, to not include non-expert or trainee endoscopists in this part of the analysis. All seven experts who undertook the lesion recognition analysis and who had prior knowledge and training in the use the iScan MV classification system from previous studies took part in this section of the study.

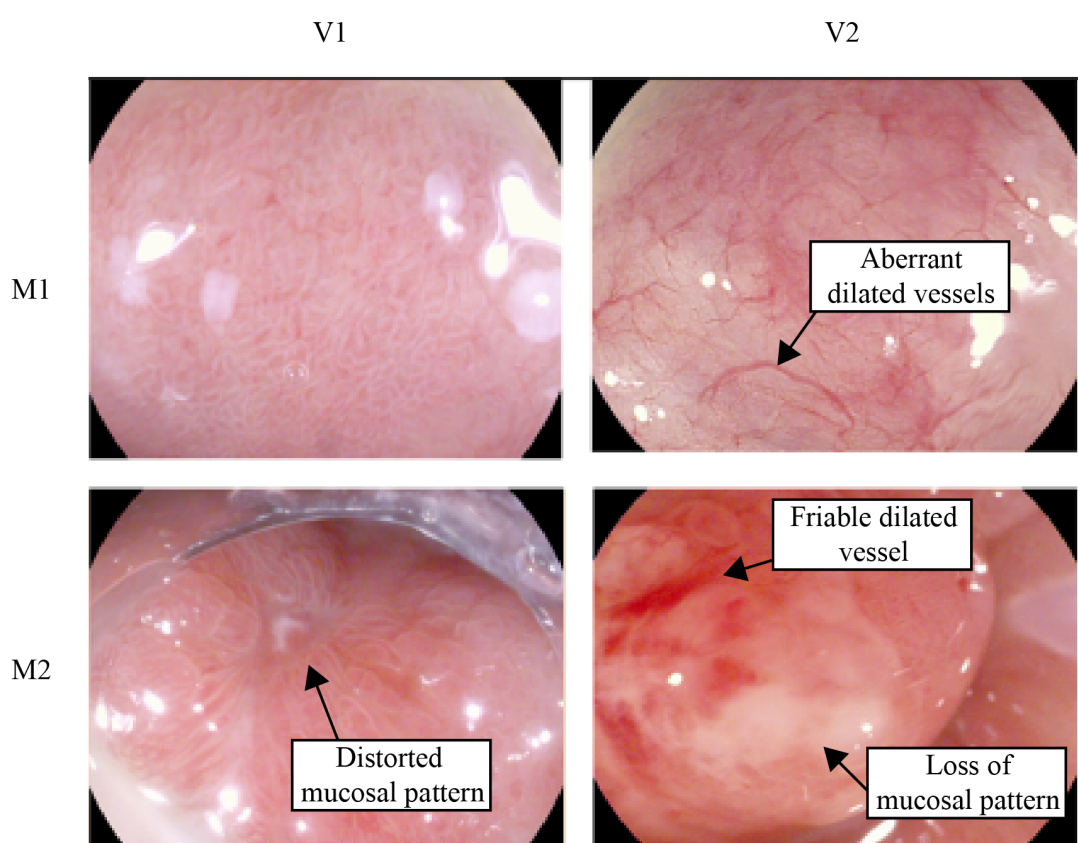


Figure 32 Representative magnification images of Barrett's oesophagus classified with the MV classification as non-dysplastic (top left) and dysplastic (other images), using HD-WLE

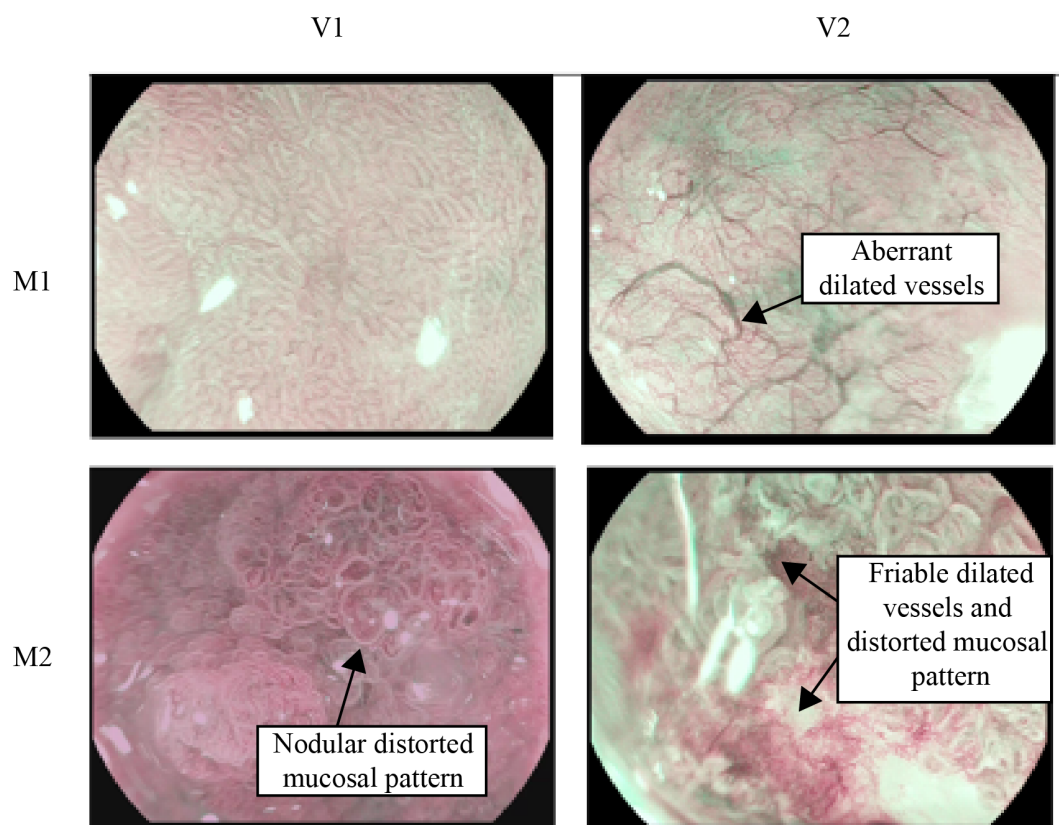


Figure 33 Representative magnification images of Barrett's oesophagus classified with the MV classification as non-dysplastic (top right) and dysplastic (other images), using iScan OE

	Mucosal features (M)	Vascular features (V)
1	Regular, gyriform pit pattern	Regular, fine calibre vessels between pits
2	Irregular, disordered pit pattern or featureless mucosa	Irregular, tortuous or dilated vessels not confined by pits

Figure 34 An overview of the iScan MV classification for use with the iScan magnification endoscopy system. A score of M1V1 represents non-dysplastic Barrett's oesophagus. Any classification containing at least one of an M2 or a V2 represents Barrett's tissue suspicious for dysplasia.

Statistical analysis

Diagnostic performance measures were calculated for each endoscopist on a per image basis in part one of this study. For each endoscopist, both expert and trainee, we calculated; dysplasia detection accuracy, sensitivity, specificity, positive predictive value, negative predictive values. Since there were assessments of multiple images from the same patient it was acknowledged that the data may have been non-independent. We therefore used multi-level logistic regression using a cross classified structure (measurements were nested within both patients and observers). Improvements in diagnostic performance measures between HD-WLE and OE were therefore reported as an odds ratio (the odds of there being an improvement when using OE compared with HD-WLE).

Interobserver variability was calculated using the Kappa statistic. K values and their standard errors were used to perform a Z test to assess whether there was any statistical variation in the agreement of endoscopists using the two modalities. A modified Likert scale developed by Landis and Koch was used to interpret K values which were interpreted according to the following scale: poor <0.20; fair =0.21-0.40, moderate =0.41-0.60, substantial =0.61-0.80; very good =0.81-1.00.

The same methodology was utilised in part two of the study where endoscopists assess ME images, in order to calculate the same diagnostic measures on a per image basis.

Sample size calculations

A previous study by our group demonstrated that for HD-WLE a panel of experts had a pooled dysplasia detection accuracy of 76%. We powered our study to assess if OE would confer up to a 6% improvement in accuracy up to 82% in a cohort of expert endoscopists. For the purposes of this power calculation we considered OE and HD-WLE as two different modalities and therefore based our calculations on comparing two independent groups of data.

As discussed above we acknowledge that due to the fact that individual endoscopists made multiple assessments of images from the same patients, the results may not be independent. The degree of clustering between these repeated measurements in the same patient are unknown. Using 80% power and a 5% significance level, we calculated that 723 individual measurements per imaging modality were required. To ensure that a reasonable margin was allowed to account for the unknown degree of clustering, we chose to double the sample size.

We therefore calculated that 1446 measurements for each of the modalities (HD-WLE and OE) would be required. Assuming that an average of 3 images per modality, per patient were assessed by 14 endoscopists, this would generate a total of 42 image assessments per modality, per patient. We therefore calculated that 35 patients were required to be recruited to this study.

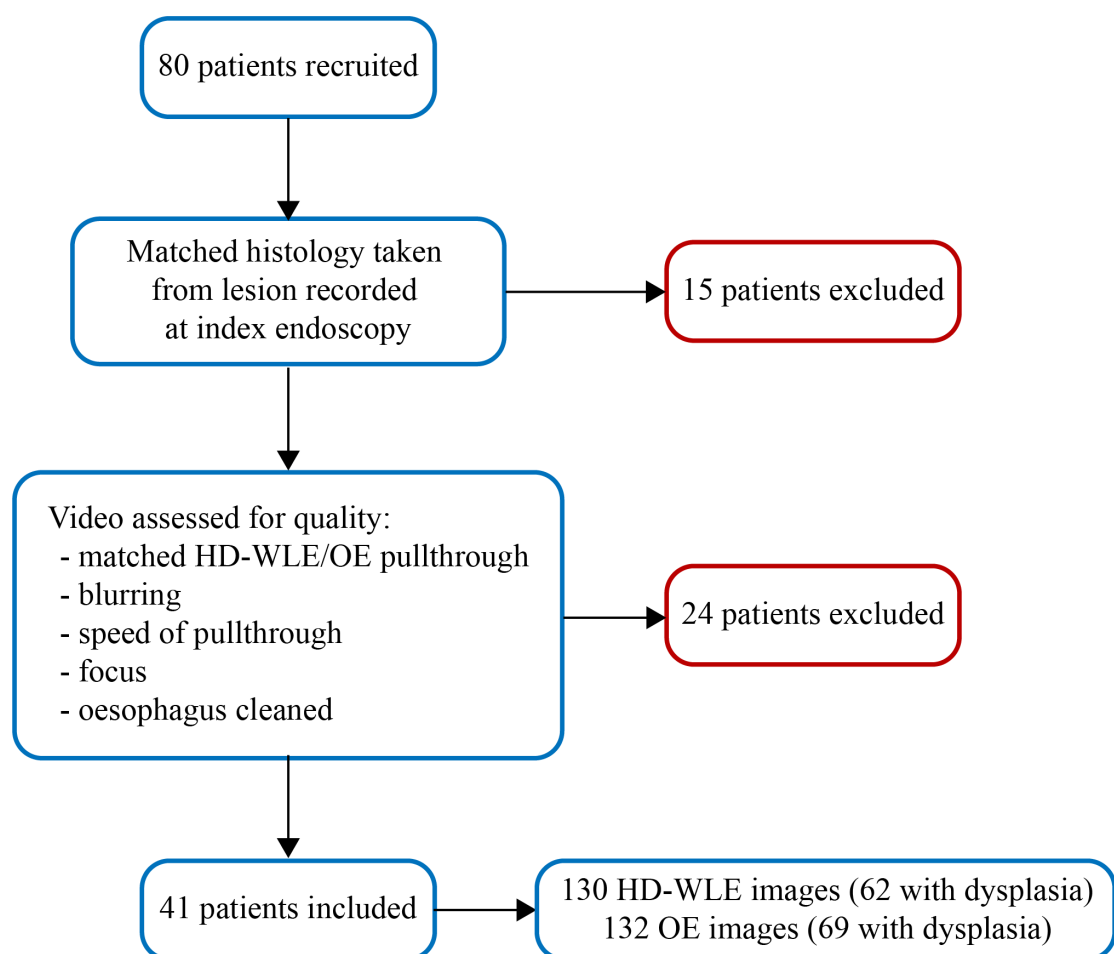


Figure 35 Study workflow for investigating the utility of iScan OE in the detection of Barrett's neoplasia

6.4 Results

Patient characteristics

80 patients were recruited to the study, all of whom had provided valid consent and had undergone histologic sampling at the time of the index endoscopy. Videos were excluded if they were deemed to be of poor quality by a study clinician (MAE). Causes for exclusion included blurred images, excessively bleeding mucosa, or recordings where the pullthrough was out of focus. Similarly, recordings were excluded if matched histology samples (either forceps biopsy or EMR/ESD specimens) corresponding to the imaged mucosa were not retrieved at the index endoscopy.

262 images from 41 patients were included after quality control (**figure 35**). 62/130 HD-WLE images contained visible dysplasia and 69/132 OE images contained visible dysplasia. The proportion of images containing visible dysplasia was not significantly different between the WLE and OE groups ($p = \text{NS}$). The histology of the lesions assessed within our patient cohort are summarised in (**figure 36**).

Lesion characteristics		
Number of patients per histologic grade	NDBE	15
	LGD	2
	HGD	11
	M1-3 adenocarcinoma	12
	≥ SM1 adenocarcinoma	1

Figure 36 Summary of lesion histology for patients recruited to the iScan OE study

Dysplasia detection rates in expert and trainee endoscopists using iScan OE

The accuracy of dysplasia detection improved in all trainees from 63% using HD-WLE compared to 76% using OE (OR: 2.00, 95%CI: 1.61-2.49, $p=0.001$). Sensitivity improved in 6 out of 7 trainees. The pooled improvement in sensitivity rose from 71% with HD-WLE to 81% with OE (OR: 1.93, 95%CI: 1.33-2.81, $p < 0.001$). The use of OE also improved specificity in 6 of 7 trainees from 55% to 70% (OR: 2.12, 95%CI: 1.58-2.85, $p < 0.001$). PPV improved from 59% with HD-WLE to 75% with OE (OR: 2.07, CI 1.58-2.71, $p < 0.001$), as did NPV from 68% to 77% (OR: 1.60, 95%CI: 1.17-2.20, $p =$

0.004). Trainees made significantly more correct diagnoses overall using OE vs WLE ($p<0.001$) (**figure 37**).

Performance measure	HD-WLE	OE	Odds ratio (95% CI)	P value
Trainees				
Sensitivity	71% (309/434)	81% (379/469)	1.93 (1.33-2.81)	0.001
Specificity	55% (261/476)	70% (301/427)	2.12 (1.58-2.85)	<0.001
PPV	59% (309/524)	75% (379/505)	2.07 (1.58-2.71)	<0.001
NPV	68% (261/386)	77% (301/391)	1.60 (1.17-2.20)	0.004
Accuracy	63% (570/910)	76% (680/896)	2.00 (1.61-2.49)	<0.001

Figure 37 Trainee diagnostic performance using OE compared to HD-WLE for dysplasia detection

The accuracy of dysplasia detection improved in all experts when using OE compared to HD-WLE. We observed an improvement from 77% to 84% (OR: 1.74, 95%CI: 1.34-2.25, $p = <0.001$). Sensitivity of dysplasia detection improved in 6 of 7 experts from 67% with HD-WLE to 77% with OE (OR: 2.26, 95%CI: 1.55-3.29, $p = <0.001$). Specificity improved in 5 of 7 experts when using OE compared to WLE. Specificity increased from 86% to 92% (OR: 2.13, 95%CI: 1.34-3.39, $p = <0.001$). PPV and NPV improved with the use of OE. PPV improved from 81% to 91% (OR: 3.37, 95%CI: 1.51-3.73, $p = <0.001$) and NPV improved from 74% to 78% (OR: 1.27, 95%CI: 0.95-1.69, $p = 0.1$). As a group, experts made significantly more correct diagnoses using OE than WLE ($p = <0.001$). **(figure 38)** summarises the pooled diagnostic performance of the expert panel.

Performance measure	HD-WLE	OE	Odds ratio (95% CI)	P value
Experts				
Sensitivity	67% (291/434)	77% (360/469)	2.26 (1.55-3.29)	<0.001
Specificity	86% (407/476)	92% (393/427)	2.13 (1.34-3.39)	0.001
PPV	81% (291/360)	91% (360/394)	3.37 (1.51-3.73)	<0.001
NPV	74% (407/550)	78% (502/693)	1.27 (0.95-1.69)	0.1
Accuracy	77% (698/910)	84% (753/896)	1.74 (1.34-2.25)	<0.001

Figure 38 Expert diagnostic performance using OE compared to HD-WLE for dysplasia detection

Performance measure	HD-WLE	OE	Odds ratio (95% CI)	P value
All endoscopists				
Sensitivity	69% (600/868)	78% (739/938)	2.03 (1.57-2.63)	<0.001
Specificity	70% (407/476)	81% (393/427)	2.10 (1.64-2.70)	<0.001
PPV	68% (600/884)	82% (739/899)	2.14 (1.70-2.69)	<0.001
NPV	71% (668/936)	78% (694/893)	1.41 (1.13-1.74)	0.002
Accuracy	70% (1268/1820)	80% (1433/1792)	1.84 (1.56-2.18)	<0.001

Figure 39: Diagnostic performance for pooled endoscopists using HD-WLE compared to OE for Barrett's associated dysplasia detection

Assessing the MV classification system for lesion detection using OE-ME compared to HD-WLE

Phase two of our study aimed to explore the diagnostic performance of a previously validated and published mucosal and vascular classification for use with iScan and magnification endoscopy. We aimed to compare its performance with HD-WLE compared to with OE. We envisage that such a system for use with magnification endoscopy could potentially facilitate the delineation of resection margins when planning endoscopic resection therapy for early neoplastic lesions

In total, 63 HD-WLE and 90 OE still images of magnified mucosa from 54 patients were obtained. Where possible magnified mucosal images were matched between both imaging modalities and there was a non-significant difference in the proportion of images in each group containing dysplastic tissue (29/63 vs 49/90). All images obtained were correlated with histology as described previously.

Using the MV classification, the panel of expert endoscopists correctly classified tissue as dysplastic or non-dysplastic with 66.7% (95%CI: 62.7-70.8%) accuracy using HD-WLE. We observed a significant improvement to 79.9% (95%CI: 77.8-82%)

using iScan OE, the use of which enabled significantly more correct diagnoses to be made ($p < 0.001$). The sensitivity of dysplasia detection also improved from 83.4% (95%CI: 76.5-88.3%) using HD-WLE to 86.3% (95%CI: 81.5-91%) using OE. Specificity improved using OE; increasing from 53.6% (95%CI: 43.5-63.7%) in HD-WLE to 71.2% (95%CI: 67.5-74.8%) using iScan OE (**figure 42**).

We observed an improvement in interobserver agreement between experts attempting to classify ME images of Barrett's mucosa as dysplastic or non-dysplastic based on our described MV classification. Interobserver agreement was fair us HD-WLE at 0.30 (95%CI: 0.10-0.49). This improved to moderate at 0.53 when experts attempted to classify OE images (95%CI: 0.33-0.66). We also note that interobserver agreement improved with the use of OE to classify either mucosal features or vascular features in isolation as shown in (**figure 43**).

Observer	WLE accuracy (%)	OE accuracy (%)	WLE sensitivity (%)	OE sensitivity (%)	WLE specificity (%)	OE specificity (%)	WLE PPV	OE PPV	WLE NPV	OE NPV
1	69.8	75.6	79.3	75.6	61.8	73.0	63.9	79.2	77.8	71.1
2	65.6	77.9	70.0	79.6	61.7	75.7	63.9	81.3	70.0	73.7
3	68.3	79.1	83.0	81.6	54.6	75.7	62.5	81.6	78.3	75.7
4	65.0	79.1	82.8	91.8	50.0	62.2	58.5	76.3	77.3	85.2
5	76.2	83.7	79.3	89.8	73.5	73.7	71.9	81.5	80.7	84.9
6	63.5	81.4	96.6	91.8	35.3	67.6	56.0	79.0	92.3	86.2
7	58.7	82.6	85.7	91.8	38.2	70.3	53.3	80.4	76.5	86.7
Mean (±SD)	66.7 (±5)	79.9 (±2)	83.4 (±8)	86.3 (±6)	53.6 (±13)	71.2 (±2)	61.4 (±6)	79.9 (±2)	79.0 (±6)	80.5 (±7)

Figure 40 Diagnostic performance of expert endoscopists using magnification endoscopy and the MV classification to detect Barrett's associated neoplasia using HD-WLE and OE

	Overall assessment (NDBE v DBE) (95%CI)	M classification (95%CI)	V classification (95%CI)
WLE	0.30 (0.10-0.49)	0.33 (0.13-0.52)	0.38 (0.20-0.56)
OE	0.53 (0.34-0.70)	0.50 (0.33-0.66)	0.52 (0.35-0.68)

Figure 41 Interobserver agreement for dysplasia detection using a mucosal pattern assessment and vascular pattern assessment using HD-WLE compared to OE.

6.5 Discussion

The Seattle Protocol, a defined procedure for taking levelled, circumferential biopsies during Barrett's surveillance endoscopies has inherent deficiencies. These deficiencies include endoscopist error, and low total sampling area introduce a potential risk for early, mucosal oesophageal cancers to be missed during endoscopic assessment. Studies demonstrate that up to 36% of early lesions are not detected through endoscopic surveillance in the year preceding diagnosis¹⁴⁵. Since early mucosal lesions can be treated with good effect through endoscopic therapies, prior to them invading deeper into the oesophageal wall, it is vital that endoscopic technologies are targeted at reducing this miss rate. Advanced endoscopic imaging platforms may improve the early detection of such lesions. This study examined two main concepts, with the aim of investigating whether advanced endoscopic imaging platforms may mitigate some of these risks.

Firstly, we examined whether virtual chromoendoscopy with iScan OE improve dysplasia detection during the endoscopic assessment of BE. Importantly we have investigated the diagnostic benefit of iScan OE in both an expert and trainee cohort. Such an imaging system may improve visualisation of the oesophageal mucosa, without the need for additional chromoendoscopy agents such as acetic acid.

This study showed that when using iScan OE, there was a significant improvement in the diagnostic performance of a cohort of both trainee and expert endoscopists in the detection of Barrett's oesophagus associated neoplasia. We observed statistically significant improvements in trainee endoscopist's sensitivity, specificity, PPV and NPV for the detection of dysplasia. Overall there was a significant increase in their diagnostic accuracy from 63% with HD-WLE to 76% with OE. We reported a similar increment in an expert endoscopist cohort. There was a statistically significant improvement in diagnostic accuracy for dysplasia detection by experts from 77% to 84% when using HD-WLE and OE respectively.

Interestingly, we noted that although there was a significant difference between trainees and expert using HD-WLE, when they used optical

enhancement imaging their accuracy improved to a level comparable with those of expert endoscopists. In this way we could also assess whether OE has a role in the endoscopy training environment for trainees undertaking Barrett's lesion recognition training. Based on the results of this study, we infer that the use of this advanced imaging modality therefore shows promise if it were used within the training environment.

Secondly, a previously proposed magnification endoscopy classification system, designed for use with the iScan system, has been validated for use with the iScan OE platform by this study²²⁴. We report that when using magnification endoscopy in conjunction with the proposed MV classification system to assess lesions, OE offers a significant improvement in accuracy of dysplasia detection compared to HD-WLE alone. Our analysis demonstrates that in expert hands, diagnostic accuracy improved from 66.7% to 79.9% when using OE. Sensitivity improved from 83.4% to 86.3% with OE. Improvements were also seen in specificity, PPV and NPV. We also observed an improvement in interobserver agreement ($k=0.53$ vs 0.30) when endoscopists used OE compared to HD-WLE. Our classification system is robust for use with OE in this regard, both constituent features that we define as representative of dysplasia, 'mucosal' and 'vascular', demonstrated improved agreement when assessed with OE.

This study compares favourably to other published work in this field. Since other advanced endoscopic imaging systems have been investigated previously, most notably Narrow Band Imaging (NBI), we used these to benchmark the performance of OE. A large, well-designed trial validated a classification system based on abnormalities in mucosal and vascular patterns for use with NBI magnification endoscopy. Expert endoscopists using the BING classification allowed dysplasia to be detected with up to 85% accuracy, 80% sensitivity and specificity of 88%. The group undertaking this study also observed a PPV and NPV of 81% and 88% respectively²²¹. We noted that although the OE system has better sensitivity and comparable accuracy in this study, our specificity, PPV and NPV were lower.

Interestingly a follow up study by Nogales et al, which used a larger number of images for testing the BING criteria than the above study, demonstrated accuracies for

dysplasia detection of 81.1%²²⁵. These results are more comparable to our result of 79.9%, although larger studies to validate our work may be needed to make direct comparisons. In our study OE attained higher sensitivities than NBI (86.3% vs 48.4%), but lower specificity and NPV. Comparison of our results suggest that OE may improve the detection of dysplasia compared to NBI, but remains a less specific modality. This may be idiosyncratic and due to the diagnostic characteristics of the endoscopists included in this study. Assuming that OE does in fact allow greater sensitivity, but lower specificity, it is arguable that in lower volume centres, sensitive detection of early neoplasia is of paramount importance since it permits early access to referral to tertiary centres and to endoscopic eradication therapy.

Our group has previously published work which assessed the iScan 1,2 and 3 magnification endoscopy advanced imaging system, an early iteration from which OE was developed, for used in BO neoplasia detection. Compared to the improvement in diagnostic performance offered by iScan 1-3 versus HD-WLE alone, we suggest that OE may be a preferable modality. Lipman et al assessed only expert endoscopists, but reported accuracy, sensitivity and specificity of dysplasia detection using the MV classification of 76%, 91% and 57% respectively²²⁴. We report a comparatively better diagnostic performance; 79.9% accuracy, 86.3% sensitivity and 71.2% specificity. The accuracy of dysplasia detection on endoscopic pullthrough using OE was greatly improved at 85.6% compared to detection accuracy of 76% reported in the Lipman iScan 1 study, furthermore we demonstrate higher detection accuracy without the addition of acetic acid. This improved diagnostic performance without the need for the application of additional topical agents is favourable, since it could potential reduce endoscopy time, reduce the consumption of resources and discomfort for patients.

We envisage that a classification with sufficient accuracy, sensitivity and specificity could both improve neoplasia detection and rationalise how biopsies are taken during BE surveillance from the random biopsies of the Seattle protocol to a more targeted approach. Although the OE system does not meet the threshold for incorporation into routine Barrett's surveillance, these results demonstrate promise that with further

refinement of the endoscopic technology, endoscopist training and further investigation it may be demonstrate sufficient accuracy to do so. With further development and investigation, abnormal areas could be detected on withdrawal of the endoscope through the BE segment, with abnormal areas and potential resection margins interrogated further with magnification endoscopy. A more targeted approach could potentially reduce procedure times and streamline workflow in endoscopy suites and pathology departments.

Limitations of this study and potential further work

The primary limitation of this study is that it utilised still images rather than real time videos. This does not represent clinical practice but is consistent with other published works in this area. The use of still images enables easier and more direct comparison of clinician assessments, but clearly represents a more artificial and controlled situation than might be expected in clinical practice. Though we chose to utilise still images, we attempted mitigate the effect this may have had on the 'real world' applicability of the results. To recreate the slow endoscope withdrawal procedure performed in clinical practice during Barrett's surveillance, we used sequential still images taken throughout the BO segment. However, it is recommended that further studies assessing the improvement in diagnostic performance that the OE system offers, should focus on assessing dysplasia detection using videos. The prevalence of early neoplasia in our cohort introduces a potential bias, our cohort is an enriched population with around 50% of our subjects exhibiting early neoplasia. This is in line with other studies and logistically it would be difficult to achieve a sufficiently powered study with a cohort prevalence reflective of day-to-day practice. Anecdotally, the prevalence of dysplasia in patients referred to tertiary centres with suspected dysplasia is high. As such it may be that the results achieved in this study may accurately reflect the potential diagnostic benefit OE may offer in a tertiary setting. Further work should aim to assess the role of OE in a patient cohort with a prevalence of Barrett's associated dysplasia that is representative of the average patient population seen in secondary care (estimated at between 3-14% in patients with symptomatic GORD²²⁶, to as low as 1.2% in asymptomatic patients¹¹³).

We would acknowledge that logistically it would be very difficult to adequately power such a study.

The baseline performance required for the incorporation of new endoscopic technologies is defined by Preservation and Incorporation of Endoscopic Innovations (PIVI) guidelines. Currently the OE system does not reach the thresholds defined by PIVI for the incorporation of new technologies into endoscopy²²⁷. The recommended PIVI thresholds for the adoption of a new endoscopic technology are defined as sensitivity, NPV and specificity for dysplasia detection of >90%, >98% and >80% respectively. We demonstrate that the sensitivity of dysplasia detection using OE approaches this for magnification endoscopy (86.3%), but the technology is not specific enough. Of note, other advanced endoscopic imaging technologies such as NBI have also not consistently exceeded PIVI thresholds, nor does HD-WLE in this study. We therefore cannot suggest that OE should routinely replace HD-WLE for use in BE surveillance. However, given these promising results, particularly those demonstrated in a trainee cohort, it may serve as a useful adjunct to improve the early detection of neoplastic tissue in these endoscopists; further studies should be conducted to evaluate this.

Another limitation of this study is that all of the involved endoscopists, both experts and trainees, practice within tertiary referral centres for endoscopy and so potentially may have more expertise in the assessment and management of early BO associated neoplasia than their equivalently trained clinicians practicing in a more general setting. In future work, the utility of OE and our MV classification system should therefore be assessed in a wider range of settings by clinicians with more varied experience. A significant limitation of this analysis was that it did not assess the diagnostic performance of consultant endoscopists who practice in centres treating a lower volume of Barrett's oesophagus and associated neoplasia. While speculative it may be that their diagnostic performance falls somewhere between trainees and consultant endoscopists practising in high volume tertiary centres. It would be valuable to assess whether OE confers an advantage over HD-WLE in these settings as

this is likely where the majority of Barrett's surveillance is being conducted in a low prevalence population.

6.6 Summary of this chapter

This study investigated the role of a new advanced endoscopic imaging system, iScan OE, for the detection of Barrett's oesophagus associated dysplasia by both an expert and trainee endoscopist cohort.

1. We demonstrate that iScan OE significantly improves the detection of Barrett's associated dysplasia on endoscopic pullthrough compared to HD-WLE alone. This improvement was observed in both trainee and expert endoscopists.
2. We have assessed whether a previously published²²⁴ consensus driven magnification endoscopy classification system for use with OE can improve diagnostic accuracy for expert endoscopists assessing Barrett's associated dysplasia using magnification endoscopy.
3. iScan OE does not meet PIVI thresholds and so at present, while a promising technology for further refinement and investigation, it should not be routinely used for Barrett's surveillance over HD-WLE.

Chapter 7– Developing a clinically interpretable convolutional neural network to aid in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof of concept study.

The work presented in this chapter formed the basis of a peer reviewed publication. Text and figures were adapted for publication. Citation:

Everson, M., Herrera, L.G.P., Li, W., Luengo, I.M., Ahmad, O., Banks, M., Magee, C., Alzoubaidi, D., Hsu, H.M., Graham, D. and Vercauteren, T., 2019. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. *United European gastroenterology journal*, 7(2), pp.297-306.

This study was a collaborative work with Dr Luis Garcia Peraza Herrera, who designed, developed and coded the CNN architecture. Clinical requirements, conceptualisation and software requirements were developed by ME

7.1 Introduction

Gastroscopy remains the investigation of choice for the diagnosis of early squamous cell neoplasia (ESCN). Typically these early lesions are flat with a subtle appearance on endoscopic assessment and so may be easily missed by clinicians. Previous work shows a significant miss rate for UGI cancers on endoscopies undertaken within three years of diagnosis¹⁹⁹. There are also geographic variations in the incidence of oesophageal squamous neoplasia; with a far higher incidence across the so-called ‘cancer belt’ comprising the Middle East, Central Asia and extending into China^{13,41,228}. Clinicians working in these regions are more likely to see a higher number of such lesions in their practice and so are more likely to be familiar with the endoscopic appearance of ESCN compared to physicians practicing in lower risk lesions. Regardless of clinician experience, the early detection and accurate characterisation of ESCN lesions is essential; in order to predict lesion histology and guide the most appropriate intervention. In this chapter we propose a novel methodology for the development of a convolutional neural network to assist both experienced and inexperienced endoscopists recognise ESCN lesions endoscopically.

ESCN lesions confined to the mucosa have been demonstrated to have low rates of local lymph node metastasis (<2%). This compares favourably to lesions which invade

the submucosa where the rate of distant spread increases dramatically (8-45.9%). This is significant since lesions confined to the mucosal layer are amenable to endoscopic therapy^{229–233}. Endoscopic eradication therapy of ESCN, is largely delivered by either endoscopic mucosal resection (EMR) or submucosal dissection (ESD). Endoscopic eradication therapy is associated with excellent rates of 5 year survival and spares patients the attendant morbidity and mortality of oesophagectomy after a late diagnosis. The potential efficacy of endoscopic eradication therapy in the treatment of ESCN, means it is highly important that clinicians undertaking upper gastrointestinal endoscopies are able to recognise early lesions, particularly in low volume clinical settings where the subtly abnormal appearance of ESCN could be easily missed.

As outlined in **chapter 2**, endoscopically visible markers aid endoscopists in identifying ESCN lesions during assessment. Intrapapillary capillary loops (IPCLs) are oesophageal microvessels, that were first characterised using magnification endoscopy^{62,63,234}. IPCLs are a validated and well characterised endoscopic marker of ESCN. Architectural changes which occur in the tissue of the oesophagus as an ESCN progressive invades through the layers correlates with stereotyped and sequential changes in the morphology of the visualised IPCLs^{69,194,235}.

Under normal conditions IPCLs arise from the submucosal vessels and run adjacent to the basal layer of the oesophageal epithelium. They are visualised as fine calibre looped structures which branch from the deeper, submucosal vessels on magnified endoscopy. As an ESCN progresses there is stepwise destruction of the architecture of the oesophageal wall, beginning in the epithelium and mucosa and gradually extending into the submucosa and beyond with time. In early ESCN or high-grade dysplastic lesions confined to the mucosa, IPCLs become more tortuous, dilated and more readily appreciated on magnification endoscopy assessment of the mucosa. As the mucosal layer is further interrupted as the ESCN invades through the mucosal layer, the IPCLs lose their looped structure and form linear dilated vessels of increasing calibre. The extent and severity of these morphologic changes is more pronounced the deeper a lesion invades. Studies have also noted that as part of this process, avascular areas (AVAs) become visible between these non-looped vessels – the presence of these AVAs

also correspond to ESCN lesions that have invaded the oesophageal wall deep to the mucosal layer²³⁶. In established ESCN lesions, which are often visible on normal endoscopy due to disruption to the oesophageal mucosal contour, the IPCLs are almost completely obliterated. In their place neovascularisation occurs with the formation of tortuous, grossly dilated and non-looped vessels – this typically represents lesion advance into the submucosal layer and beyond^{194,237–239,240}.

Advanced endoscopic imaging modalities, such as narrow band imaging (NBI) in combination with magnification endoscopy, afford improved visualisation of subtle microvascular patterns in the oesophageal mucosa of patients with ESCN²⁴¹. Although several classifications have been proposed to define the abnormal ICPL morphologies that correlate with histological invasion depth of ESCN lesions we have elected to use the Japanese Endoscopic Society (JES) classification for this study.

Although other validated classification systems exist, as discussed above, we believe that the proposed Japanese Endoscopic Society (JES) IPCL classification is a simplified system, allowing easy recognition of ESCN by both endoscopists experienced and inexperienced in the assessment of ESCN^{194,236}. Since the JES classification is widely accepted and utilised in routine clinical practice in areas of high prevalence such as China and Japan, we are further encouraged to use it as the classification system for this work. We are conducting this study in a Western setting and ultimately aim to develop a convolution neural network designed to aid less experienced endoscopists identify ESCN, we selected the JES classification for its ease of interpretability, fewer subdivisions of IPCL patterns and good correlation with histologic findings²³⁶. Each IPCL subgroup identified in the JES classification corresponds with high accuracy to a given histological grade and invasion depth of ESCN. In common with other classifications, IPCL patterns with increasing irregularity and vessel dilation represent more advanced, invasive disease²³⁶ (**figures 44 and 45**). The accuracy of the JES classification as reported by Oyama et al is high compared to other classifications – with the overall accuracy for histology prediction 90.5% across type B1-3. Overall accuracies for histology prediction were 91.9%, 93.4% and 95.9% for type B1, B2 and B3 IPCL patterns respectively²³⁶. Kim et al also report excellent interobserver

agreement using the JES classification for the prediction of histology in based on IPCL patterns²⁴².

In **chapter 2**, IPCL classifications advanced by Inoue et al.^{62,63} and Arima et al.²³⁹ were discussed; both of these classifications also associate progressive morphological abnormalities in IPCLs with deeper invasion of an ESCN, but we suggest they are complex and require a higher degree of interpretation by endoscopists, particularly those unfamiliar with the endoscopic assessment of ESCN

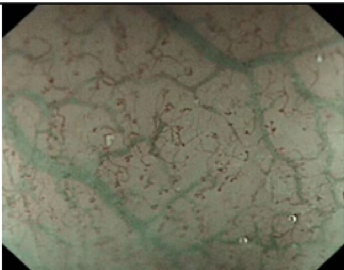

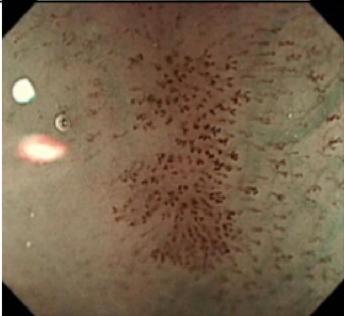

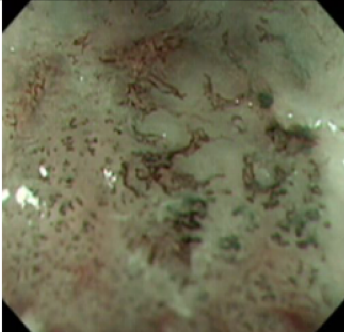

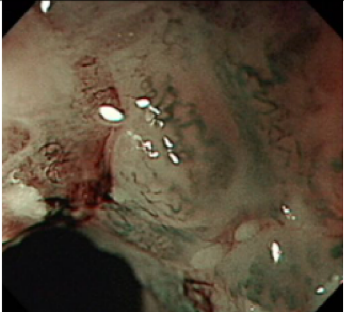

JES type		Typical morphology	Typical histology
A		Submucosal vessels visible with small regular obliquely branched IPCLs 	Normal/LGD
B1		Mild dilation of IPCLs with increased tortuosity but preservation of normal loop structure 	HGD/LP
B2		Gross dilation of IPCL patterns. Increased tortuosity with destruction of some to form dilated linear vessels interspersed with avascular areas (AVAs). 	MM/SM1
B3		Complete loss of normal IPCL morphology. Vessels up to 3x dilation of type B2 IPCLs. Mucosal contour may be abnormal with evidence of neovascularisation. 	SM2 or deeper

Figure 42 Representative examples of the different IPCL morphologies and the associated ESCN invasion depths as classified by the Japanese Endoscopic Society^{69,194}

IPCL classification	Neoplastic	Typical histology	Amenable to EET?
A	No	Normal/LGD	Yes
B1	Yes	HGD/LP	Yes
B2	Yes	MM/SM1	Possibly
B3	Yes	SM2 or deeper	No

Figure 43 How the findings of particular IPCL morphologies at endoscopy correlate with the presence of neoplasia and whether endoscopic eradication therapy is indicated. Adapted from²⁴³

As outlined above, due to geographic disparities in the incidence of ESCN, some clinicians, particularly those in the Western world, may be less experienced in recognising the endoscopic features of ESCN. We proposed in this study that computer-aided diagnosis (CAD), through the use of convolutional neural networks (CNNs) could provide a useful adjunct to endoscopists assessing ESCN lesions. As will be discussed in more detail in this chapter, CNNs often rely on ground truth data; that is to say data which is pre-labelled according to the current gold standard. We propose that the distinct patterns observed in IPCL morphology could provide the input data to train a CNN to classify the oesophageal squamous mucosa as normal or abnormal. In this study the gold standard was derived from images extracted from endoscopic videos of an ESCN assessment; the IPCL patterns observed in these images and videos were then correlated with the assessment of an expert panel of endoscopists and the histology obtained from tissue sampling at the time of the endoscopy. Once this ground truth was established the labelled images were used to train a network to accurately classify previously unseen images as neoplastic or normal.

A CNN developed for the purpose of accurately classifying squamous lesions, based on visualised IPCL patterns intra-endoscopy, could have a number of potential benefits. In this study we attempt to provide proof of concept that IPCLs can be used as endoscopic markers for a CNN to predict dysplasia, we envisage that with further refinement this CNN could be used as an adjunct for endoscopists to improve their lesion recognition, characterisation and prediction of ESCN invasion depth. Such a system could achieve this in a number of ways. Firstly, it could serve as a red flag system; alerting clinicians to the presence of abnormal IPCL patterns that may represent dysplasia and warrant a biopsy – thereby improving the pick-up rate of dysplastic lesions. Conversely, if clinically validated, it could reassure clinicians that suspect areas of mucosa are likely normal and thereby reduce the need for excessive biopsies. This could serve to streamline endoscopic assessments and could potentially reduce procedure times and conserve resources by facilitating better targeting of biopsies. Secondly, a downstream application of such a system could have a role as a decision support tool for endoscopists to triage lesions as either amenable to EET or

not, allowing patients with those lesions that cannot be managed endoscopically to be more promptly referred for surgical intervention or chemotherapy. These later applications of such a system should be regarded with caution at present, since for a CNN designed for this purpose to be used in clinical practice it would require much greater validation; this study currently serves only as a proof of concept.

7.2 Aims of this study

At the time this study was conducted, the use of convolutional neural networks to characterise early squamous cell neoplasia of the oesophagus was entirely novel. IPCL morphologies had been used by endoscopists to assess ESCN but had not been used as the input data for the development of CNNs. The aim of this chapter is broadly to provide proof of concept, a baseline methodology and a benchmark for the diagnostic performance of a CNN for the prediction of histology in ESCN, based on observed IPCL patterns. The specific aims are:

1. To develop a novel AI system that can classify squamous oesophageal mucosa as normal or abnormal in **endoscopically resectable lesions** (<SM1 invasion, **figure 44 and 45**) in real time on videos acquired during magnification endoscopy
2. To develop a methodology that can be used to develop further AI systems, for use in vivo, that accurately predict the grade of a lesions histology and invasion depth based on IPCL patterns

7.3 Methods

Patient recruitment, inclusion and exclusion criteria

Patients attending for upper gastrointestinal endoscopy at one of two high volume referral centres for ESCN in Taiwan [(National Taiwan University Hospital and -Da Hospital/I-Shou University) were recruited with consent. Both patients presenting incidentally with a macroscopically normal oesophagus and those attending for assessment of a suspected ESCN were recruited.

Patients with a suspected finding of ESCN were only included if pathological samples (EMR, ESD, oesophagectomy specimens) were acquired at the time of their endoscopy or subsequent surgery. Patients with active oesophageal ulceration were excluded, as were those who had had previous oesophageal surgery or ablative therapy. Our study complied with the Declaration of Helsinki. The Institutional Review Board of the E-Da Hospital approved this study (IRB number: EMRP-097-022).

Endoscopic procedures and video acquisition

Gastrosopies were performed by two expert endoscopists (WLW, HPW) at the referral centres described above; an expert endoscopist was defined as a gastroenterologist who had completed all formal training and currently performed >50 ESCN assessments per year.

All endoscopies were performed using a high definition ME-NBI GIF-H260Z endoscope, capable of imaging the oesophagus with both narrow band imaging and magnification endoscopy simultaneously. An Olympus Lucera CV-290 processor (Olympus, Tokyo, Japan) was used. All endoscopic procedures were recorded in high definition.

After giving consent, patients underwent a full upper gastrointestinal endoscopy with either topical anaesthesia (oral xylocaine) or conscious sedation with intravenous sedatives. The oesophageal mucosa was cleaned with a solution of Simethicone to

remove food residue, mucus and bubbles, prior to the interrogation of the IPCL pattern using magnification endoscopy with narrow band imaging (ME-NBI). Magnification endoscopy was performed on areas of interest at between 80-100x magnification. After assessment of the lesion pathological samples of the imaged area were acquired through endoscopic mucosal resection or endoscopic submucosal dissection. Histologic samples were formalin fixed, in line with routine practice and analysis was undertaken by two expert gastrointestinal pathologists, All histologic samples were reported according to the Vienna classification system²⁴⁴.

Labelling ME-NBI videos and establishing an expert consensus

High-definition videos of each endoscopic procedure were reviewed independently by three expert upper gastrointestinal endoscopists (WLW, RJH and HPW). All reviewing endoscopists had extensive experience in the endoscopic assessment and treatment of ESCN lesions. For each video, the visualised IPCL patterns viewed on magnification endoscopy were classified by consensus based on the JES classification system. For videos where more than one IPCL subtype was visible the worst case IPCL pattern was used to classify the video. The visual IPCL classifications were then correlated with histology taken from the imaged area before a final, concordant classification was assigned to each group of images. Type A IPCLs were considered normal, type B1, 2 and 3 IPCLs were considered abnormal and indicative of the previously described grade of neoplasia.

In order to generate training images for the neural network, frames were extracted from each endoscopic video. Frames were sampled from the recordings at 2 frames per second (fps), and stored in the lossless .png format as sequential still images.

Rationale for selecting images as a data input

For an AI system to be effective at characterising squamous lesions based on IPCL patterns in-vivo, it needs to operate at video rate. A video is simply a rapid sequence of images and so it is acceptable to break videos down into individual frames as a data input upon which to train a CNN. By fragmenting the videos into still images, more data upon which the network can train is generated and the images can be quality

controlled as outlined below. This is a generally accepted process for training AI networks for video classification. The selection of 30fps as the frame rate at which images were sampled from videos was somewhat arbitrary. It was agreed amongst study members that to sample images at higher frequency would give rise to a very large number of very similar images; since all included images in this study were manually quality controlled by a study member, there needed to be a trade-off between sampling a high number of images to improve training, while not sampling so many that they could not be processed in a feasible amount of time. As long as our system is able to classify IPCL patterns accurately at video rate, it will have potential for the real-time identification of ESCN in vivo and so the relevance of the number of images sampled for training is less important.

Image quality control

As images were sampled at a rate of 30fps, they were of variable quality, with some more informative than others. Prior to inclusion in the training datasets, images were manually quality controlled by an experienced clinician study member (MAE). As is the case in all endoscopic assessment some images are not of sufficient quality for an endoscopist to make an adequate characterisation of the IPCL patterns. Those images that were blurred, contained no visible IPCLs or where the mucosa was obscured by blood or mucus were deemed uninformative and removed from both the training and testing datasets. After quality control all images were cropped to remove all black borders and any identifiable or discriminative patient or clinician demography labels. This prevented the CNN identifying discriminative features on the images to distinguish between normal and abnormal cases that were not related to the IPCL patterns. This additional step in the method was included after preliminary work demonstrated that the black border around the endoscopic image was by chance larger in more patients with neoplastic lesions (due to a display setting on the recording device). The CNN then identified a larger black border as indicative of neoplasia. Once the borders were removed from all images, the informative feature in the image that enabled classification were the IPCLs.

Formation of image datasets

After sampling and quality control the full image dataset used in this study consisted of 7046 images, with resolutions ranging from 458x308 to 696x308 pixels. As a proof of concept, our study uses a large number of images from a relatively small number of patients.

We therefore employed 5-fold cross validation to generate five distinct datasets with different combinations of images (**figure 46**). The use of a 5-fold cross-validation method served two purposes in this study. The first was to maximise the number of images that could be used for training, validation and testing. To achieve this, each fold of images was subdivided into a training set of images, a validation set and a testing set. Images included in the training set were used to train the CNN to recognise IPCL features that were particularly represented in either normal or neoplastic tissues. Images in the validation set were used to refine network hyperparameters prior to testing. Images in the testing set were unseen in either the training or validation set of images for that fold and so when the network classified them as normal or abnormal, the process and result could be regarded as analogous to how it would perform in images from a new patient.

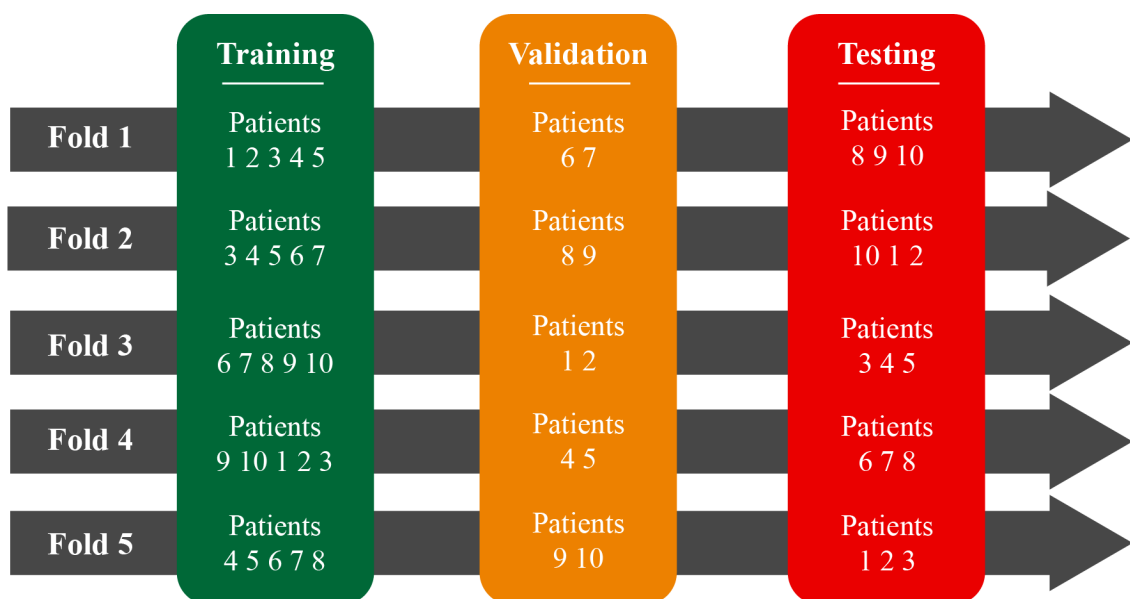


Figure 44 Schematic representation of how patients were assigned to allow five-fold cross validation with training, validation and testing image datasets in this study

For any given fold, images of patients used within the training dataset were not used in the validation or testing dataset, since the CNN would simply classify based on the ground truth label applied to that patient's images during training. On average, each fold used 3962 images for training, and 1637 unseen images (846 normal and 791 abnormal) for validation and testing. The composition of images in each fold used in this study is summarised in *(figures 47 and 48)* .

Fold	Training	Validation	Testing
1	2620	201	577
2	1792	891	715
3	1822	685	891
4	1792	715	891
5	1559	685	1154
Average	1917	635	646

Figure 45 Number of frames in each fold used for cross validation containing normal IPCL patterns

Fold	Training	Validation	Testing
1	2803	258	587
2	2205	739	704
3	1549	1360	739
4	1912	961	775
5	1754	743	1151
Average	2045	812	791

Figure 46 Number of frames in each fold used for cross validation containing abnormal IPCL patterns

Secondly each fold used in this study essentially generates a distinct neural network, with a different diagnostic performance and feature recognition abilities for the classification of abnormal mucosa. As a result some folds are expected to perform better than others to some extent. As we had developed multiple networks, with the same overall architecture simultaneously, we were able to conclude that the overall architecture of our network was robust since relatively similar results were achieved in each fold of data used to train and test the CNN. This refutes the notion that our CNN was performing well by 'fluke' on one particular combination of patient's images, with no indication of how it would perform with different combinations of patients. Clearly it will be required in further studies to assess whether this performance

stability is maintained once higher numbers of patients and images are used for training and testing the CNN.

Convolutional neural network design

The design of our convolutional neural network was a collaborative effort between our group consisting of clinicians and biomedical engineers specialising in interventional image computing. More detailed technical work on the architecture of this neural network is beyond the scope of this thesis, but forms a constituent chapter in another doctoral thesis by one of the study members (LCGPH). A full description of the technical methodology utilised for this proof of concept study has been previously reported by our group²⁴⁵ in an open preprint repository (Arxiv); a summary of the workflow and network architecture is provided in **(figure 49)**.

Quality controlled ME-NBI images of ESCNs were used as the input data. Each image was labelled according to IPCL subtype as assessed by the expert panel and correlated with known histology. In training, input images are passed through layers of the CNN, which develops filters for features on the image, at various resolutions. As images pass through the layers of the CNN, the network develops filters to detect various visual features such as edges and borders, colours and the shape of IPCL patterns. In our CNN images were passed through five such layers, with a batch size of 1, momentum of 0.9 and a fixed learning rate of 1^{-6} . The optimiser was stochastic gradient descent. All folds were trained with 4x the number of images in the training set as the maximum number of iterations. Global average pooling (GAP) was used after the decoder, prior to classification. Once training is complete new, unseen images are provided to the CNN and its classification performance is assessed. To enable the network output to be clinically interpretable, explicit class activation maps (eCAMs) were produced. As discussed in the results section, eCAMs are used visually display which discriminative features in each image the CNN utilises to classify them as normal or abnormal.

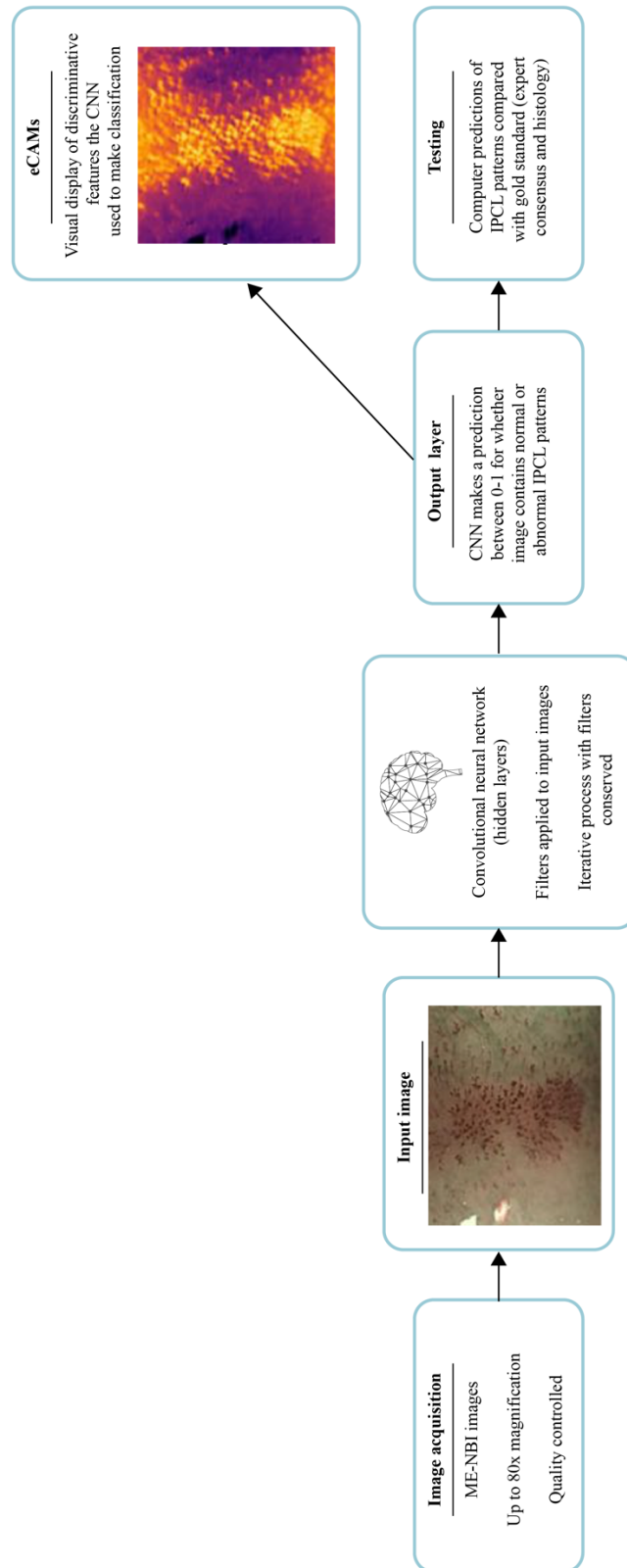


Figure 47 Schematic representation of the workflow and convolutional neural network architecture used in this study

Statistical analysis

Accuracy, F1 scores (a weighted average of precision and sensitivity, used to assess diagnostic accuracy in binary classification problems), sensitivity and specificity for the identification of normal or abnormal squamous mucosa based on the visualised IPCL patterns were calculated. Power calculations were not deemed necessary for this study since it is entirely novel and a proof of concept work. We are attempting only to benchmark the currently achievable diagnostic performance of AI systems in the classification of IPCL patterns.

7.4 Results

Patient characteristics

17 patients were included in our dataset; 10 with histologically proven ESCN and 7 with a normal squamous oesophagus. All included patients with dysplasia or ESCN were deemed to have endoscopically resectable lesions (defined in this study as a lesion invading no deeper than SM1), a summary of the histological invasion depth given in (**figure 50**).

ESCN characteristic		
IPCL patterns	Type A	7
	Type B1	5
	Type B2	5
Histology	Normal	7
	HGIN (M1)	1
	Lamina propria (M2)	4
	Muscularis mucosa (M3)	4
	Submucosa (SM1)	1

Figure 48 Summary of demographics and lesion information for patients recruited

Explicit class activation maps (eCAMs) provide a clinically interpretable neural network output

For a CNN to have clinical utility, it must be clinically interpretable. Our CNN generated class activation maps to visually demonstrate which image features CNN was using to discriminate healthy and unhealthy tissue. The distribution of abnormal IPCL patterns and image artifacts like bubbles, glare and mucous are heterogeneous between images. Similarly, within lesions there could be heterogenous IPCL morphologies visible, since our images were classified based on the highest grade of IPCL morphologies seen. The CAMs generated by our CNN enabled us to check that it was using the visualised IPCL patterns as a discriminant feature when making its classifications, which in turn could then be visualised by clinicians. In doing so, CAMs

provide a clinically interpretable network output. They also served as a validation step to ensure that the CNN was using the IPCL patterns in order to classify images as normal or abnormal, rather than other subtle discriminative features.

Representative images of the CAMs generated by our CNN are displayed in **(figure 51)**. Visual assessment of the generated CAMs provided insights into the decision-making process of our CNN. Firstly, we could confirm that our CNN was utilising visualised IPCL patterns as its most discriminative feature when classifying. We also noted with interest that it was able to accurately distinguish discrete areas of abnormal IPCL patterns that were found within otherwise normal mucosa, containing normal IPCL patterns. Interestingly, our CNN did not appear to use avascular areas that were visible between the abnormal B2 IPCLs seen in **(figure 51)**. Despite not using this feature it remained able to discriminate the abnormal vessels highly selectively and still make correct classifications. Our CNN appeared to show good discrimination of image artifacts such as specular reflections, glare, bubbles and blood as uninformative. We did however note that some images which contained healthy tissue and only type A IPCLs were incorrectly classified as abnormal. On further interrogation it appeared to be caused by the CNNs tendency to classify the deep oesophageal submucosal vessels as abnormal. The clinical consequence of such over diagnosis would be false positive classifications of normal tissue as abnormal.

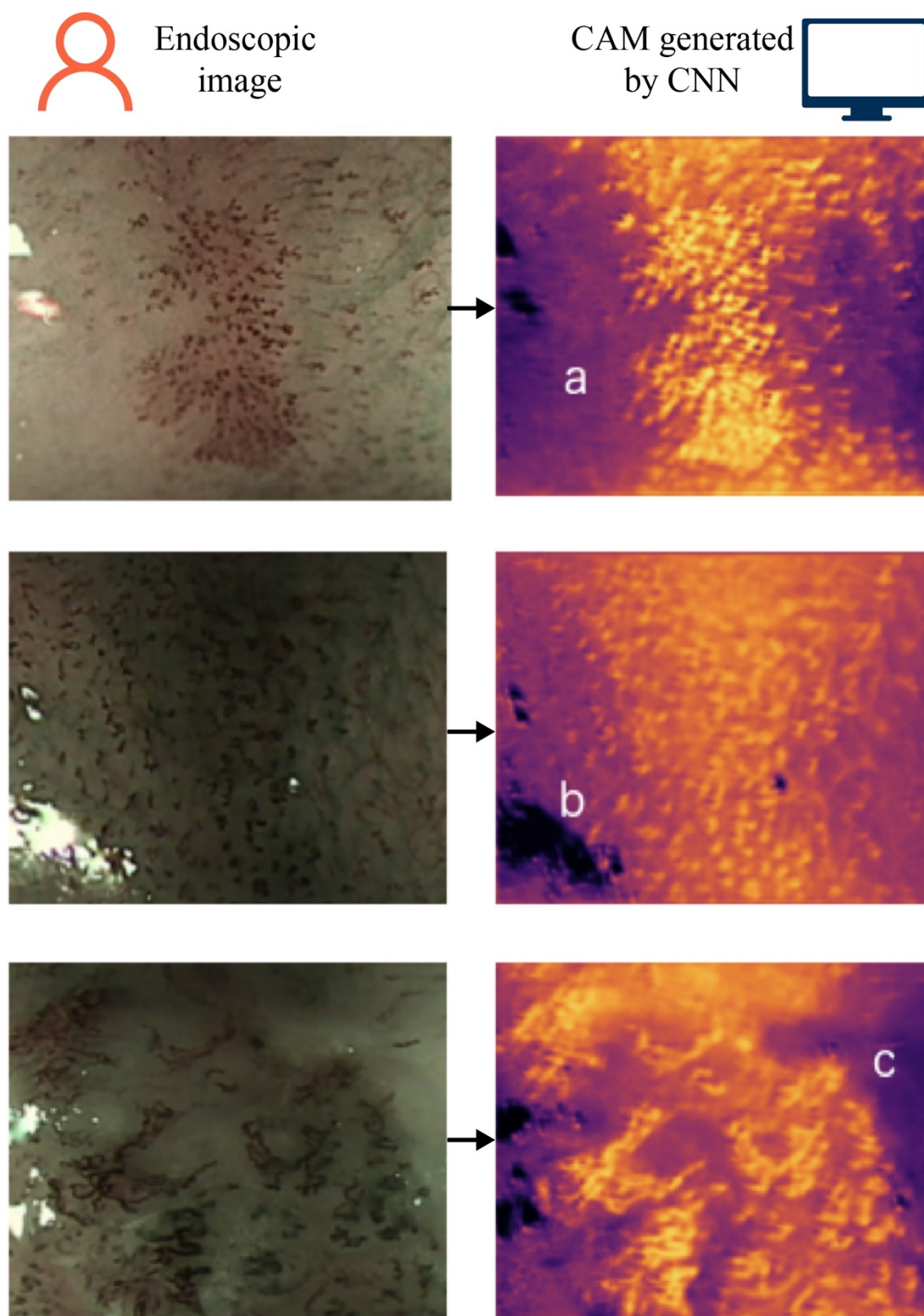


Figure 49 Input images (left column) with corresponding eCAMs (right column), illustrating visual features recognised by the CNN when classifying images. a) recognition of abnormal IPCLs patterns. b) specular reflections are ignored by the CNN c) high selectivity between normal mucosa and abnormal IPCLs.

CNN performance for IPCL classification

Our CNN operates at video rate, and is capable of classifying sequential HD-images that were acquired from fragmented endoscopic videos. The classification interval time varied according to the size of the images analysed, but ranged from 26.17ms to 37.48ms. Given these classification times it would be feasible for our CNN to be used in real-time endoscopic classification.

We demonstrate a mean accuracy for the differentiation of normal compared to abnormal mucosal areas based on IPCL patterns (type A vs B1/B2/B3) of 93.3% (range 86.2-98.3%). The average F1 score (a measure of accuracy for binary classification CNNs) for identifying ESCN based on the IPCL pattern was similarly high at 92.7% (range 85.4-98.2%). Our algorithm achieved a sensitivity and specificity of 89.7% (range 78.1-100%) and 96.9% (range 92-99.7%) respectively.

We noted some variability in the performance statistics of our network between folds. Within any fold there will be images that are simpler to classify as normal or neoplastic by a clinician, hence the same applies to the CNN. Our results showed some variability in the diagnostic performance of our CNN between folds. This is an indicator that suggested that it may struggle to generalise to classify the variations in IPCL morphologies that represent abnormal tissue (Type B1/2/3). Although this was the largest reported image dataset used for this purpose in the published literature at the time, we acknowledge that this finding suggests that more training images are required to train the network to be able to classify mucosa based on the full spectrum of variability in IPCL patterns. Our CNN performance statistics are summarised in (**figure 52**).

Fold	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 score (%)
1	86.2	80.4	92	85.4
2	89.0	78.1	99.7	87.6
3	97.7	100	95.9	97.6
4	98.3	99.4	97.3	98.2
5	95.1	90.6	99.6	94.9
Average	93.3	89.7	96.9	92.7

Figure 50 Summary of CNN performance statistics for detection of abnormal squamous mucosa based on IPCL patterns

7.5 Discussion

This study introduced the first application of computer aided diagnosis and artificial intelligence for the accurate, real-time characterisation of ESCN lesions, through the identification of intrapapillary capillary loops. In this study IPCL patterns were identified on magnification endoscopy combined with narrow band imaging, and classified based on the widely used Japanese Endoscopic Society classification system.

As outlined in the introduction to this chapter, the accurate assessment and characterisation of ESCN lesions is vital, in order to predict histology and invasion depth. Mucosal lesions have low rates of lymph node metastasis compared to lesions invading the submucosa and so are amenable to minimal invasive endoscopic therapy^{229–232}. If delivered promptly, prior to distant spread of squamous lesions, endoscopic eradication therapies facilitate resection with 5 year survivals of 75–100%²⁴⁶.

Prior to undertaking this study, we needed to select an IPCL classification system that would lend itself to developing a neural network. To allow this we identified that we would need a system that offered high diagnostic accuracy and that was not too complex or subdivided, since the more classes that a CNN is required to subdivide, the larger the volume of data required for each class. The rationale for selecting the Japanese Endoscopic Society classification is outlined below:

In **chapter 2.7** the validated IPCL classification systems used in current clinical practice were outlined. These systems were developed to standardise reporting and to assist clinicians with the accurate characterisation of IPCL patterns, with the aim of improving ESCN lesion recognition. Inoue et al. first proposed a five-part classification of IPCL patterns⁶⁴. In summary progressive morphologic changes in IPCL patterns correspond with progressive ESCN invasion through the oesophageal wall. Type I-III IPCL patterns are associated with normal mucosa, inflammation or LGIN. Type IV IPCLs were suggestive of HGD. Type V IPCL patterns were subdivided into V₁ (M1 carcinoma in situ), V₂ (M2 carcinoma in situ), V₃ (M3 or early SM1 invasion)

and V₄ (invasion to at least SM2) based on the progression of abnormal IPCL morphology. In one study, the Inoue classification enabled identification of early mucosal lesions (M1-2) with 89.5% sensitivity²³⁵. One of the potential weaknesses of the Inoue classification was its comparably reduced sensitivity for prediction of ESCN invasion into SM1 and SM2; which was reported at 58.7% and 55.8% respectively²³⁵. It is important to distinguish between mucosal and submucosal lesions since the most appropriate endoscopic treatment options and likelihood of distant spread differs significantly. Arima et al. subsequently proposed a four-part classification system with type 1, 2, 3 and 4 IPCLs representing normal mucosa, inflammatory changes, M1-2 carcinoma in situ and >M3 invasion respectively²³⁹. When deciding which classification system to adopt for use with our image dataset and CNN training, we felt that the multiple subdivisions of the Inoue classification are unnecessarily complex for a non-expert audience and could have reduced our ability to acquire enough data to adequately populate the datasets for this study. We also had to consider downstream applications of our CAD system; including that it would be needed to assist endoscopists less experienced in the assessment of ESCN to make clinical decisions. We therefore selected the JES classification, since it is now widely adopted in high prevalence Asian settings, but also offers ease of interpretability more suited to a less experienced audience.

The JES IPCL classification is the most contemporary and accurate classification system that has been proposed and is now widely used in areas of high ESCN prevalence. Furthermore its diagnostic accuracy and sensitivity persists even when classifying more advanced lesions that invade beyond the mucosa²³⁶. We believe that a CNN that functions using this classification will remain clinically relevant as it becomes more widely used. From a Western perspective, the JES classification represents a simplified classification system compared to the systems proposed by Inoue or Arima – with less interpretation required of subtle differences in IPCL morphologies, particular in more invasive ESCN lesions. We suggest that for this reason, this system is the most likely to be adopted and utilised in a Western setting for this reason and so our CNN is ‘future proofed’ by using it as a basis for classification.

This study reports, to our knowledge, the first use of a convolutional neural network for the purpose of classifying oesophageal squamous mucosa as normal or dysplastic, using IPCL patterns as a discriminant feature. We provided proof of concept and reported a novel methodology which can serve as the benchmark for further work aiming to develop CNNs for this purpose. This neural network uses sequential still images captured from HD videos of endoscopic examinations, to train a CNN to characterise ESCN lesions based on the IPCL patterns visualised at endoscopic assessment. This CAD platform achieved an average accuracy of 93.7% for the classification of dysplastic mucosa based on the appearance of abnormal IPCL patterns. The sensitivity and specificity of our CNN were 89.7% and 96.9% respectively. The accuracy of our network compares very favourably with results reported in the Oyama study, which claimed an overall pooled accuracy for histology prediction of 90.5% across neoplastic IPCL subtypes (type B1-3). Overall accuracies for histology prediction were 91.9%, 93.4% and 95.9% for type B1, B2 and B3 IPCL patterns respectively²³⁶. Since our system provided a binary classification of normal or neoplastic, it is difficult to make direct comparisons of our systems diagnostic performance with other studies, which typically report accuracies per IPCL subgroup, but we note that our results are similar for a normal compared to neoplastic distinction.

Furthermore, our CNN can operate at video rate, with rapid prediction times meaning that if clinically validated and successfully integrated, it could be used for real-time in vivo classification of endoscopic videos of possible ESCN.

Wang et al. demonstrate that in a non-expert panel of endoscopists accuracy of histology prediction using the JES classification ranged from 48-57% after a short training program²⁴⁷. Kim et al (2017) also demonstrated an overall accuracy of 78.6% for the correct prediction of histology²⁴². We note with interest that the diagnostic performance of the JES IPCL classification system is lower when used in a non-expert cohort of endoscopists. Since most of the validation studies of the various IPCL classifications are performed by expert endoscopists in high volume centres, we suggest that the true diagnostic performance by endoscopists outside of these

settings may fall short of this. We therefore conclude that with further development our system could yield an additional diagnostic benefit for non-expert endoscopists or in centres where the endoscopic assessment of ESCN is less common and so clinicians may feel less confident in recognising and classifying IPCL patterns. This is particularly relevant for endoscopists practising in areas of the world with a low prevalence of ESCN, where miss-rates for oesophageal squamous cancer are high.

Limitations of this study and potential further work

While highly promising, this proof of concept study has some limitations. We attempted to identify and mitigate some of these limitations in order to inform further work. Relative to most complex classification problems, our patient sample size is small. We attempted to generate as many images as possible using multiple frame sampling per patient, generating 7046 images from 17 patients for use in training, validation and testing datasets for the development of the CNN. While we acknowledge the relatively small image count, we note that at the time this study was undertaken, ours is the largest reported image dataset used to train a network for this application. Large amounts of labelled data is required to produce generalisable neural networks that are capable of accurate classification. These networks must be trained and tested using images acquired under a range of patient related and external conditions so that in real world practice they are able to give accurate predictions. We suggest that out of the more than 7000 images utilised in the development of this network, there was likely a lot of similarity in images extracted from the endoscopic assessment of each patient as they were sequential. In addition to the low patient numbers, this further reduces the true variety of images upon which the network was trained and so could reduce its ability to generalise to other unseen patient images. We identified this early on as a potential source of inaccuracy in our network. Our use of CAMs therefore acts as an additional form of validation in this regard. This study demonstrates that when classifying discriminative features within images, the IPCL patterns are the visual features the network bases its predictions on. We would expect that our network should still maintain reasonable diagnostic performance when tested on new unseen images, given the reasonably stereotyped nature of IPCL morphologies. As reported in **chapter 8**, increasing the size of our image dataset, in order to promote the ability of this network to generalise was a key priority for further work.

To further mitigate the low patient numbers included in this study, we utilised 5-fold cross validation, as outlined in (*figure 52*), in order to benchmark the ability of the CNN to generalise under different operating conditions. This method ensures that our network was trained and tested on all available images, whilst ensuring that no images

that were used for training were used for testing. Given that we observed similar levels of diagnostic performance between folds, we can infer that the underlying architecture of our CNN is robust and that further work should focus on increasing training data to enable it to generalise to classify more diverse input data.

Currently our system is able to differentiate normal (type A) compared to abnormal (type B1-3) IPCL patterns when making its classification of squamous mucosa. As outlined above, even for clinicians attempting to classify lesions invading the submucosa, using the Inoue and Arima classifications, there remains a decrease in diagnostic performance – most studies assessing these classifications reported less accurate characterisation of >SM1 lesions based on IPCL patterns. Preliminary work demonstrated that our network was not able to accurately subclassify IPCL patterns into the individual type A-B3 subgroups; likely due to insufficiently sized image datasets, particularly with regard to lesions with type B3 IPCLs. We therefore chose to include only images of type A, B1 and B2 IPCLs as an input to our CNN for this study. This choice was made since lesions with type B3 IPCLs have typically invaded deep into the submucosal layer, and are no longer endoscopically resectable lesions as a result. This study therefore aimed to classify mucosa as normal or abnormal but still endoscopically resectable. Our further work, reported below, aimed to incorporate endoscopic images of normal and abnormal IPCLs ranging from type A to type B3; a requirement for if our CNN was to be adopted into a clinical setting. Developing a multiclass classifier requires further work and robust validation before it could be used routinely, since the potential outcome of misclassifications in this case could have a devastating impact on patients.

The final main limitation of this study is that our gold standard used only three expert clinicians. Since this was a pilot study, and all of the consensus classifications were correlated with histological results, we argue that this is sufficient to provide accurate ground truth to develop a CNN. Future work will need to externally validate this CNNs performance against a larger panel of clinicians of varying levels of experience before it could be used in a clinical environment. We note however that published work on the JES classification reports interobserver variability

of between 0.6²⁴⁷ and 0.86²⁴² – this should be the benchmark against which a CNNs classifications should be assessed when compared to the classifications of a panel of clinicians during external validation.

In the following chapter we will outline a further study which developed as a result of the study in this chapter. This next body of work aimed to produce, and validate against clinicians, a CNN capable of more precise binary classification based on all of the individual JES subgroups – type A, B1, B2 and B3.

7.6 Summary of this chapter

This study aimed to provide proof of concept that a convolutional neural network could be developed to classify endoscopic images of oesophageal mucosa as normal or neoplastic, based on visualised intrapapillary capillary loop patterns.

1. We have developed a novel AI system that can classify oesophageal squamous mucosa as normal or dysplastic in **endoscopically resectable lesions** (<SM1 invasion) in real time on videos acquired during magnification endoscopy. The main discriminant visual feature on which the CNN based it's decision on was the IPCL patterns.
2. We provide a baseline methodology and performance by which further work on developing a convolutional neural network for this purpose can be conducted and assessed.

Chapter 8 – Validating a clinically interpretable convolutional neural network for the prediction of early squamous cell neoplasia of the oesophagus; comparing diagnostic performance with a panel of expert European and Asian endoscopists

The work presented in this chapter formed the basis of peer reviewed publications. Text and figures were adapted for publication. Citation:

García-Peraza-Herrera, L.C., **Everson, M.**, Lovat, L., Wang, H.P., Wang, W.L., Haidry, R., Stoyanov, D., Ourselin, S. and Vercauteren, T., 2020. Intrapapillary capillary loop classification in magnification endoscopy: open dataset and baseline methodology. *International journal of computer assisted radiology and surgery*, pp.1-9.

Everson, M., Garcia-Peraza-Herrera, L., Wang, H. P., Lee, C. T., Chung, C. S., Hsieh, P. H., ... & Haidry, R. J. A clinically interpretable convolutional neural network for the real-time prediction of early squamous cell cancer of the esophagus: comparing diagnostic performance with a panel of expert European and Asian endoscopists. *Gastrointestinal Endoscopy*

This study was a collaborative work with Dr Luis Garcia Peraza Herrera, who designed, developed and coded the CNN architecture. Clinical requirements, conceptualisation and software requirements were developed by ME

8.1 Introduction

The application of artificial intelligence (AI) in diagnostic endoscopy, as well as other fields of medicine, is gathering pace. One such application of AI is in the endoscopic diagnosis of early squamous cell neoplasia of the oesophagus (ESCN). Endoscopic detection of ESCN remains challenging and there is still a significant miss rate in the identification of early lesions¹⁹⁹. As outlined in previous chapters, early detection is vital, since squamous lesions confined to the oesophageal mucosa are amenable to endoscopic eradication therapy due to their low rates of lymph node metastasis^{80,81,229,230,248}.

Computer-aided endoscopic diagnosis (CAD), using convolutional neural networks (CNNs) has the potential to be used as an adjunct during endoscopy. In the context of our clinical problem, a CNN requires input data (endoscopic images) where specific visual features (IPCL microvessels) correspond with a validated classification (JES IPCL

classification). Stereotyped morphological changes in IPCL patterns correlate with the progressive invasion of ESCN lesions, and as outlined in **chapter 7** have been shown to provide suitable input data for training a CNN to act as a classifier for whether oesophageal tissue is normal or dysplastic. Repetitive training of the CNN, along with increasing volumes and variability of this input data, should allow the development of this feature recognition, in order for the network to make predictions on the histology of a lesion with increasing accuracy. We envisage that such a clinically validated CNN, could provide a useful diagnostic adjunct for endoscopists, particularly those in low volume settings where experience assessing and triaging ESCN lesions to appropriate therapies, may be limited.

As outlined in **chapter 7** we have previously reported a proof of concept study for the use of convolutional neural networks in the real time classification of ESCN lesions based on IPCL patterns^{243,245}. In this study we provided a baseline methodology and diagnostic performance upon which subsequent work, including the body of work discussed in this chapter can be based and assessed. The design of our neural network included a number of important and novel features that provide a good foundation for the development of a clinically applicable system.

Firstly, our results were clinically interpretable. The use of class activation maps (CAMs) confirmed visually that the CNN used IPCL patterns as the discriminative feature upon which it made its classifications, as a clinician would during endoscopy. While a clinician may use other features such as motility of the oesophageal wall, surface texture of the mucosa and intuition, our network performed comparably to human assessors classifying IPCL patterns in other studies^{234,236,239,242,247}. The CAMs also provide a visual representation of the abnormal IPCL patterns associated with the classification of tissue as normal or neoplastic. The use of these CAMs for this purpose could be used as either a 'red flag' type system to highlight abnormal areas, or as a training aid, to familiarise inexperienced clinicians with the appearance of abnormal IPCLs.

Secondly our neural network was able to classify images as normal or neoplastic at video rate, with classification times ranging from 26.17ms to 37.48ms. For such a system to have clinical utility, particularly for in-vivo endoscopic classification such speed is required in order for endoscopists to make real-time decisions.

Thirdly our neural network demonstrated a high degree of diagnostic accuracy. As summarised in (*figure 52*), we observed a mean accuracy for the differentiation of neoplastic compared to normal IPCL patterns of 93.3% (range 86.2-98.3%). We also demonstrated promising sensitivity and specificity for neoplasia classification of 89.7% (range 78.1-100%) and 96.9% (range 92-99.7%) respectively. All of these results compare favourably to studies using human assessors to classify IPCL patterns and as such prompted the continuation of this work.

Despite the promise of the first iteration of our neural network, we identified several limitations in the design of our neural network and dataset in our previous work. This study sought to ameliorate these limitations and expand on the work undertaken previously. To this end we expanded our dataset significantly in order to capture a wider spectrum of ESCN pathology and variability in IPCL patterns; generating a CNN that is both clinically valid and interpretable.

We also identified that of the few published studies assessing the validity of the JES classification, most are conducted in an expert, predominantly Asian high volume setting, and even less is known about the utility of the JES classification in non-expert or Western endoscopists. In addition to further developing our neural network, this body of work also sought to explore the diagnostic performance of European endoscopists, with Asian endoscopists and our convolutional neural network in predicting the histology of ESCN lesions based on their IPCL patterns.

8.2 Aims of this study

This study sought to extend the proof of concept study outlined in **chapter 7**, in order to produce a more clinically robust neural network that was better able to make histology predictions in a more generalised dataset. We also aimed to validate its performance against human endoscopists. Our specific aims were:

1. To train a convolutional neural network capable of classifying squamous mucosa as normal or dysplastic, based on the appearance of intrapapillary capillary loop patterns. The JES IPCL classification¹⁹⁴ was used to characterise visualised IPCLs.
2. To expand the patient numbers and image count in our dataset significantly in order to produce a neural network better able to generalise its predictions on a novel patient population. In addition to images of normal oesophageal mucosa (type A IPCLs), this study's dataset was now to include images of lesions which were both endoscopically resectable (type B1 and B2 IPCLs) and non-endoscopically resectable (type B3 IPCLs).
3. To quantify the diagnostic performance of a cohort of expert Asian endoscopists against that of a cohort of expert European endoscopists using the JES classification; a comparison currently unreported in the literature. The diagnostic performance of these two cohorts could then be compared to the diagnostic performance of our neural network.

8.3 Methods

Patient recruitment

Patients attending for endoscopic assessment at two ESCN tertiary referral centres in Taiwan were recruited (National Taiwan University Hospital and E-Da Hospital Kaosiung). Patients were required to give full written consent prior to the procedure. Patients could be enrolled via one of two routes; those presenting for a routine endoscopy who subsequently were found to have a macroscopically normal oesophagus or those with known or suspected squamous dysplasia or carcinoma who were presenting for assessment or treatment. In all included patients, pathological samples were acquired to confirm histologic diagnosis either by forceps biopsy, EMR, ESD or oesophagectomy where dysplasia or neoplasia were suspected. Patients with active oesophageal ulceration were excluded, as were those who had previously had ablative therapies to the oesophagus, or previous oesophageal surgery. Our study complied with the Declaration of Helsinki. The Institutional Review Board of E-Da Hospital approved this study (IRB number: EMRP-097-022. July 2017).

Endoscopic procedures and video acquisition

Gastrosopies were performed under conscious sedation or local anaesthesia by two expert endoscopists (WLW, HPW) at one of the two referral centres listed above. For the purposes of this study, an expert endoscopist was defined as a consultant gastroenterologist who having completed their formal endoscopic training continues to undertake >50 ESCN assessments and resections per year. A solution of simethicone and water was applied to the oesophageal mucosa prior to recording to remove mucus, food residue and blood. An area of the cleaned oesophageal mucosa was identified as either normal or neoplastic and magnification endoscopy was undertaken on this area to allow clear visualisation of the oesophageal mucosa and microvasculature. All endoscopies were performed using a HD ME-NBI GIF-H260Z endoscope (Olympus, Japan), in combination with an Olympus Lucera CV-290 processor (Olympus, Japan).

Correlating imaged areas with histology

In patients where the endoscopist identified a lesion suspicious for dysplasia on assessment, the lesion border was marked by the endoscopist using cautery forceps, prior to resection or biopsies being taken. This ensured that neoplastic areas were well demarcated to allow magnification endoscopy to be undertaken in the intended area, and in the area from which histologic sampling would take place after imaging. Magnification was undertaken using narrow band imaging (ME-NBI) at 80-100x magnification with a distal attachment in order to interrogate the IPCL patterns and generate a high definition video recording.

For patients with normal mucosa, an area was selected, imaged under ME-NBI and IPCL patterns were classified. A forceps biopsy of the imaged area was then acquired to confirm the imaging findings. Neoplastic lesions were resected by either endoscopic mucosal resection (EMR) or endoscopic submucosal dissection (ESD). Resected samples were formalin fixed and prepared for review by two expert gastrointestinal histopathologists, who reported the histologic findings in accordance with the Vienna classification system²⁴⁴. The worst case histology that was reported was regarded as the histology present within the whole lesion, in line with routine practice.

Labelling ME-NBI videos and establishing an expert consensus

The video recorded from each patient was viewed independently, then classified according to the JES IPCL classification system¹⁹⁴ by consensus of three expert endoscopists (WLW, HPW, RJH). IPCLs were classified as type A, B1, B2 or B3 in order to predict the worst case histology for the entire lesion. Therefore, where there was variation in the microvascular pattern, the highest grade of IPCL pattern observed in each video was used as the overall classification for the lesion as a whole.

The expert's visual IPCL classifications were then correlated with histology taken from the imaged area before a final, concordant classification was assigned to each group of images. For the purposes of this study and in line with the JES classification, type A

IPCLs were considered normal, type B1, 2 and 3 IPCLs were considered indicative of neoplasia of increasing histologic invasion depth.

Generation of image datasets

After each video had been labelled according to the visualised IPCL patterns and histology, single frames were extracted at a rate of 2fps. All HD images were stored in the lossless .png format. Since not all sampled frames were deemed informative, as would be the case in routine practice, all images were quality controlled prior to inclusion in the dataset. Frames were individually quality controlled by a clinician with experience in the endoscopic imaging of oesophageal cancer and several endoscopic imaging studies (MAE). Frames that were degraded by lighting or motion artefact, excessive mucus or blood were removed from the dataset where it was felt that it would not be possible for a clinician to make a decision based on IPCL patterns (**figure 55**). The aim of quality control in this case was to remove images that could safely be regarded as uninformative, rather than to produce a dataset of images that are of higher quality than those seen in daily endoscopic practice.

Following quality control our total dataset comprised 114 patients (45 with normal histology, 69 with neoplastic histology). 67742 images were included in our dataset (28,078 normal and 39,662 dysplastic) with an average of 593 frames per patient. The differential number of patients between the normal and neoplastic groups is the neoplastic subgroup consists of three subgroups (type B1 – B3 IPCLs). To ensure a dataset that included as closely balanced a number of images as possible between all four IPCL subtypes, it was necessary to recruit more patients and as a result images into the neoplastic group.

To maximise the available images for network training and testing, as well as to ensure that no patients included in the training group of images would then be included in the testing group of images, five-fold cross validation was used. For each fold (a set of unique patient images) patients were randomly assigned into a training group,

validation group (used for hyperparameter training) and testing group in the ratio of 80%, 10% and 10% respectively.

Convolutional neural network and class activation maps

We have previously reported the technical methodology involved in the generation of this neural network. Technical development of the CNN was a collaborative work with other study members (LCGPH) and forms part of the work in his doctoral thesis. It is beyond the scope of this thesis to discuss the development of this network in detail but note that the full technical methodology is available in an open access preprint repository (Arxiv)²⁴⁵ and in the published literature^{243,249}. For additional validation of our methodology, our dataset is available via an open access repository (GitHub). For the purposes of this chapter, we have detailed additional technical information relevant to this latest iteration of our network.

We selected to use ResNet-18 as the underlying model for our proposed CNN and was also used as a baseline for result comparison. Based on our pilot study we acknowledge that for our CNN to have clinical utility, it must be clinically interpretable. ResNet-18 provides good classification performance, but is not a clinically interpretable model so adaptations to our methodology were required. The utility of CAMs was evident in our previous work, hence we attempted to produce them as an output of the network for this study.

In our model²⁴⁹, the fully connected layer is removed as described by Zhou et al²⁵⁰. The computation of the class score predictions is reformulated (**figure 53**), which is mathematically equivalent to the solution proposed by Zhou et al, but during the forward pass of the network produces class activation maps. Input images were 256x256, thereby leading to an 8x8 feature tensor at the output of the encoder. Such low resolution would not permit clinical interpretability as the IPCL patterns which informed the classification cannot be identified by a clinician operator. Hence, the feature demonstrated in (**figure 53**) is connected as a side output (connected to the loss) to all encoder resolutions²⁴⁹ in order to produce class activation maps (CAMs).

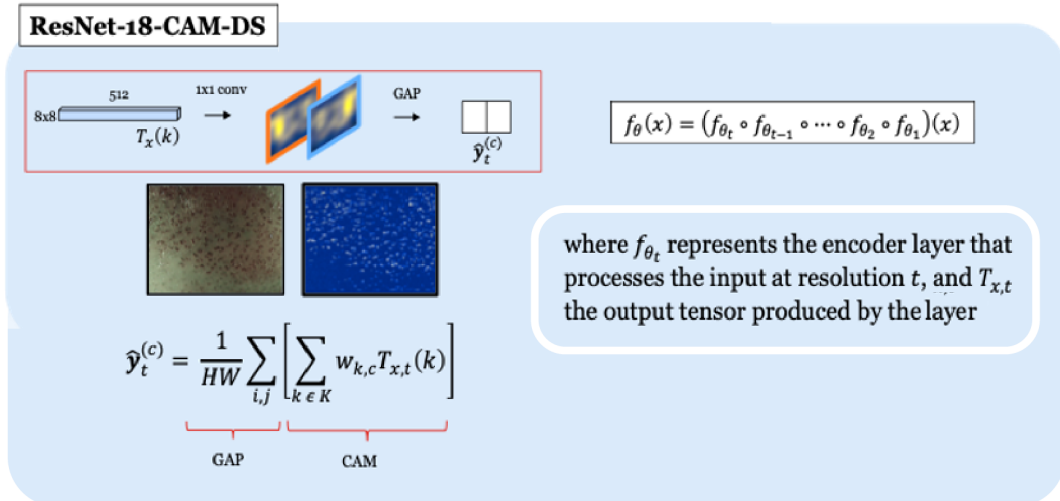


Figure 51 Summary of the CNN side output incorporated to allow clinical interpretability of class activation maps (CAMs), adapted from²⁴⁹

Clinician classification of images and statistical analysis

In addition to the development of our neural network, we sought to investigate the comparable diagnostic performance of a cohort of expert Asian and expert European endoscopists using the JES system to classify images based on IPCL patterns. Using a random number generator frames were sampled from all patients with dysplastic lesions at a rate of 1:500, this generated 79 random images. The same number of frames were taken from all patients with normal oesophageal mucosa to ensure a 50:50 split of normal to dysplastic lesions in the testing set assessed by experts. The sample of images was reviewed by a clinician (MAE) to ensure none were too similar; to ensure that all images were fair representation of the IPCL variability within that patients lesion.

All the selected images were assessed independently by a group of 9 expert endoscopists (5 Asian, 4 EU experts) using a bespoke online platform (**figure 54**) to preserve image quality. Endoscopists were asked to classify the predominant IPCL pattern as type A, or type B1 to B3. This classification was assessed against the expert consensus classification, which correlated with the histology for each IPCL subtype. Experts were also asked to state whether they could visualise submucosal vessels in each image.

ching-tai.lee

< Previous

Question 157 of 158

Next >

Continue >>

Sign Out



Using the JES IPCL **AB** classification, please predict the lesion histology based on the observed patterns of IPCLs seen in this image:

☐ Normal/Low grade IN (A)
☐ High grade IN/lamina propria (B1)
☒ Muscularis mucosa/Submucosa 1 (B2)
☐ Submucosa 2 or deeper (B3)

☐ Were deep submucosal vessels visible in this image?

Submit

Figure S2 HD image assessment portal used by the expert panels in this study to classify IPCLs

Diagnostic performance measures were calculated for individual endoscopists and for each geographical group. Accuracy, sensitivity and specificity were calculated; we report the average diagnostic performance for Asian endoscopists, EU endoscopists and our CNN. An F1 score, a measure of the diagnostic accuracy of binary classification algorithms, was also calculated. Interobserver agreement for clinicians was calculated using Krippendorff's alpha and assessed using a modified Likert scale (Landis and Koch). Multilevel regression analysis was performed to test the association between endoscopists identifying submucosal vessels and the presence of neoplasia.

We calculated the average per-frame diagnostic performance for our CNN (F1 score, accuracy, sensitivity and specificity) as well as an analysis of per-patient diagnostic performance. To establish a per-patient classification as normal or dysplastic, images were grouped by patient. Our network outputs a probability that each image shows dysplastic tissue, we set a per image threshold >0.5 as being positive that an image contains dysplasia. To establish a per-patient classification the probability outputs of

each frame from that patient were used to compute an overall average probability of the presence of dysplasia. A threshold of >0.5 average probability was used to classify that patient as dysplastic and <0.5 as normal, CNN predictions were compared with the ground truth of histologic analysis.

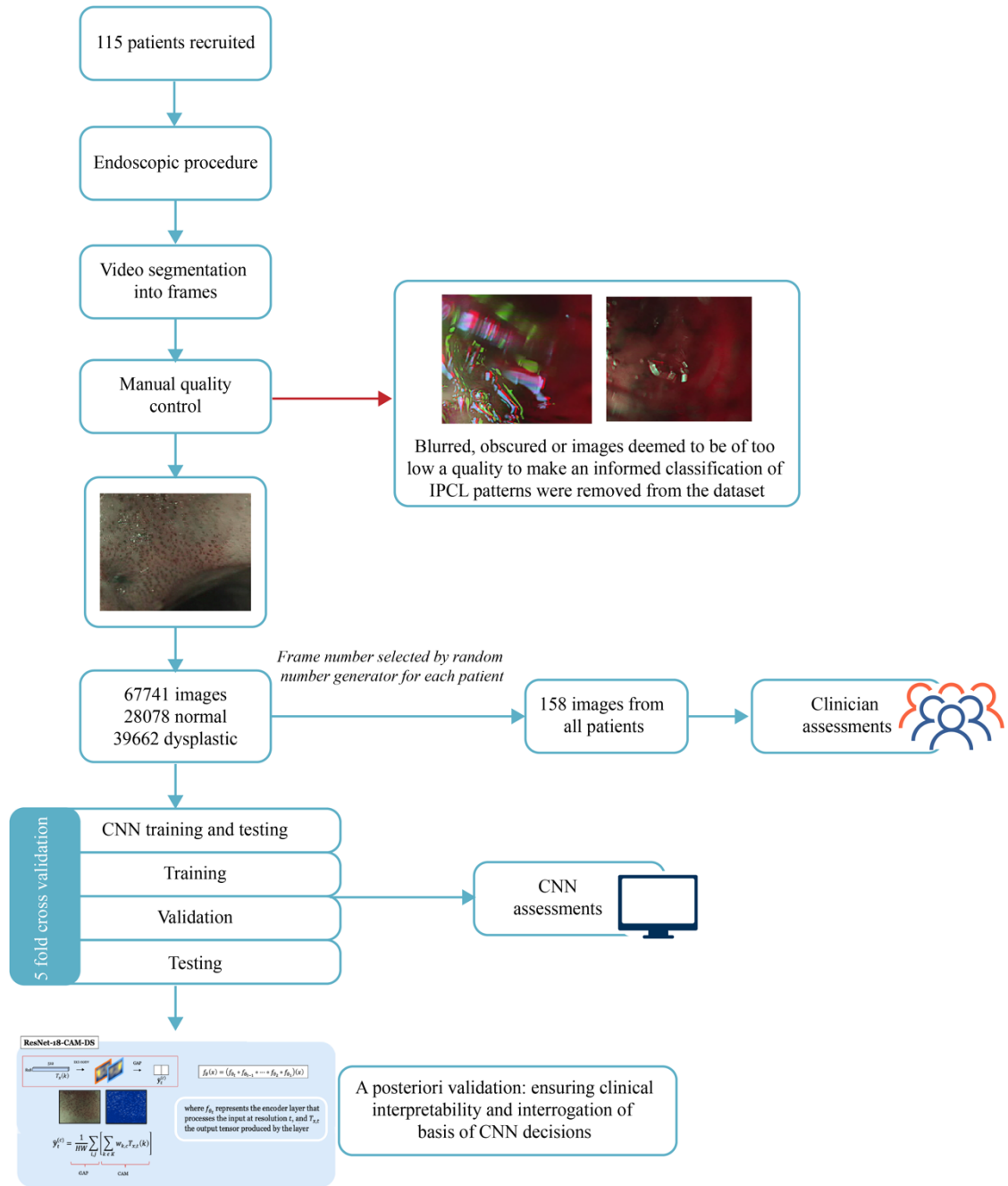


Figure 53 Schematic representation of the chapter 8 study workflow

8.4 Results

Patient characteristics

115 patients were included. 45 of these were determined to have a normal oesophageal mucosa. 70 patients were determined to have a neoplastic lesion, with IPCL patterns ranging from type B1-B3. The breakdown of patient numbers by histologic grade of ESCN is reported in **(figure 56)**.

Patient histology	Number of patients in dataset
Normal	45
HGD/LPM	35
MM/SM1	17
>SM2	18

Figure 54 Summary of patient numbers recruited to the study by histologic stage of ESCN

CNN performance for the classification of IPCL patterns as normal or dysplastic

We assessed both the per-frame and per-patient diagnostic performance of our CNN. On a per-frame basis the diagnostic performance results of our CNN are summarised in **(figure 57)**. Although there was some variation between folds, we report an average F1 score, accuracy, sensitivity and specificity of 94.0%, 91.7%, 93.7% and 92.4% respectively **(figure 57)**.

Fold	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 score (%)
1	92.5	99.6	81.3	94.1
2	92.4	91.3	95.1	94.4
3	97.4	98.3	96.6	97.0
4	94.5	98.9	89.3	95.1
5	81.9	80.5	99.8	89.2
Average	91.7	93.7	92.4	94.0

Figure 55 The diagnostic performance measures for the detection of dysplastic tissue based on IPCL patterns for all folds of the CNN developed in this study

We also assessed the per-patient diagnostic performance of our CNN. For this analysis we used 12 patients per fold for testing, using five folds therefore gave a total of 60 independent patients that were used to test the CNN classification. Our results are promising, on a per-patient level we demonstrate that our CNN only failed to correctly classify one patient as normal or dysplastic (patient 158). In this case it reported a false positive result for dysplasia in two independent folds. To further interrogate potential image features that may have posed a diagnostic challenge to our CNN, we extracted the images that were classified with the highest probability as normal or dysplastic in each fold, but were subsequently found to be either a true positive (TP), true negative (TN), false positive (FP) or false negative (FN). These diagnostically challenging images are displayed in **(figure 58)**.

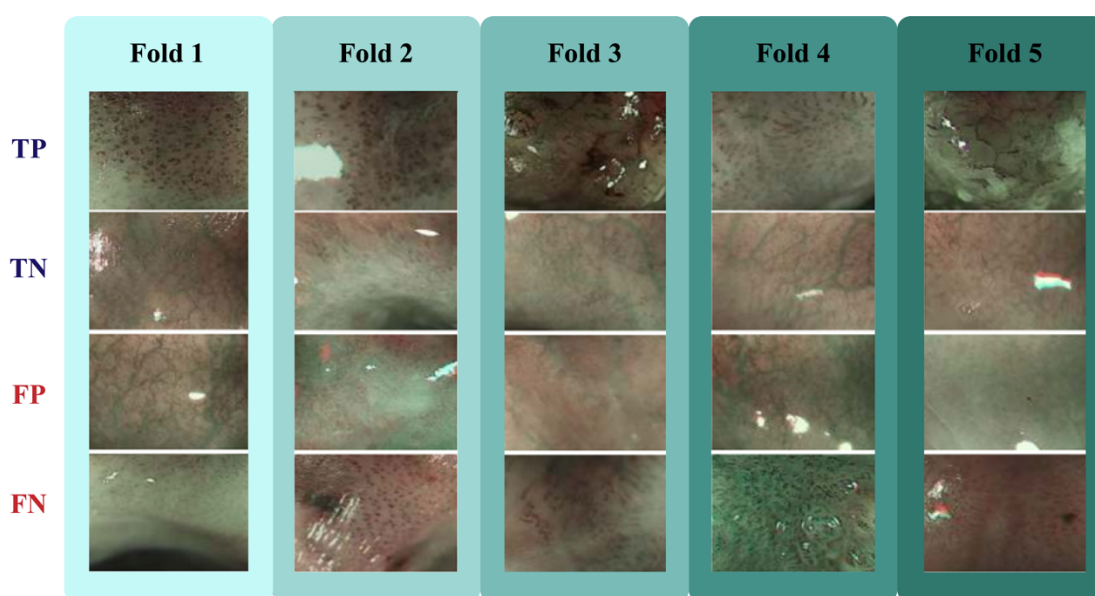


Figure 56 Images classified with the highest probability by our CNN which were subsequently found to be a true positive (TP) or negative (TN) or a false positive (FP) or negative (FN)

Comparative pooled diagnostic performance of EU and Asian expert endoscopists

158 still images were reviewed by a panel of 9 expert endoscopists (5 based in Asian centres, 4 based in European centres). All endoscopists assessing images were blinded to the endoscopy procedure and histology results. The diagnostic performance measures of Asian and EU endoscopists were calculated; the two groups of endoscopists achieved F1 scores of 98% and 97%; accuracy of 97.1% and 96.9%; sensitivity of 96.9% and 98.9% and specificity of 97.6% and 91.5% respectively.

The diagnostic performance measures of all expert endoscopists were pooled to generate a mean performance for human expert assessors. The pool expert endoscopist F1 score, accuracy, sensitivity and specificity was 96.5%, 94.7%, 97% and 88% respectively. Performance measures are summarised in **(figure 59)**. A receiver operating characteristic curve demonstrates an AUC of 0.96, suggesting a high level of diagnostic accuracy **(figure 60)**.

Interestingly we note that Asian endoscopists had significantly higher specificity than EU endoscopists (97.6% vs 91.5% $p=0.01$) whereas EU endoscopists had significantly higher sensitivity than their Asian counterparts (98.9% vs 96.9% $p=0.01$). This finding may relate to a tendency of European endoscopists, likely less familiar with the endoscopic assessment of ESCN to over-call neoplasia.

Endoscopist	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 score (%)
EU average	96.9	98.9	91.5	97.0
Asian average	97.1	96.9	97.6	98.0
Pooled average	94.7	97.0	88.0	96.5

Figure 57 Summary of the pooled diagnostic performance for both European and Asian expert endoscopists when assessing and classifying abnormal IPCL patterns

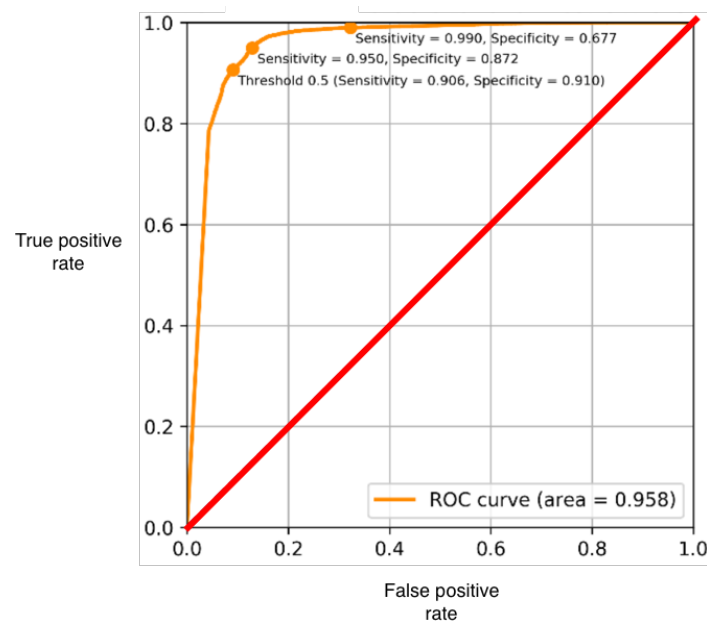


Figure 58 Receiver operant characteristic (ROC) curve for the CNNs diagnostic performance

Interobserver variation between experts classifying IPCL patterns

Both panels of experts demonstrated very good interobserver variation when classifying IPCL patterns as either normal (type A) or abnormal (type B). There was no statistically significant difference between the two groups when classifying ($P=0.29$), but we did note that a European cohort achieved comparable agreement with an Asian cohort (**figure 61**).

Expert panel	Kappa (95% CI)
European	0.90 (0.85-0.95)
Asian	0.85 (0.79-0.92)

Figure 59 Interobserver agreement for the classification of normal and abnormal IPCL patterns in a European and Asian expert endoscopist cohort

Expert predictions of histologic invasion depth based on IPCL patterns

This study assessed the diagnostic performance of two panels of expert endoscopists when classifying IPCL patterns according to the JES classification. (**figures 62, 63, 64 and 65**) outline the diagnostic performance measures for each panel of experts assessing images containing IPCLs of each type. Endoscopists were asked to classify the IPCL pattern in each image assessed, where the classified IPCL pattern and associated histology prediction made by the expert matched that of the reported histology, a diagnosis was considered correct.

Type A vessels	European endoscopists	Asian endoscopists	P value
Accuracy	97.0%	97.1%	0.91
Sensitivity	98.9%	96.9%	0.01
Specificity	91.5%	97.6%	0.01
PPV	97.1%	99.1%	0.56
NPV	97.0%	91.7%	0.06

Figure 60 Summary of expert endoscopist performance statistics for the classification of Type A IPCL patterns

Type B1 vessels	European endoscopists	Asian endoscopists	P value
Accuracy	92.1%	90.6%	0.26
Sensitivity	90.0%	79.0%	0.002
Specificity	92.8%	94.6%	0.17
PPV	80.9%	83.2%	0.57
NPV	96.5%	93.0%	0.06

Figure 61 : Summary of expert endoscopist performance statistics for the classification of Type B1 IPCL patterns

Type B2 vessels	European endoscopists	Asian endoscopists	P value
Accuracy	81.3%	79.2%	0.24
Sensitivity	62.8%	79.5%	<0.001
Specificity	87.4%	79.2%	<0.001
PPV	62.0%	55.6%	0.19
NPV	87.8%	92.2%	0.02

Figure 62 Summary of expert endoscopist performance statistics for the classification of Type B2 IPCL patterns

Type B3 vessels	European endoscopists	Asian endoscopists	P value
Accuracy	84.0%	82.7%	0.36
Sensitivity	63.2%	41.1%	<0.001
Specificity	90.6%	95.8%	<0.001
PPV	68.1%	75.7%	0.56
NPV	88.6%	83.7%	0.79

Figure 63 Summary of expert endoscopist performance statistics for the classification of Type B3 IPCL patterns

Deep submucosal vessels may be a discriminative feature for neoplasia

All experts reviewed each image and stated whether or not they could visualise the deep submucosal vessels in addition to the IPCL patterns. We demonstrated a statistically significant difference in the visualisation of these structures [$p < 0.001$]; in normal tissue submucosal vessels were visible in 57.7% of assessments (213/369), whereas in dysplastic tissue they were visible in only 6.3% of images (66/1053) (**figure 66**).

This finding was consistent in both the European and Asian groups who identified submucosal vessels less frequently in dysplastic tissue compared to normal tissue

(7.5% and 5.3% compared to 70.7% and 47.3% respectively. These findings indicate that the absence of visible submucosal vessels may be another endoscopic finding that should raise the suspicion of ESCN.

Endoscopist group	Histology	Visible submucosal vessels % (N/N)	Odds ratio (95% CI)	P value
Pooled	Normal	57.7% (213/369)	0.05 (0.04-0.07)	<0.001
	Dysplastic	6.3% (66/1053)		
European	Normal	70.7% (116/164)	0.03 (0.02-0.05)	<0.001
	Dysplastic	7.5% (35/468)		
Asian	Normal	47.3% (97/205)	0.06 (0.04-0.10)	<0.001
	Dysplastic	5.3% (31/585)		

Figure 64 Absence of visible submucosal vessels may indicate the presence of ESCN

Assessing the clinical interpretability of this CNN

As identified in our pilot study in **chapter 7**, any CNN designed for the purpose of histologic classification of ESCN using IPCL morphologies must be clinically interpretable for it to have utility in an in vivo endoscopy environment. We previously developed class activation maps (CAMs) in order to provide a degree of clinical interpretability. This study refined these maps to ensure that they provide higher resolution demarcation of IPCLs classified by the CNN.

CAMs are visual representations of what the CNN ‘sees’. When making a classification of an image certain visual features, represented by clusters of pixels, are regarded as discriminative. Over the thousands of iterations the CNN requires to train and then test as a classifier it will find some clusters of pixels more discriminative than others. In **(figure 67)** we demonstrate representative images used during testing, alongside the CAM revealing which areas of the image informed the classification. We observed that the CNN found the IPCL patterns to be informative, rather than other spurious features seen in the image such as submucosal vessels, specular reflections and bubbles. This is consistent with the CAMs we produced in our pilot study. CAMs serve two purposes; they serve as form of internal validation, increasing the confidence that

our CNN classifies based on IPCL patterns and secondly the allow our network outputs to be clinically interpretable.

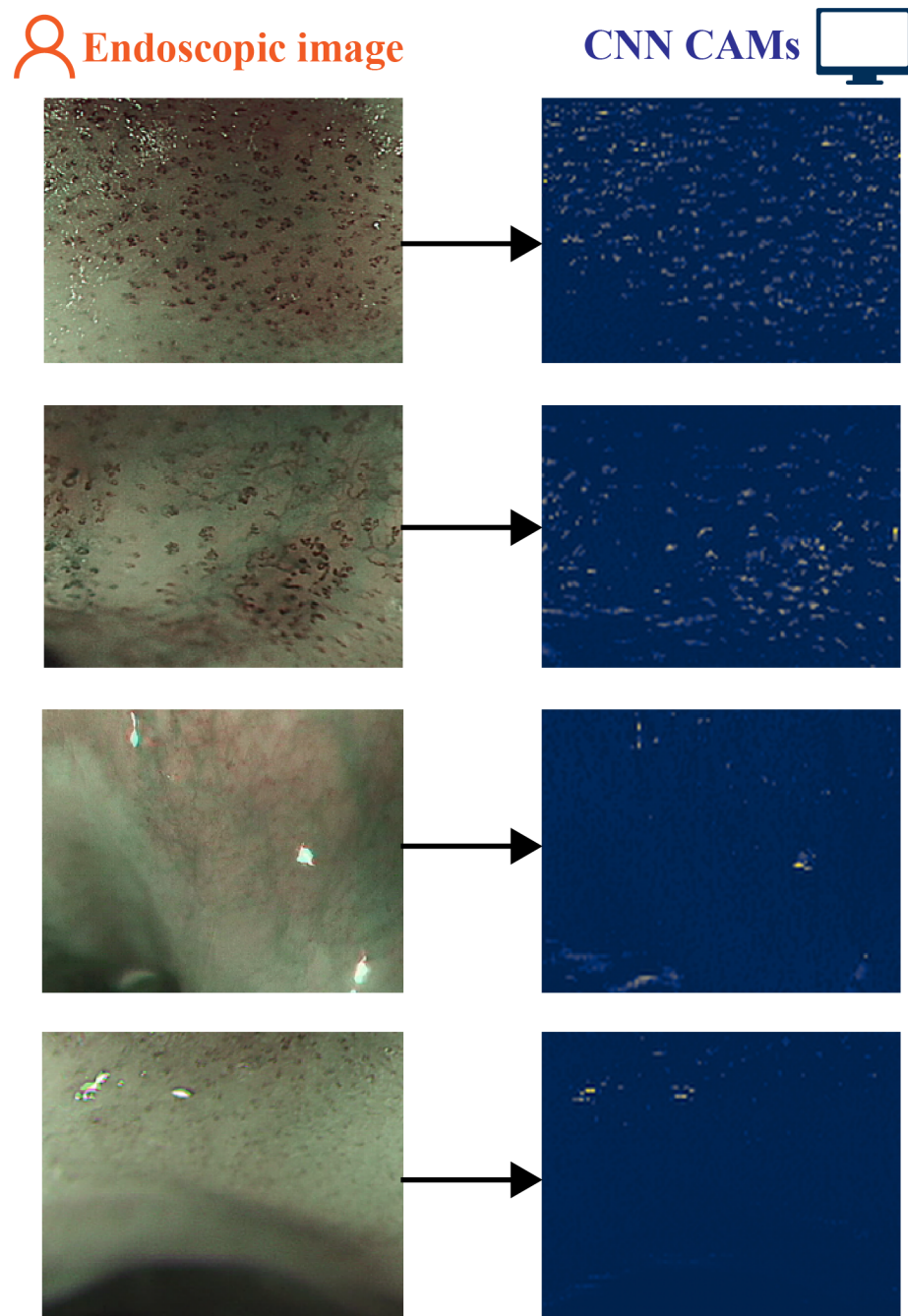


Figure 65 Representative examples of class activation maps generated by this CNN

8.5 Discussion

This study reports a comparison in the diagnostic performance between a cohort of expert endoscopists based in Europe and Asia, with that of a clinically interpretable convolutional neural network in the classification of IPCL patterns in early squamous cell neoplasia of the oesophagus.

ESCN lesions are subtle and can be easily missed; hence several classification systems exist to assist clinicians in the identification and characterisation of ESCN based on IPCL morphology^{236,238,239}. Early ESCN lesions, confined to the mucosa, can be treated successfully endoscopically without the need for more invasive surgery. Endoscopic adjuncts to detect and identify these IPCL patterns and thereby ESCN lesions is vital to improve patient outcomes. This study sought to address the lack of data on the diagnostic performance of Western clinicians assessing ESCN; almost all studies to date have assessed the use of IPCL classifications within high-volume referral centres, predominantly in Asia. This study has provided a benchmark for the performance of the JES IPCL classification when used by Western clinicians.

In chapter 7 we reported a proof of concept study which outlines the use of an early iteration of our CNN to classify oesophageal mucosa as squamous or dysplastic^{243,245}. We therefore aimed to further develop this CNN using an expanded dataset to assess whether it could maintain or improve its ability to classify IPCL patterns in a dataset with greater variability of endoscopic images. We then aimed to compare its performance with expert clinicians to assess whether it could provide an incremental improvement in diagnostic performance.

Testing of our CNN revealed promising results; we demonstrate accuracy for the prediction of dysplasia of 91.7%, an F1 score of 94% and sensitivity of 93.7%. This compares favourably with an analysis of the diagnostic performance of expert clinicians using the JES classification by Oyama et al., which reported accuracies of 91.9 - 95.9% and sensitivities of 55% - 97.5%. This CNN was able to classify images at video rate and was trained using consecutive, segmented frames from endoscopy videos, so has the potential for real-time use.

We note that since the development of our proof of concept study, other groups have published work in this area. Guo et al. propose a CNN for the classification of dysplastic mucosa compared to normal tissue. This group used 6671 images to train their system; reporting sensitivities of 98%. We noted, including in our own proof of concept study, that relatively small datasets such as this may struggle to achieve such diagnostic performance when trained and tested on larger, more variable datasets²⁵¹. If the CNN is not exposed to a wide variety of operating conditions, image resolutions or patient variability it will likely struggle to perform in a wider, unselected cohort.

Zhao et al. report another CNN trained on a small dataset of 1383 images, with reported accuracies of 87% for the detection of dysplastic IPCL patterns¹⁹³. This dataset was heavily skewed towards images containing type B1 IPCL patterns. We therefore argue that this network is unlikely, based on our experience, to classify lesions continue B2 or B3 IPCLs with such accuracy. A system with poor diagnostic performance in identifying B2 and B3 IPCLs as abnormal could have serious clinical consequences, since these patients may require more radical treatment options such as surgery. In contrast to these studies, the CNN we propose has been tested and trained on a larger dataset, and despite the inclusion of a balance of IPCL subtypes maintains a high level of diagnostic accuracy.

This study assessed the diagnostic performance of a CNN compared to expert clinicians. We recruited a cohort of Asian and EU endoscopists and assessed their performance using a representative sample of images from the same patient cohort used to train our CNN. We selected both EU and Asian endoscopists to assess whether there may be geographic variations in performance, based on exposure to assessing and managing ESCN lesions; this is the first such reported study to assess this. We observed that all endoscopists achieved high diagnostic accuracy using the JES system. The Asian endoscopists achieved an accuracy of 97.1%, F1 score 98%, sensitivity of 96.9% and specificity of 97.6%. This was comparable with EU endoscopists who achieved an accuracy of 96.9%, F1 score 97%, sensitivity of 98.9% and specificity of 91.5%. It is likely that the unexpectedly higher sensitivity of EU endoscopists for

identifying dysplasia is likely related to them over-diagnosing dysplasia, as evidenced by the concomitant lower specificity.

Our CNN achieves lower but comparable diagnostic performance when compared with the expert endoscopists in this study. In some instances it performed better than some individual clinicians for individual performance measures. We note that the diagnostic performance of our expert panel was higher than that reported in other studies using both the JES and other IPCL classifications, which suggests that with further refinement our CNN shows promise as a diagnostic adjunct to clinicians involved in the endoscopic management of ESCN. Given that this study only compared our network to expert clinicians in both geographic settings, we propose that further work should assess our CNN against the performance of less experienced endoscopists.

In **chapter 4.3** we outlined work by other groups on the development of a CNN to classify IPCL patterns as normal or dysplastic. We note that our study uses the largest reported dataset of images reported at present, but does not report the largest number of patients. Estimating the effect of higher image counts but lower patient numbers on CNN performance is difficult. We suggest that the large numbers of sequential images used may mean that the estimates of diagnostic performance for our CNN may be closer to what is currently achievable by a CNN, compared to studies trained using smaller datasets. Smaller image datasets are prone to overfitting, meaning that they may over-exaggerate diagnostic performance, which cannot be replicated on larger and more varied image datasets.

While not directly comparable, consideration should also be given to the performance of our CNN relative to other studies using a range of diagnostic classifications. Oyama et al. demonstrated an overall accuracy for histology prediction of 90.5% using the JES classification²³⁶. In a retrospective analysis of patients with histologically confirmed ESCN who underwent ME-NBI assessment, Mizumoto et al. report diagnostic accuracy of 82% for the differentiation of lesions superficial to the LPM compared to those invading deeper than the MM using the JES classification²⁵². Kim et al. report an overall

accuracy for identifying dysplastic lesions using the JES system of 78.6%²⁴². Our study demonstrates that our CNN classified lesions with a higher average accuracy of 91.7% for the prediction of dysplastic tissue.

We identify some limitations of this study, which we have attempted to mitigate. There is a degree of selection bias in that videos identified by two expert clinicians as either dysplastic or non-dysplastic were included. We acknowledge that this may dispose our dataset to more obvious lesions, but attempted to include lesions across the spectrum of pathology (HGD-SM3+) to mitigate this. Another limitation is that we cannot correlate a specific IPCL with a histology result rather we classified imaged areas in their entirety; in keeping with other studies in this field and in keeping with clinical practice. To mitigate this we classified lesions by their most abnormal and predominant IPCL pattern and correlated this with the most serious pathology identified under microscopy, since this is what occurs in clinical practice and is most appropriate.

Further work should focus on training CNNs with datasets of increasing size and variability. While this remains in the future, in order to be used in clinical settings CNNs need to be rigorously validated against clinicians. We suggest that further studies should aim to characterise the diagnostic performance of non-expert endoscopists based in high and low volume centres for the treatment of ESCN. This could provide a benchmark against which to assess the potential utility of our CNN as an adjunct to endoscopic assessment of ESCN by non-experts.

Chapter 9 - Discussion and further work

This body of work aimed to present novel methods to improve the endoscopic detection of early oesophageal neoplasia. At the outset this thesis highlighted that oesophageal cancer is a significant health burden worldwide. Given the evolving epidemiologic trends in the incidence of different histological subtypes of oesophageal cancer, improving detection is likely to require a different focus in a Western setting (where adenocarcinoma predominates) and an African or Asian setting (where squamous cell carcinoma predominates).

9.1 Minichromosomal maintenance component complex 5 (MCM5) as a marker of Barrett's oesophagus related neoplasia – a feasibility study

The first study in this thesis was presented in chapter 5. It is well recognised that the endoscopic detection of oesophageal cancer is variable, with a significant mis-rate for detection on endoscopy in the year preceding a diagnosis. In the UK a significant number of patients are referred for urgent access to endoscopy on the basis of 'red flag' symptoms, an overwhelming majority of these patients are found to have a normal oesophagus and for those with pathological findings, most are benign etiologies. The resource burden on secondary care services, from endoscopy to pathology, associated with these referrals is significant. There is an unmet need to streamline the patients referred to endoscopy, in order to improve cancer detection rates, reduce patients being referred for unnecessary investigations and to reduce the increasing burdens on endoscopy services. We investigated whether a novel biomarker, MCM5 could be used to better identify patients with early oesophageal cancer, particularly in the context of BO.

At the outset of this study, we envisaged that a validated biomarker used in this context, could serve two purposes. A biomarker that could be obtained from blood or gastric fluid could reduce the need for endoscopy in patients with 'red flag' symptoms. In the case of MCM5 we demonstrate that it can be obtained and quantified from aspirated gastric fluid samples; in our study this was achieved via gastroscopy, but this could feasibly be performed through the less invasive insertion of a nasogastric tube.

Such a procedure could be performed in a primary care by appropriately trained staff and used to identify patients with red flag symptoms in whom the likelihood of it being caused by a cancer was low. These patients could then be given a trial of alternative treatments or less invasive tests to look for other benign causes. Conversely patients in whom the likelihood of cancer was high could be triaged to endoscopy services.

The second potential use of such a biomarker, could be to rationalise the way in which patients with BO undergo surveillance. As discussed the majority of patients with BO do not progress to dysplasia or neoplasia, yet they continue to undergo surveillance over a long period of time. A biomarker that could be quantified from gastric aspirates undertaken during BO surveillance endoscopy could highlight patients who have early neoplasia that was missed on macroscopic inspection; these patients could then be recalled for further surveillance at a shorter interval than previously planned.

As a feasibility study our work into the value of MCM5 for the detection of adenocarcinoma presents an interesting signal for further interrogation. We also report a methodology to use as a benchmark for further work. This study demonstrated that MCM5 levels are significantly raised in patients with oesophageal adenocarcinoma compared to patients with a macroscopically normal oesophagus (154.1 vs 14.4 [$p = 0.04$]). This suggests that there may be some value in the use of MCM5 as a screening test for patients presenting to primary or secondary care physicians with red flag symptoms. Further work should seek to further evaluate the discriminative performance of MCM5 levels for the detection of cancer in a larger, powered study. If there is evidence that MCM5 quantification can differentiate unselected patients with cancer from those without it, more work into the logistical aspects and cost effectiveness of obtaining gastric aspirates without the need for endoscopy should be considered.

Our second objective was to investigate whether MCM5 could be used to differentiate patients with non-dysplastic BO, from those with dysplasia or neoplasia; if this was the case it MCM5 levels could be used to inform surveillance intervals. We reported a stepwise association in the mean MCM5 expression levels between patients with a

normal oesophagus, NDBO, HGD and oesophageal adenocarcinoma (14.4, 49.8, 112.3, 154.1 respectively). In this study there was no statistically significant difference in MCM5 expression levels between these three patient groups. Since this was a feasibility study and hence not formally powered, there is a risk of type II errors. Further work in this area should involve a formally powered study with increased patient numbers, in order to further interrogate whether there is a relationship between increased MCM5 expression and the presence of BO, BO associated dysplasia and subsequently oesophageal cancer. We suggest that patients included within an oesophageal cancer cohort should include patients both with BO associated adenocarcinoma, as well as squamous cell carcinoma.

9.2 Assessing whether virtual chromoendoscopy using iScan Optical Enhancement improves the detection of Barrett's oesophagus associated neoplasia in expert and non-expert endoscopists.

In chapter 6 we presented a study which aimed to establish whether a novel advanced imaging system, iScan OE, could provide an improvement in the diagnostic performance of expert endoscopists assessing for BO associated neoplasia compared to HD-WLE. This study demonstrated that iScan OE is a promising technology for use in the endoscopic detection of BO associated neoplasia, both on macroscopic assessment and during magnification endoscopy.

Our study demonstrated a significant improvement in both trainee and expert endoscopists for the detection of BO associated dysplasia and adenocarcinoma, when using iScan OE compared to HD-WLE. We observed an improvement in the diagnostic accuracy of dysplasia detection in trainees from 63% to 76% when using OE and an improvement in the accuracy of experts from 78% to 86%. We observed significant improvements in the sensitivity and specificity of both groups using OE compared to HD-WLE.

The relevance of this finding is difficult to interpret. At the outset of this study, it would have seemed more likely that trainee endoscopists using HD-WLE would exhibit a diagnostic performance closer to that of experts, compared to when they used OE,

since most of their training would have been conducted using HD-WLE endoscopes. There are two possible explanations for this observation we suggest; firstly that OE may reduce the learning curve for the detection of dysplasia in trainees and so allow them to perform closer to an expert level, secondly that the experts used in our study may still be on their learning curve when using OE and so their diagnostic performance may be expected to improve further with additional training. Further work, with a larger cohort of expert and trainee endoscopists, with multiple assessments separated temporally may provide more granular detail on how the use of OE improves performance in these two groups

This study also validated a previously proposed magnification endoscopy classification system for use with the iScan OE platform²²⁴. Our iScan MV classification system, when used to assess the mucosa under OE magnification endoscopy, confers a significant improvement in accuracy of dysplasia detection using OE compared to WLE, with accuracy improving from 66.7% to 79.9%. We also observed an improvement in sensitivity ifrom 83.4% to 86.3% with OE. In this part of the study only expert endoscopists were assessed, since trainee or non-expert endoscopists are less likely to have access to or experience with magnification endoscopy.

We identify several limitations to our study. The first is that the prevalence of BO associated neoplasia in our study population, does not match that of the unselected population attending for either routine endoscopy, or BO surveillance endoscopy. Due to the relative rarity of BO neoplasia it would be logistically challenging to conduct a study such as this in an unselected population. Our methodology is in keeping with other large studies in this area, who also used an enriched population^{167,221}.

A second limitation is that we did not use video recordings in this study; this was a conscious decision. We recognise that endoscopists rely on a number of visual cues when assessing the oesophageal mucosa other than mucosal and vascular patterns; the contours of the luminal wall as well as abnormalities in its motility all contribute to the detection of neoplastic tissue. Previous work from our group demonstrated that asking endoscopists to assess videos posed significant challenges when trying to

standardise what tissue they felt was dysplastic – for example, would their localisation of the dysplasia be based on time seen in the video, segmented areas of dysplasia or pinpoint selections in various frames of the video. To allow direct comparisons between the assessments of endoscopists we therefore opted to use images of the oesophageal lumen and mucosa. To mitigate the fact we had not used videos we used multiple sequential images from each patient to simulate an endoscopic ‘pullthrough’ manoeuvre. This meant that the assessing endoscopists were able to make an assessment on multiple areas of the oesophagus for each patient and to get some appreciation of the variation in mucosa throughout. This methodology is not dissimilar to other reported work in the field. In future work it would be recommended to record assessments on video footage, possibly with the use of a bespoke computer portal to facilitate a more standardised comparison between assessors selections.

The Preservation and Incorporation of Valuable endoscopic Innovations (PIVI) thresholds recommended by the ASGE, outline the diagnostic performance a new endoscopic technology should meet prior to incorporation into routine endoscopic practice. At present no advanced endoscopic imaging system has been demonstrated to meet PIVI thresholds for the incorporation into routine endoscopic practice. We note that our system is comparable to other advanced imaging systems in the improvement it offers expert endoscopists for dysplasia detection. Our study, as well as others in the reported literature also demonstrated that by the standards set by PIVI, HD-WLE would also not reach the threshold required to be incorporated into routine practice. Further work should look to investigate whether the thresholds set by PIVI are attainable and whether potentially useful innovations are not being investigated further because they do not appear to meet PIVI thresholds in initial testing. The validity of PIVI thresholds is therefore debatable. If our study alone had demonstrated that the expert endoscopists diagnostic performance for the detection of dysplasia with HD-WLE was below PIVI thresholds, then it raises questions surrounding their expertise. However given that most reported work demonstrates that in expert endoscopists the use of HD-WLE only meets PIVI thresholds inconsistently, we argue that further work should be completed to establish the validity of the PIVI thresholds.

9.3 Developing a clinically interpretable convolutional neural network to aid in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof of concept study.

Chapter 7 presents the feasibility study and preparatory work undertaken for the development of a novel CNN for the characterisation of ESCN based on IPCL patterns. The follow up study to this work is presented in Chapter 8. This small study developed the architecture required for this CNN. As a result of this work we identified further refinements to improve the performance and validity of our system, as well as to develop features that could give it clinical utility in downstream applications.

Early iterations of our network were ‘black-boxed’. This meant that our system was producing classifications based on image features that were not available to scrutinise by human operators. During our quality control process images were cropped to remove any patient identifiable data or settings information that was included within the black margin that surrounds the endoscopic images. We identified very high levels of diagnostic accuracy based on an early test run of our system which would usually not be expected when training CNNs. Upon further interrogation it was found that the CNN was classifying based on the size of the remaining black margins, which after quality control by a human clinician were larger in patient images which displayed dysplastic lesions. This prompted two revisions to our methodology. Firstly we developed an SOP to ensure that during quality control only the part of the image extracted from the videos which contained the endoscope view was included in training. This ensured that no additional, non-discriminative image features were used by our CNN to classify images. Secondly, this led to our development of CAMs. The class activation maps ensured that only relevant image features were used by the neural network to classify tissue as dysplastic or normal. The CAMs allowed clinicians to scrutinise the architecture of the network and its decision making, to ensure that it was behaving in a manner that mimicked that of a clinician.

The results obtained in this study were entirely novel and so we were unable to compare our results to other work available in the literature. When assessing our results we noted significant variability between folds. This is often an indicator that

the CNN is ‘overfitting’ data, meaning that it becomes very good at classifying a small subset of patients, but is less able to generate accurate classifications when presented with a greater variety of patients and operating conditions. The solution to this was to expand our patient dataset to include a greater number and variety of patients, so that the CNN trains using a wider range of image features.

9.4 Validating a clinically interpretable convolutional neural network for the prediction of early squamous cell neoplasia of the oesophagus; comparing diagnostic performance with a panel of expert European and Asian endoscopists

The final and largest study in this body of work was presented in chapter 8. This study built upon the feasibility study reported in chapter 7. We aimed to develop a clinically interpretable CNN, capable of classifying magnification endoscopy images of oesophagus squamous mucosa as normal or dysplastic, based on IPCL patterns. From this study we identified that larger numbers of patients were required to ensure that the features that our neural network was recognising as dysplastic could ‘generalise’ to accurately detect dysplasia in a wider, more varied patient population. Our second study therefore increased the number of patients recruited nearly seven-fold and increased the number of included images nearly ten-fold. Having achieved this we aimed to validate it against a cohort of Asian and European expert endoscopists in order to benchmark its diagnostic performance.

This study represents a significant contribution to the field of artificial intelligence in endoscopic diagnosis. Unlike other ‘black box’ type CNNs, we specifically designed our system to be clinically interpretable. The development of CAMs, visual representations of the image features that the CNN has identified as discriminative for the identification of cancer, could facilitate the use of this system intra-endoscopy. We envisage that with validation and further development, the CAMs produced by our network and demonstrated in (*figure 67*) could be used to highlight to clinicians during endoscopy which areas of the endoscopic image contain abnormal tissue. In the case of magnification endoscopy this could permit the more accurate delineation of lesion resection margins; areas containing abnormal IPCLs could be displayed endoscopically and marked for resection. Our system was also developed using

sequential video frames and can classify at video rate, another vital ability if it is to have clinical utility.

Our methodology is robust, we report one of the largest image datasets for this purpose in the literature. The ‘non-black box’ nature of our system allows us to interrogate that it is predicting the presence of dysplasia on the basis of IPCLs and not on other obscure, non-discriminative features. Similarly during endoscopy, endoscopists utilise an array of visual features to detect dysplasia, often subconsciously. Our use of sequential still images, which when run in sequence are essentially a video, should provide a strong foundation for the use of our CNN to classify endoscopic videos in real time.

Since the study in chapter 8 was conducted, other groups have reported similar work. Horie et al.¹⁹², using 8248 images from 384 patients with a mixture of OAC and ESCN, demonstrated sensitivities of neoplasia detection of 98%. This system was based on macroscopic feature detection to allow the identification of areas suspicious for neoplasia. It should be noted that this is a relatively small number of images to use for a computer aided detection algorithm, particularly as it was split between two different subtypes of oesophageal cancer. The PPV of this system was also 40%, suggesting that the CNN may be disposed to over diagnosing cancers on the basis of shadows and other non-discriminative features.

Another system, developed by Nakagawa et al.²⁵³ most closely replicates our own. To develop their CNN, they used 5768 magnification endoscopy images, and 8660 non-magnified images of 804 ESCN lesions. Following training, 509 ME images from 155 patients were used to test the diagnostic performance of their CNN. Interestingly this group demonstrated the ability of their network to distinguish between mucosal and submucosal cancers with a promising degree of accuracy and sensitivity. For the detection of mucosal ESCN this study demonstrated accuracy and sensitivity for neoplasia detection of 92.3% and 93.7% respectively; for the detection of submucosal cancers this decreased to 89.7% and 91.5% respectively. Our CNN is not yet able to predict invasion depth of cancers; this was a conscious selection in methodology. We

opted to develop a robust and clinically interpretable CNN for the detection of dysplastic IPCLs in the first instance. Unlike the Nakagawa study we have produced CAMs which demonstrate what features the CNN is using to classify images as dysplastic; providing an additional layer of validation to our methodology. Secondly we opted to use a significantly larger dataset to first establish our networks ability to recognise IPCLs as normal or dysplastic. Based on our experience we would argue that such a small dataset would struggle to reliably generalise its predictions to a larger, unselected patient population, particularly when classifying over multiple patient subgroups. We have therefore aimed to ensure that our binary classification algorithm is of a substantial reliability and quality, before applying it to the more complex problem of multi-class classification of varying cancer invasion depths. While this is currently a major limitation of our CNN that would prevent its widespread clinical use, we aim to continue to develop a system, with larger patient numbers, capable of predicting the invasion depth of ESCN based on IPCL patterns.

This thesis has made frequent references to and acknowledges the fact that endoscopists practice in a variety of settings; expert endoscopists working in high volume centres may not be representative of the average endoscopist working in an average centre. In this study we benchmarked the performance of a group of Asian and European expert endoscopists for the detection of ESCN. The utility of the JES classification in Asian settings has been well characterised and validated, but less is known about its utility in a low volume European setting. We have demonstrated that using the JES classification expert European endoscopists achieve comparable levels of diagnostic accuracy to expert Asian endoscopists. This body of work therefore sets the benchmark for other studies assessing the JES classification, or those comparing the diagnostic performance of CNNs for classifying IPCLs patterns with clinicians in a European setting. The performance of our CNN matches that of some expert cohorts reported in the literature. In our study while the CNN performance was promising, it failed to match that of most expert endoscopists since their performance measures were so high.

Further work is needed to identify the performance of non-expert European and Asian endoscopists in the classification of abnormal IPCL patterns using the JES classification; this should then be compared with the performance of our CNN. This work may demonstrate that our CNN outperforms non-expert endoscopists. The implication of this finding is that our CNN may then have a role as a decision and diagnostic support tool for non-expert endoscopists assessing lesions that are suspicious for ESCN. Accurate characterisation of these lesions by a CNN could reduce the need for unnecessary biopsies and improve the triage of patients to prompt endoscopic or surgical therapies for their ESCN.

9.5 Conclusions

This thesis presents four studies which investigate technologies that could improve the endoscopic detection of early oesophageal neoplasia. We have successfully developed an early iteration of a convolutional neural network that can classify oesophageal mucosa as normal or dysplastic in real time. This work provides a framework for further studies which should aim to produce a clinically validated and applicable system to classify neoplastic tissue intra-endoscopy. This body of work has also assessed and validated a classification system for the detection of Barrett's oesophagus associated neoplasia using a new advanced endoscopic imaging system, as well as demonstrating a new imaging technology may be more effective than the use of HD-WLE alone.

Word count 47636

1. Ferlay, J. *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**, (2015).
2. Arnold, M., Soerjomataram, I., Ferlay, J. & Forman, D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* **64**, (2015).
3. Bytzer, P., Christensen, P. B., Damkier, P., Vinding, K. & Seersholm, N. Adenocarcinoma of the esophagus and Barrett's esophagus: a population-based study. *Am. J. Gastroenterol.* **94**, 86 (1999).
4. The histology guide - Leeds University. Layers of the oesophagus.
<https://www.histology.leeds.ac.uk/oral/oesophagus.php>.
5. Dartmouth College. Slide #DMS 127 Oesophageal mid portion. Dartmouth.edu pathology.
https://www.dartmouth.edu/~anatomy/Histo/lab_5/GI/DMS127/popup.html.
6. Rice, T. W., Rusch, V. W., Ishwaran, H., Blackstone, E. H., & Worldwide Esophageal Cancer Collaboration. Cancer of the esophagus and esophagogastric junction: data-driven staging for the seventh edition of the American Joint Committee on Cancer/International Union Against Cancer Cancer Staging Manuals. *Cancer* **116**, 3763–3773 (2010).
7. Howlader, N. *et al.* SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD. 2016-02-16]. <http://seer.cancer.gov/csr> (2016).
8. NCRAS; National Cancer Registration and Analysis Service. Stage Breakdown by CCG 2014. (2016).

9. National Cancer Registration and Analysis Service (NCRAS). Routes to diagnosis of cancer by stage, 2012-2013. (2016).
10. Howlader, N. *et al.* SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD. 2016-02-16 *Httpseer Cancer Govcsr* (2016).
11. Office of National Statistics. Cancer survival in England - adults diagnosed.
<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed>
(2013).
12. Zhang, Y. Epidemiology of esophageal cancer. *World J. Gastroenterol. WJG* **19**, 5598–5606 (2013).
13. Eslick, G. D. Epidemiology of Esophageal Cancer. *Gastroenterol. Clin. North Am.* **38**, 17–25 (2009).
14. Engel, L. S. *et al.* Population attributable risks of esophageal and gastric cancers. *J. Natl. Cancer Inst.* **95**, 1404–1413 (2003).
15. Jemal, A. *et al.* Global cancer statistics. *CA. Cancer J. Clin.* **61**, 69–90 (2011).
16. Tran, G. D. *et al.* Prospective study of risk factors for esophageal and gastric cancers in the Linxian general population trial cohort in China. *Int. J. Cancer* **113**, 456–463 (2005).
17. Nasrollahzadeh, D. *et al.* Opium, tobacco, and alcohol use in relation to oesophageal squamous cell carcinoma in a high-risk area of Iran. *Br. J. Cancer* **98**, 1857–1863 (2008).
18. Lepage, C., Rachet, B., Jooste, V., Faivre, J. & Coleman, M. P. Continuing Rapid Increase in Esophageal Adenocarcinoma in England and Wales. *Am. J. Gastroenterol.* **103**, 2694 (2008).
19. LORD, R. V. N. *et al.* Rising incidence of oesophageal adenocarcinoma in men in Australia. *J. Gastroenterol. Hepatol.* **13**, 356–362 (1998).

20. Dulai, G. S., Guha, S., Kahn, K. L., Gornbein, J. & Weinstein, W. M. Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: A systematic review. *Gastroenterology* **122**, 26–33 (2018).
21. Sikkema, M., De Jonge, P. J. F., Steyerberg, E. W. & Kuipers, E. J. Risk of esophageal adenocarcinoma and mortality in patients with Barrett's esophagus: a systematic review and meta-analysis. *Clin. Gastroenterol. Hepatol.* **8**, 235–244 (2010).
22. Chen, W. *et al.* Cancer statistics in China, 2015. *CA. Cancer J. Clin.* **66**, 115–132 (2016).
23. Abnet, C. C., Arnold, M. & Wei, W.-Q. Epidemiology of esophageal squamous cell carcinoma. *Gastroenterology* **154**, 360–373 (2018).
24. Steevens, J., Botterweck, A. A., Dirx, M. J., van den Brandt, P. A. & Schouten, L. J. Trends in incidence of oesophageal and stomach cancer subtypes in Europe. *Eur. J. Gastroenterol. Hepatol.* **22**, 669–678 (2010).
25. Cook, M., Chow, W.-H. & Devesa, S. Oesophageal cancer incidence in the United States by race, sex, and histologic type, 1977–2005. *Br. J. Cancer* **101**, 855–859 (2009).
26. Freedman, N. D. *et al.* A prospective study of tobacco, alcohol, and the risk of esophageal and gastric cancer subtypes. *Am. J. Epidemiol.* **165**, 1424–1433 (2007).
27. Ishiguro, S. *et al.* Effect of alcohol consumption, cigarette smoking and flushing response on esophageal cancer risk: a population-based cohort study (JPHC study). *Cancer Lett.* **275**, 240–246 (2009).
28. International Agency for Research on Cancer. *A review of human carcinogens: personal habits and indoor combustions*. vol. 100 (World Health Organization, 2012).
29. Lin, Y. *et al.* Epidemiology of esophageal cancer in Japan and China. *J. Epidemiol.* JE20120162 (2013).

30. Garidou, A. *et al.* Life-style factors and medical conditions in relation to esophageal cancer by histologic type in a low-risk population. *Int. J. Cancer* **68**, 295–299 (1996).
31. Franceschi, S. *et al.* Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy. *Cancer Res.* **50**, 6502–6507 (1990).
32. Lagergren, J., Bergström, R., Lindgren, A. & Nyrén, O. The role of tobacco, snuff and alcohol use in the aetiology of cancer of the oesophagus and gastric cardia. *Int. J. Cancer* **85**, 340–346 (2000).
33. Abedi-Ardekani, B. *et al.* Polycyclic aromatic hydrocarbon exposure in oesophageal tissue and risk of oesophageal squamous cell carcinoma in north-eastern Iran. *Gut* **59**, 1178–1183 (2010).
34. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Betel-quid and areca-nut chewing and some areca-nut derived nitrosamines. *IARC Monogr. Eval. Carcinog. Risks Hum.* **85**, 1 (2004).
35. Ren, J.-S., Kamangar, F., Forman, D. & Islami, F. Pickled food and risk of gastric cancer—a systematic review and meta-analysis of English and Chinese literature. *Cancer Epidemiol. Prev. Biomark.* **21**, 905–915 (2012).
36. Jeurnink, S. *et al.* Variety in vegetable and fruit consumption and the risk of gastric and esophageal cancer in the European prospective investigation into cancer and nutrition. *Int. J. Cancer* **131**, E963–E973 (2012).
37. Qiu, S. & Yang, G. Precursor lesions of esophageal cancer in high-risk populations in Henan Province, China. *Cancer* **62**, 551–557 (1988).
38. Rubio, C., Liu, F. & Zhao, H.-Z. Histological classification of intraepithelial neoplasias and microinvasive squamous carcinoma of the esophagus. *Am. J. Surg. Pathol.* **13**, 685–690 (1989).

39. Weisenberg, E. Squamous dysplasia. Pathology Outlines.
<http://www.pathologyoutlines.com/topic/esophagussquamousdysplasia.html>.
40. Dawsey, S. M. *et al.* Squamous esophageal histology and subsequent risk of squamous cell carcinoma of the esophagus. A prospective follow-up study from Linxian, China. *Cancer* **74**, 1686–1692 (1994).
41. Wang, G. Q. *et al.* Histological precursors of oesophageal squamous cell carcinoma: results from a 13 year prospective follow up study in a high risk population. *Gut* **54**, 187–192 (2005).
42. Dawsey, S. M. *et al.* Endoscopic and histologic progression of untreated squamous esophageal neoplasia in asymptomatic patients from Linxian, China: Implications for screening. in *Gastroenterology* vol. 112 A553–A553 (WB SAUNDERS CO-ELSEVIER INC 1600 JOHN F KENNEDY BOULEVARD, STE 1800, PHILADELPHIA, PA 19103-2899 USA, 1997).
43. Muñoz, N. *et al.* PRECURSOR LESIONS OF OESOPHAGEAL CANCER IN HIGH-RISK POPULATIONS IN IRAN AND CHINA. *The Lancet* **319**, 876–879 (2018).
44. Lu, X.-J. *et al.* Endoscopic survey of esophageal cancer in a high-risk area of China. *World J. Gastroenterol. WJG* **10**, 2931–2935 (2004).
45. He, Z. *et al.* Prevalence and risk factors for esophageal squamous cell cancer and precursor lesions in Anyang, China: a population-based endoscopic survey. *Br. J. Cancer* **103**, 1085–1088 (2010).
46. Etemadi, A., Abnet, C. C., Golozar, A., Malekzadeh, R. & Dawsey, S. M. Predicting esophageal squamous cell carcinoma and squamous dysplasia: risk modeling in a high risk area in Iran. *Arch. Iran. Med.* **15**, 18–21 (2012).

47. Dawsey, S. M. *et al.* Squamous dysplasia and early esophageal cancer in the Linxian region of China: distinctive endoscopic lesions. *Gastroenterology* **105**, 1333–1340 (1993).
48. Goda, K. *et al.* Narrow-band imaging magnifying endoscopy versus Lugol chromoendoscopy with pink-color sign assessment in the diagnosis of superficial esophageal squamous neoplasms: a randomised noninferiority trial. *Gastroenterol. Res. Pract.* **2015**, (2015).
49. Nagami, Y. *et al.* Usefulness of Non-Magnifying Narrow-Band Imaging in Screening of Early Esophageal Squamous Cell Carcinoma: A Prospective Comparative Study Using Propensity Score Matching. *Am. J. Gastroenterol.* **109**, 845–845 (2014).
50. Ide, E. *et al.* Endoscopic detection of early esophageal squamous cell carcinoma in patients with achalasia: narrow-band imaging versus Lugol's staining. *J. Oncol.* **2013**, (2013).
51. Morita, F. H. A. *et al.* Narrow band imaging versus lugol chromoendoscopy to diagnose squamous cell carcinoma of the esophagus: a systematic review and meta-analysis. *BMC Cancer* **17**, 54–54 (2017).
52. Takahashi, M. *et al.* Endoscopic diagnosis of early neoplasia of the esophagus with narrow band imaging: correlations among background coloration and iodine staining findings. *J. Gastroenterol. Hepatol.* **29**, 762–768 (2014).
53. Dawsey, S. M. *et al.* Mucosal iodine staining improves endoscopic visualization of squamous dysplasia and squamous cell carcinoma of the esophagus in Linxian, China. *Cancer* **83**, 220–231 (1998).

54. Hashimoto, C. L. *et al.* Lugol's Dye Spray Chromoendoscopy Establishes Early Diagnosis of Esophageal Cancer in Patients with Primary Head and Neck Cancer. *Am. J. Gastroenterol.* **100**, 275 (2005).
55. Shiozaki, H. *et al.* Endoscopic screening of early esophageal cancer with the Lugol dye method in patients with head and neck cancers. *Cancer* **66**, 2068–2071 (1990).
56. Shao, Y. *et al.* Lugol chromoendoscopic screening for esophageal dysplasia/early squamous cell carcinoma in patients with esophageal symptoms in low-risk region in China. *Oncol. Lett.* **10**, 45–50 (2015).
57. Chai, T.-H., Jin, X.-F., Li, S.-H., Du, R.-L. & Zhang, J. A tandem trial of HD-NBI versus HD-WL to compare neoplasia miss rates in esophageal squamous cell carcinoma. *Hepatogastroenterology.* **61**, 120–124 (2014).
58. Muto, M. *et al.* Early detection of superficial squamous cell carcinoma in the head and neck region and esophagus by narrow band imaging: a multicenter randomized controlled trial. *J. Clin. Oncol.* **28**, 1566 (2010).
59. Goda, K. *et al.* Magnifying endoscopy with narrow band imaging for predicting the invasion depth of superficial esophageal squamous cell carcinoma. *Dis. Esophagus* **22**, 453–460 (2009).
60. Sato, H. *et al.* Utility of intrapapillary capillary loops seen on magnifying narrow-band imaging in estimating invasive depth of esophageal squamous cell carcinoma. *Endoscopy* **47**, 122–128 (2015).
61. Guo, J. *et al.* Diagnostic value of probe-based confocal laser endomicroscopy and high-definition virtual chromoendoscopy in early esophageal squamous neoplasia. *Gastrointest. Endosc.* **81**, 1346–1354 (2015).

62. INOUE, H. *et al.* Ultra-high magnification endoscopy of the normal esophageal mucosa. *Dig. Endosc.* **8**, 134–138 (1996).
63. INOUE, H. *et al.* Ultra-high magnification endoscopic observation of carcinoma in situ of the esophagus. *Dig. Endosc.* **9**, 16–18 (1997).
64. Inoue, H. Magnification endoscopy in the esophagus and stomach. *Dig. Endosc.* **13**, (2001).
65. Goda, K. *et al.* Clinicopathological features of narrow-band imaging endoscopy and immunohistochemistry in ultraminute esophageal squamous neoplasms. *Dis. Esophagus* **27**, 267–275 (2014).
66. Arima, M., Tada, M. & Arima, H. Evaluation of microvascular patterns of superficial esophageal cancers by magnifying endoscopy. *Esophagus* **2**, 191–197 (2005).
67. Oyama, T. A new classification of magnified endoscopy for superficial esophageal squamous cell carcinoma. *Esophagus* **8**, 247–251 (2011).
68. Oyama, T., Momma, K. & Makuuchi, H. Japan esophageal society classification of superficial esophageal squamous cell carcinoma. *Endosc Dig* **24**, 466–468 (2012).
69. Oyama, T. *et al.* Prediction of the invasion depth of superficial squamous cell carcinoma based on microvessel morphology: magnifying endoscopic classification of the Japan Esophageal Society. *Esophagus* **14**, 105–112 (2017).
70. Kim, S. J. *et al.* New magnifying endoscopic classification for superficial esophageal squamous cell carcinoma. *World Journal of Gastroenterology* vol. 23 4416–4421 (Baishideng Publishing Group Inc, 2017).
71. Inoue, H. *et al.* Endoscopic mucosal resection with a cap-fitted panendoscope for esophagus, stomach, and colon mucosal lesions. *Gastrointest. Endosc.* **39**, 58–62 (1993).

72. Inoue, H. Endoscopic mucosal resection for esophageal and gastric mucosal cancers. *Can. J. Gastroenterol.* **12**, (1998).
73. Katada, C. *et al.* Clinical outcome after endoscopic mucosal resection for esophageal squamous cell carcinoma invading the muscularis mucosae-a multicenter retrospective cohort study. *Endoscopy* **39**, 779–783 (2007).
74. Yamashina, T. *et al.* Long-term outcome and metastatic risk after endoscopic resection of superficial esophageal squamous cell carcinoma. *Am. J. Gastroenterol.* **108**, 544–551 (2013).
75. He, S. *et al.* Endoscopic radiofrequency ablation for early esophageal squamous cell neoplasia: Report of safety and effectiveness from a large prospective trial. *Endoscopy* **47**, 398–408 (2015).
76. Yu, X. *et al.* Durability of radiofrequency ablation for treatment of esophageal squamous cell neoplasia: 5-year follow-up of a treated cohort in China. *Gastrointest. Endosc.* **89**, 736–748 (2019).
77. Wang, W.-L. *et al.* Lessons from pathological analysis of recurrent early esophageal squamous cell neoplasia after complete endoscopic radiofrequency ablation. *Endoscopy* **50**, 743–750 (2018).
78. Haidry, R. J. *et al.* Radiofrequency ablation for early oesophageal squamous neoplasia: outcomes form United Kingdom registry. *World J. Gastroenterol. WJG* **19**, 6011 (2013).
79. Oyama, T. *et al.* Endoscopic Submucosal Dissection of Early Esophageal Cancer. *Clin. Gastroenterol. Hepatol.* **3**, S67–S70 (2018).
80. Nagami, Y. *et al.* The five-year survival rate after endoscopic submucosal dissection for superficial esophageal squamous cell neoplasia. *Dig. Liver Dis.* **49**, 427–433 (2017).

81. Ono, S. *et al.* Long-term outcomes of endoscopic submucosal dissection for superficial esophageal squamous cell neoplasms. *Gastrointest. Endosc.* **70**, 860–866 (2009).
82. Berger, A. *et al.* Long-term follow-up after endoscopic resection for superficial esophageal squamous cell carcinoma: a multicenter Western study. *Endoscopy* **51**, 298–306 (2019).
83. Powell, J., McConkey, C. C., Gillison, E. W. & Spychal, R. T. Continuing rising trend in oesophageal adenocarcinoma. *Int. J. Cancer* **102**, 422–427 (2002).
84. Levi, F. & La Vecchia, C. Adenocarcinoma of the esophagus in Switzerland. *Jama* **265**, 2960–2960 (1991).
85. Møller, H. Incidence of cancer of oesophagus, cardia and stomach in Denmark. *Eur. J. Cancer Prev.* **1**, 159–164 (1992).
86. Edgren, G., Adami, H.-O., Weiderpass, E. & Nyrén, O. A global assessment of the oesophageal adenocarcinoma epidemic. *Gut* **62**, 1406–1414 (2013).
87. Xie, S.-H. & Lagergren, J. The male predominance in esophageal adenocarcinoma. *Clin. Gastroenterol. Hepatol.* **14**, 338–347 (2016).
88. Blot, W. J., Devesa, S. S., Kneller, R. W. & Fraumeni, J. F. Rising incidence of adenocarcinoma of the esophagus and gastric cardia. *Jama* **265**, 1287–1289 (1991).
89. Yang, P. C. & Davis, S. Incidence of cancer of the esophagus in the US by histologic type. *Cancer* **61**, 612–617 (1988).
90. Devesa, S. S., Blot, W. J. & Fraumeni Jr, J. F. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. *Cancer Interdiscip. Int. J. Am. Cancer Soc.* **83**, 2049–2053 (1998).

91. Lagergren, J., Bergström, R., Lindgren, A. & Nyrén, O. Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *N. Engl. J. Med.* **340**, 825–831 (1999).
92. Cook, D. J. *et al.* Stress ulcer prophylaxis in critically ill patients: resolving discordant meta-analyses. *Jama* **275**, 308–314 (1996).
93. Whiteman, D. C. *et al.* Combined effects of obesity, acid reflux and smoking on the risk of adenocarcinomas of the oesophagus. *Gut* **57**, 173–180 (2008).
94. Lagergren, J., Bergström, R. & Nyrén, O. Association between body mass and adenocarcinoma of the esophagus and gastric cardia. *Ann. Intern. Med.* **130**, 883–890 (1999).
95. Hoyo, C. *et al.* Body mass index in relation to oesophageal and oesophagogastric junction adenocarcinomas: a pooled analysis from the International BEACON Consortium. *Int. J. Epidemiol.* **41**, 1706–1718 (2012).
96. Turati, F., Tramacere, I., La Vecchia, C. & Negri, E. A meta-analysis of body mass index and esophageal and gastric cardia adenocarcinoma. *Ann. Oncol.* **24**, 609–617 (2013).
97. Petrick, J. L. *et al.* Body weight trajectories and risk of oesophageal and gastric cardia adenocarcinomas: a pooled analysis of NIH-AARP and PLCO Studies. *Br. J. Cancer* **116**, 951–959 (2017).
98. Coleman, H. G., Xie, S. H. & Lagergren, J. The Epidemiology of Esophageal Adenocarcinoma. *Gastroenterology* **154**, (2018).
99. Cook, M. B. *et al.* Cigarette smoking and adenocarcinomas of the esophagus and esophagogastric junction: a pooled analysis from the international BEACON consortium. *J. Natl. Cancer Inst.* **102**, 1344–1353 (2010).

100. Coleman, H. G. *et al.* Tobacco smoking increases the risk of high-grade dysplasia and cancer among patients with Barrett's esophagus. *Gastroenterology* **142**, 233–240 (2012).
101. Pohl, H. *et al.* Risk factors in the development of esophageal adenocarcinoma. *Am. J. Gastroenterol.* **108**, 200–207 (2013).
102. Nie, S., Chen, T., Yang, X., Huai, P. & Lu, M. Association of helicobacter pylori infection with esophageal adenocarcinoma and squamous cell carcinoma: a meta-analysis. *Dis. Esophagus* **27**, 645–653 (2014).
103. Xie, F.-J. *et al.* Helicobacter pylori infection and esophageal cancer risk: an updated meta-analysis. *World J. Gastroenterol. WJG* **19**, 6098 (2013).
104. Islami, F. & Kamangar, F. Helicobacter pylori and esophageal cancer risk: a meta-analysis. *Cancer Prev. Res. (Phila. Pa.)* **1**, 329–338 (2008).
105. Barrett, N. R. Chronic peptic ulcerz of the œophagus and 'œsophagitis'. *Br. J. Surg.* **38**, 175–182 (1950).
106. Barrett, N. R. The lower esophagus lined by columnar epithelium. *Surgery* **41**, 881–894 (1957).
107. Naef, A., Savary, M., Ozzello, L. & Pearson, F. G. Columnar-lined lower esophagus: an acquired lesion with malignant predisposition: report on 140 cases of Barrett's esophagus with 12 adenocarcinomas. *J. Thorac. Cardiovasc. Surg.* **70**, 826–835 (1975).
108. Fitzgerald, R. C. *et al.* British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* **63**, 7 LP – 42 (2014).
109. Shaheen, N. J., Falk, G. W., Iyer, P. G. & Gerson, L. B. Corrigendum: ACG Clinical Guideline: Diagnosis and Management of Barrett's Esophagus. *Am. J. Gastroenterol.* **111**, 1077 (2016).

110. Bennett, C. *et al.* BOB CAT: a large-scale review and Delphi consensus for management of Barrett's esophagus with no dysplasia, indefinite for, or low-grade dysplasia. *Am. J. Gastroenterol.* **110**, 662 (2015).
111. Chandrasoma, P. *et al.* Columnar-lined esophagus without intestinal metaplasia has no proven risk of adenocarcinoma. *Am. J. Surg. Pathol.* **36**, 1–7 (2012).
112. Kelty, C. J., Gough, M. D., Van Wyk, Q., Stephenson, T. J. & Ackroyd, R. Barrett's oesophagus: intestinal metaplasia is not essential for cancer risk. *Scand. J. Gastroenterol.* **42**, 1271–1274 (2007).
113. Ronkainen, J. *et al.* Prevalence of Barrett's Esophagus in the General Population: An Endoscopic Study. *Gastroenterology* **129**, 1825–1831 (2017).
114. Zagari, R. M. *et al.* Gastro-oesophageal reflux symptoms, oesophagitis and Barrett's oesophagus in the general population: the Loiano–Monghidoro study. *Gut* **57**, 1354–1359 (2008).
115. Yousef, F. *et al.* The incidence of esophageal cancer and high-grade dysplasia in Barrett's esophagus: a systematic review and meta-analysis. *Am. J. Epidemiol.* **168**, 237–249 (2008).
116. Hvid-Jensen, F., Pedersen, L., Drewes, A. M., Sørensen, H. T. & Funch-Jensen, P. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N. Engl. J. Med.* **365**, 1375–1383 (2011).
117. Desai, T. K. *et al.* The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut* **61**, 970–976 (2012).
118. Wani, S. *et al.* Patients with nondysplastic Barrett's esophagus have low risks for developing dysplasia or esophageal adenocarcinoma. *Clin. Gastroenterol. Hepatol.* **9**, 220–227 (2011).

119. Hvid-Jensen, F., Pedersen, L., Drewes, A. M., Sørensen, H. T. & Funch-Jensen, P. Incidence of adenocarcinoma among patients with Barrett's esophagus. *N. Engl. J. Med.* **365**, 1375–1383 (2011).
120. Schlemper, R. J. *et al.* The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* **47**, 251–255 (2000).
121. Voltaggio, L. & Montgomery, E. A. Diagnosis and management of Barrett-related neoplasia in the modern era. *Surg. Pathol. Clin.* **10**, 781–800 (2017).
122. Karamchandani, D. Barrett esophagus. Pathology Outlines.
<http://www.pathologyoutlines.com/topic/esophagusBarrettsgeneral.html>.
123. Duits, L. C. *et al.* Barrett's oesophagus patients with low-grade dysplasia can be accurately risk-stratified after histological review by an expert pathology panel. *Gut* *gutjnl-2014* (2014).
124. Reid, B. J., Levine, D. S., Longton, G., Blount, P. L. & Rabinovitch, P. S. Predictors of progression to cancer in Barrett's esophagus: baseline histology and flow cytometry identify low-and high-risk patient subsets. *Am. J. Gastroenterol.* **95**, 1669 (2000).
125. Sharma, P. *et al.* Dysplasia and cancer in a large multicenter cohort of patients with Barrett's esophagus. *Clin. Gastroenterol. Hepatol.* **4**, 566–572 (2006).
126. Shaheen, N. J. *et al.* Durability of Radiofrequency Ablation in Barrett's Esophagus with Dysplasia. *Gastroenterology* **141**, 460–468 (2011).
127. Spechler, S. J., Sharma, P., Souza, R. F., Inadomi, J. M. & Shaheen, N. J. American Gastroenterological Association technical review on the management of Barrett's esophagus. *Gastroenterology* **140**, e18–e52 (2011).
128. Spechler, S. J. Screening and surveillance for Barrett's esophagus—an unresolved dilemma. *Nat. Clin. Pract. Gastroenterol. Hepatol.* **4**, 470–471 (2007).

129. Wang, K. K. & Sampliner, R. E. Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *Am. J. Gastroenterol.* **103**, 788–797 (2008).
130. Kahrilas, P. J., Shaheen, N. J. & Vaezi, M. F. American Gastroenterological Association Institute technical review on the management of gastroesophageal reflux disease. *Gastroenterology* **135**, 1392–1413 (2008).
131. Corley, D. A. *et al.* Impact of Endoscopic Surveillance on Mortality From Barrett's Esophagus–Associated Esophageal Adenocarcinomas. *Gastroenterology* **145**, 312–319.e1 (2013).
132. Ferguson, M. K. & Durkin, A. Long-term survival after esophagectomy for Barrett's adenocarcinoma in endoscopically surveyed and nonsurveyed patients. *J. Gastrointest. Surg.* **6**, 29–36 (2002).
133. Fountoulakis, A. *et al.* Effect of surveillance of Barrett's oesophagus on the clinical outcome of oesophageal cancer. *Br. J. Surg.* **91**, 997–1003 (2004).
134. Rubenstein, J. H., Sonnenberg, A., Davis, J., McMahon, L. & Inadomi, J. M. Effect of a prior endoscopy on outcomes of esophageal adenocarcinoma among United States veterans. *Gastrointest. Endosc.* **68**, 849–855 (2008).
135. McCabe, C., Claxton, K. & Culyer, A. J. The NICE cost-effectiveness threshold. *Pharmacoeconomics* **26**, 733–744 (2008).
136. Kastelein, F. *et al.* Surveillance in patients with long-segment Barrett's oesophagus: a cost-effectiveness analysis. *Gut* **64**, 864–871 (2015).
137. Hirst, N. G., Gordon, L. G., Whiteman, D. C., Watson, D. I. & Barendregt, J. J. Is endoscopic surveillance for non-dysplastic Barrett's esophagus cost-effective? Review of economic evaluations. *J. Gastroenterol. Hepatol.* **26**, 247–254 (2011).

138. Gordon, L. G. *et al.* Cost-effectiveness of endoscopic surveillance of non-dysplastic Barrett's esophagus. *Gastrointest. Endosc.* **79**, 242-256.e6 (2014).
139. Inadomi, J. M. *et al.* Screening and surveillance for Barrett esophagus in high-risk groups: a costutility analysis. *Ann. Intern. Med.* **138**, 176–186 (2003).
140. Everson, M. A., Ragunath, K., Bhandari, P., Lovat, L. & Haidry, R. How to Perform a High-Quality Examination in Patients With Barrett's Esophagus. *Gastroenterology* (2018) doi:10.1053/j.gastro.2018.03.001.
141. Sharma, P. *et al.* The Development and Validation of an Endoscopic Grading System for Barrett's Esophagus: The Prague C & M Criteria. *Gastroenterology* **131**, 1392–1399 (2017).
142. Workshop, P. in the P. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon. *Gastrointest. Endosc.* **58**, S3–S43 (2017).
143. Axon, A. *et al.* Update on the Paris classification of superficial neoplastic lesions in the digestive tract. *Endoscopy* **37**, 570–578 (2005).
144. van Putten, M. *et al.* 'Missed' oesophageal adenocarcinoma and high-grade dysplasia in Barrett's oesophagus patients: A large population-based study. *United Eur. Gastroenterol. J.* **6**, 519–528 (2018).
145. Visrodia, K. *et al.* Magnitude of Missed Esophageal Adenocarcinoma After Barrett's Esophagus Diagnosis: A Systematic Review and Meta-analysis. *Gastroenterology* **150**, 599-e15 (2016).
146. Tschanz, E. R. Do 40% of Patients Resected for Barrett Esophagus With High-Grade Dysplasia Have Unsuspected Adenocarcinoma? *Arch. Pathol. Lab. Med.* **129**, 177–180 (2005).

147. Gupta, N. *et al.* Longer inspection time is associated with increased detection of high-grade dysplasia and esophageal adenocarcinoma in Barrett's esophagus. *Gastrointest. Endosc.* **76**, 531–538 (2017).
148. ABRAMS, J. A. *et al.* Adherence to Biopsy Guidelines for Barrett's Esophagus Surveillance in the Community Setting in the United States. *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.* **7**, 710–736 (2009).
149. Canto, M. I. F. *et al.* Methylene blue-directed biopsies improve detection of intestinal metaplasia and dysplasia in Barrett's esophagus. *Gastrointest. Endosc.* **51**, 560–568 (2017).
150. Ngamruengphong, S., Sharma, V. K. & Das, A. Diagnostic yield of methylene blue chromoendoscopy for detecting specialized intestinal metaplasia and dysplasia in Barrett's esophagus: a meta-analysis. *Gastrointest. Endosc.* **69**, 1021–1028 (2009).
151. Longcroft-Wheaton, G., Duku, M., Mead, R., Poller, D. & Bhandari, P. Acetic Acid Spray Is an Effective Tool for the Endoscopic Detection of Neoplasia in Patients With Barrett's Esophagus. *Clin. Gastroenterol. Hepatol.* **8**, 843–847 (2017).
152. Bhandari, P., Kandaswamy, P., Cowlshaw, D. & Longcroft-Wheaton, G. Acetic acid-enhanced chromoendoscopy is more cost-effective than protocol-guided biopsies in a high-risk Barrett's population. *Dis. Esophagus* **25**, 386–392 (2012).
153. Coletta, M. *et al.* Acetic acid chromoendoscopy for the diagnosis of early neoplasia and specialized intestinal metaplasia in Barrett's esophagus: a meta-analysis. *Gastrointest. Endosc.* **83**, 57-67.e1 (2016).
154. Curvers, W. *et al.* Chromoendoscopy and Narrow-Band Imaging Compared With High-Resolution Magnification Endoscopy in Barrett's Esophagus. *Gastroenterology* **134**, 670–679 (2017).

155. Kandiah, K. *et al.* International development and validation of a classification system for the identification of Barrett's neoplasia using acetic acid chromoendoscopy: the Portsmouth acetic acid classification (PREDICT). *Gut* **67**, 2085 (2018).
156. Gono, K. *et al.* Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J. Biomed. Opt.* **9**, 568–577 (2004).
157. Sharma, P. *et al.* The utility of a novel narrow band imaging endoscopy system in patients with Barrett's esophagus. *Gastrointest. Endosc.* **64**, 167–175 (2017).
158. Kara, M. A., Ennahachi, M., Fockens, P., ten Kate, F. J. W. & Bergman, J. J. G. H. M. Detection and classification of the mucosal and vascular patterns (mucosal morphology) in Barrett's esophagus by using narrow band imaging. *Gastrointest. Endosc.* **64**, 155–166 (2017).
159. Herrero, L. A. *et al.* Zooming in on Barrett oesophagus using narrow-band imaging: an international observer agreement study. *Eur. J. Gastroenterol. Hepatol.* **21**, 1068–1075 (2009).
160. Mannath, J., Subramanian, V., Hawkey, C. J. & Ragunath, K. Narrow band imaging for characterization of high grade dysplasia and specialized intestinal metaplasia in Barrett's esophagus: a meta-analysis. *Endoscopy* **42**, 351–359 (2010).
161. Sharma, P. *et al.* Development and Validation of a Classification System to Identify High-Grade Dysplasia and Esophageal Adenocarcinoma in Barrett's Esophagus Using Narrow-Band Imaging. *Gastroenterology* **150**, 591–598 (2017).
162. Helmutt, N., Iaccuci, M. & Robles-Medrand, C. No Title. *Special report: iScan and the new OE optical enhancement technology in the GI tract*
https://cdn2.hubspot.net/hubfs/2164988/Paper_Content/Images/i-SCAN and OE Special Report.pdf (2016).

163. Kodashima, S. & Fujishiro, M. Novel image-enhanced endoscopy with i-scan technology. *World J. Gastroenterol. WJG* **16**, 1043–1049 (2010).
164. Lipman, G. *et al.* Systematic assessment with I-SCAN magnification endoscopy and acetic acid improves dysplasia detection in patients with Barrett's esophagus. *Endoscopy* (2017). doi:10.1055/s-0043-113441.
165. Verna, C. *et al.* I-SCAN targeted versus random biopsies in Barrett's oesophagus. *Dig. Liver Dis.* **46**, 131–134 (2014).
166. Togashi, K. *et al.* Blue laser imaging endoscopy system for the early detection and characterization of colorectal lesions: a guide for the endoscopist. *Ther. Adv. Gastroenterol.* **9**, 50–56 (2016).
167. Subramaniam, S. *et al.* Development and validation of the international Blue Light Imaging for Barrett's Neoplasia Classification. *Gastrointest. Endosc.* **91**, 310–320 (2020).
168. Overholt, B. F. *et al.* Photodynamic therapy with porfimer sodium for ablation of high-grade dysplasia in Barrett's esophagus: international, partially blinded, randomized phase III trial. *Gastrointest. Endosc.* **62**, 488–498 (2005).
169. Prasad, G. A. *et al.* Long-term survival following endoscopic and surgical treatment of high-grade dysplasia in Barrett's esophagus. *Gastroenterology* **132**, 1226–1233 (2007).
170. Shaheen, N. J. *et al.* Radiofrequency Ablation in Barrett's Esophagus with Dysplasia. *N. Engl. J. Med.* **360**, 2277–2288 (2009).
171. Haidry, R. & Lovat, L. Long-term durability of radiofrequency ablation for Barrett's-related neoplasia. *Curr. Opin. Gastroenterol.* **31**, (2015).
172. Shaheen, N. J. *et al.* Safety and efficacy of endoscopic spray cryotherapy for Barrett's esophagus with high-grade dysplasia. *Gastrointest. Endosc.* **71**, 680–685 (2010).

173. Gosain, S., Mercer, K., Twaddell, W. S., Uradomo, L. & Greenwald, B. D. Liquid nitrogen spray cryotherapy in Barrett's esophagus with high-grade dysplasia: long-term results. *Gastrointest. Endosc.* **78**, 260–265 (2013).
174. Alzoubaidi, D. *et al.* Comparison of two multiband mucosectomy devices for endoscopic resection of Barrett's esophagus-related neoplasia. *Surg. Endosc.* 1–8.
175. Pouw, R. E. *et al.* Stepwise radical endoscopic resection for eradication of Barrett's oesophagus with early neoplasia in a cohort of 169 patients. *Gut* **59**, 1169–1177 (2010).
176. Ell, C. *et al.* Curative endoscopic resection of early esophageal adenocarcinomas (Barrett's cancer). *Gastrointest. Endosc.* **65**, 3–10 (2007).
177. Chennat, J. *et al.* Complete Barrett's eradication endoscopic mucosal resection: an effective treatment modality for high-grade dysplasia and intramucosal carcinoma—an American single-center experience. *Am. J. Gastroenterol.* **104**, 2684–2692 (2009).
178. Van Vilsteren, F. *et al.* Learning to perform endoscopic resection of esophageal neoplasia is associated with significant complications even within a structured training program. *Endoscopy* **44**, 4–14 (2012).
179. Konda, V. J. *et al.* Complete endoscopic mucosal resection is effective and durable treatment for Barrett's-associated neoplasia. *Clin. Gastroenterol. Hepatol.* **12**, 2002–2010 (2014).
180. Tomizawa, Y. *et al.* Safety of endoscopic mucosal resection for Barrett's esophagus. *Am. J. Gastroenterol.* **108**, (2013).
181. Small, A. J. *et al.* Comparative risk of recurrence of dysplasia and carcinoma after endoluminal eradication therapy of high-grade dysplasia versus intramucosal carcinoma in Barrett's esophagus. *Gastrointest. Endosc.* **81**, 1158–1166 (2015).

182. Pech, O. *et al.* Long-term results and risk factor analysis for recurrence after curative endoscopic therapy in 349 patients with high-grade intraepithelial neoplasia and mucosal adenocarcinoma in Barrett's oesophagus. *Gut* **57**, 1200–1206 (2008).
183. Tomizawa, Y., Konda, V., Coronel, E., Chapman, C. & Siddiqui, U. Efficacy, Durability, and Safety of Complete Endoscopic Mucosal Resection of Barrett Esophagus: A Systematic Review and Meta-Analysis. *J. Clin. Gastroenterol.* **52**, 210–216 (2018).
184. Yang, D. *et al.* Endoscopic submucosal dissection for early Barrett's neoplasia: a meta-analysis. *Gastrointest. Endosc.* **87**, 1383–1393 (2018).
185. Guo, H.-M., Zhang, X.-Q., Chen, M., Huang, S.-L. & Zou, X.-P. Endoscopic submucosal dissection vs endoscopic mucosal resection for superficial esophageal cancer. *World J. Gastroenterol. WJG* **20**, 5540 (2014).
186. Komeda, Y., Bruno, M. & Koch, A. EMR is not inferior to ESD for early Barrett's and EGJ neoplasia: an extensive review on outcome, recurrence and complication rates. *Endosc. Int. Open* **2**, E58 (2014).
187. Fitzgerald, R. C. *et al.* British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* **63**, 7 LP – 42 (2014).
188. NICE. CG106 Barrett's Oesophagus: ablative therapy.
<https://www.nice.org.uk/guidance/cg106/chapter/1-Guidance>.
189. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106 (1962).
190. Shahidi, N. & Bourke, M. J. Can artificial intelligence accurately diagnose endoscopically curable gastrointestinal cancers? *Tech. Gastrointest. Endosc.* 150639 (2019).

191. Zhang, C. *et al.* Tu1217 The Use of Convolutional Neural Artificial Intelligence Network to Aid the Diagnosis and Classification of Early Esophageal Neoplasia. A Feasibility Study. *Gastrointest. Endosc.* **85**, AB587–AB588 (2018).
192. Horie, Y. *et al.* Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest. Endosc.* **89**, 25–32 (2019).
193. Zhao, Y.-Y. *et al.* Computer-assisted diagnosis of early esophageal squamous cell carcinoma using narrow-band imaging magnifying endoscopy. *Endoscopy* **51**, 333–341 (2019).
194. Oyama, T. A new classification of magnified endoscopy for superficial esophageal squamous cell carcinoma. *Esophagus* **8**, 247–251 (2011).
195. NICE (2015) Suspected cancer: recognition and referral (NICE guideline). National Institute for Health and Care Excellence. (2015).
196. Shenbagaraj, L. *et al.* Endoscopy in 2017: a national survey of practice in the UK. *Frontline Gastroenterol.* flgastro-2018 (2018).
197. Kapoor, N., Bassi, A., Sturgess, R. & Bodger, K. Predictive value of alarm features in a rapid access upper gastrointestinal cancer service. *Gut* **54**, 40–45 (2005).
198. Cheung, D., Menon, S. & Trudgill, N. OC-012 How Commonly is Oesophageal Cancer Missed at Endoscopy (A UK Primary Care Based Study)? *Gut* **62**, A5 LP-A6 (2013).
199. Menon, S. & Trudgill, N. How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis. *Endosc. Int. Open* **2**, E46–E50 (2014).
200. Phoa, K. N. *et al.* Multimodality endoscopic eradication for neoplastic Barrett oesophagus: Results of an European multicentre study (EURO-II). *Gut* **65**, (2016).

201. ABRAMS, J. A. *et al.* Adherence to Biopsy Guidelines for Barrett's Esophagus Surveillance in the Community Setting in the United States. *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.* **7**, 710–736 (2009).
202. Gordon, L. G. & Mayne, G. C. Cost-effectiveness of Barrett's oesophagus screening and surveillance. *Best Pract. Res. Clin. Gastroenterol.* **27**, 893–903 (2013).
203. Van der Burgh, A., Dees, J., Hop, W. & Van Blankenstein, M. Oesophageal cancer is an uncommon cause of death in patients with Barrett's oesophagus. *Gut* **39**, 5–8 (1996).
204. Macdonald, C. E., Wicks, A. C. & Playford, R. J. Final results from 10 year cohort of patients undergoing surveillance for Barrett's oesophagus: observational study. *Bmj* **321**, 1252–1255 (2000).
205. Hameeteman, W., Tytgat, G. N. J., Houthoff, H. J. & van den Tweel, J. G. Barrett's Esophagus; Development of Dysplasia and Adenocarcinoma. *Gastroenterology* **96**, 1249–1256 (1989).
206. Bell, S. P. & Dutta, A. DNA Replication in Eukaryotic Cells. *Annu. Rev. Biochem.* **71**, 333–374 (2002).
207. Sakwe, A. M., Nguyen, T., Athanasopoulos, V., Shire, K. & Frappier, L. Identification and Characterization of a Novel Component of the Human Minichromosome Maintenance Complex. *Mol. Cell. Biol.* **27**, 3044–3055 (2007).
208. Maiorano, D., Lutzmann, M. & Méchali, M. MCM proteins and DNA replication. *Curr. Opin. Cell Biol.* **18**, 130–136 (2006).
209. Williams, G. H. *et al.* Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14932–14937 (1998).

210. Stoeber, K. *et al.* Immunoassay for urothelial cancers that detects DNA replication protein Mcm5 in urine. *The Lancet* **354**, 1524–1525 (1999).
211. Stoeber, K. *et al.* Diagnosis of genito-urinary tract cancer by detection of minichromosome maintenance 5 protein in urine sediments. *J. Natl. Cancer Inst.* **94**, 1071–1079 (2002).
212. Going, J. J. *et al.* Aberrant expression of minichromosome maintenance proteins 2 and 5, and Ki-67 in dysplastic squamous oesophageal epithelium and Barrett’s mucosa. *Gut* **50**, 373–377 (2002).
213. Williams, G. H. *et al.* Diagnosis of oesophageal cancer by detection of minichromosome maintenance 5 protein in gastric aspirates. *Br. J. Cancer* **91**, 714 (2004).
214. Skacel, M. *et al.* The diagnosis of low-grade dysplasia in Barrett’s esophagus and its implications for disease progression. *Am. J. Gastroenterol.* **95**, 3383–3383 (2000).
215. Vennalaganti, P. *et al.* Discordance Among Pathologists in the United States and Europe in Diagnosis of Low-Grade Dysplasia for Patients With Barrett’s Esophagus. *Gastroenterology* **152**, 564-570.e4 (2017).
216. Stoeber, K. *et al.* Immunoassay for urothelial cancers that detects DNA replication protein Mcm5 in urine. *The Lancet* **354**, 1524–1525 (2018).
217. Tschanz, E. R. Do 40% of Patients Resected for Barrett Esophagus With High-Grade Dysplasia Have Unsuspected Adenocarcinoma? *Arch. Pathol. Lab. Med.* **129**, 177–180 (2005).
218. Phoa, K., Fl, van V., BM, W. & al, et. Radiofrequency ablation vs endoscopic surveillance for patients with barrett esophagus and low-grade dysplasia: A randomized clinical trial. *JAMA* **311**, 1209–1217 (2014).

219. Levine, D. S., Blount, P. L., Rudolph, R. E. & Reid, B. J. Safety of a systematic endoscopic biopsy protocol in patients with Barrett's esophagus. *Am. J. Gastroenterol.* **95**, 1152–1152 (2000).
220. Visrodia, K. *et al.* Magnitude of Missed Esophageal Adenocarcinoma After Barrett's Esophagus Diagnosis: A Systematic Review and Meta-analysis. *Gastroenterology* **150**, 599–e15 (2016).
221. Sharma, P. *et al.* Development and Validation of a Classification System to Identify High-Grade Dysplasia and Esophageal Adenocarcinoma in Barrett's Esophagus Using Narrow-Band Imaging. *Gastroenterology* **150**, 591–598 (2017).
222. Kandiah, K. *et al.* International development and validation of a classification system for the identification of Barrett's neoplasia using acetic acid chromoendoscopy: the Portsmouth acetic acid classification (PREDICT). *Gut* (2017).
223. Subramaniam, S. *et al.* OC-068 Blue light imaging for barrett's neoplasia classification (blinc): the development and validation of a new endoscopic classification system to identify barrett's neoplasia. *Gut* **66**, A36 LP-A37 (2017).
224. Lipman, G. *et al.* Systematic assessment with I-SCAN magnification endoscopy and acetic acid improves dysplasia detection in patients with Barrett's esophagus. *Endoscopy* (2017). doi:10.1055/s-0043-113441.
225. Nogales, O. *et al.* Usefulness of Non-magnifying Narrow Band Imaging in EVIS EXERA III Video Systems and High-Definition Endoscopes to Diagnose Dysplasia in Barrett's Esophagus Using the Barrett International NBI Group (BING) Classification. *Dig. Dis. Sci.* **62**, 2840–2846 (2017).

226. Eusebi, L. H., Cirota, G. G., Zagari, R. M. & Ford, A. C. Global prevalence of Barrett's oesophagus and oesophageal cancer in individuals with gastro-oesophageal reflux: a systematic review and meta-analysis. *Gut* **70**, 456–463 (2021).
227. Sharma, P. *et al.* The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on imaging in Barrett's Esophagus. *Gastrointest. Endosc.* **76**, 252–254 (2017).
228. Zhang, Y. Epidemiology of esophageal cancer. *World J. Gastroenterol. WJG* **19**, 5598–5606 (2013).
229. Cho, J. W. *et al.* Lymph Node Metastases in Esophageal Carcinoma: An Endoscopist's View. *Clin. Endosc.* **47**, 523–529 (2014).
230. Shimada, H. *et al.* Impact of the Number and Extent of Positive Lymph Nodes in 200 Patients with Thoracic Esophageal Squamous Cell Carcinoma after Three-field Lymph Node Dissection. *World J. Surg.* **30**, 1441–1449 (2006).
231. Kodama, M. & Kakegawa, T. Treatment of superficial cancer of the esophagus: a summary of responses to a questionnaire on superficial cancer of the esophagus in Japan. *Surgery* **123**, 432–439 (1998).
232. Sgourakis, G., Gockel, I. & Lang, H. Endoscopic and surgical resection of T1a/T1b esophageal neoplasms: A systematic review. *World J. Gastroenterol. WJG* **19**, 1424–1437 (2013).
233. Sgourakis, G. *et al.* Detection of lymph node metastases in esophageal cancer. *Expert Rev. Anticancer Ther.* **11**, 601–612 (2011).
234. Inoue, H. Magnification endoscopy in the esophagus and stomach. *Dig. Endosc.* **13**, (2001).

235. Sato, H. *et al.* Utility of intrapapillary capillary loops seen on magnifying narrow-band imaging in estimating invasive depth of esophageal squamous cell carcinoma. *Endoscopy* **47**, 122–128 (2015).
236. Oyama, T. *et al.* Prediction of the invasion depth of superficial squamous cell carcinoma based on microvessel morphology: magnifying endoscopic classification of the Japan Esophageal Society. *Esophagus* **14**, 105–112 (2017).
237. Kaga, M., Inoue, H., Kudo, S.-E. & Hamatani, S. Microvascular architecture of early esophageal neoplasia. *Oncol. Rep.* **26**, 1063–1067 (2011).
238. Inoue, H. *et al.* Magnification endoscopy in esophageal squamous cell carcinoma: a review of the intrapapillary capillary loop classification. *Ann. Gastroenterol. Q. Publ. Hell. Soc. Gastroenterol.* **28**, 41–48 (2015).
239. Arima, M., Tada, M. & Arima, H. Evaluation of microvascular patterns of superficial esophageal cancers by magnifying endoscopy. *Esophagus* **2**, 191–197 (2005).
240. Kumagai, Y., Toi, M., Kawada, K. & Kawano, T. Angiogenesis in superficial esophageal squamous cell carcinoma: magnifying endoscopic observation and molecular analysis. *Dig. Endosc.* **22**, 259–267 (2010).
241. Gono, K. *et al.* Appearance of enhanced tissue features in narrow-band endoscopic imaging. *J. Biomed. Opt.* **9**, 568–577 (2004).
242. Kim, S. J. *et al.* New magnifying endoscopic classification for superficial esophageal squamous cell carcinoma. *World J. Gastroenterol.* **23**, 4416–4421 (2017).
243. Everson, M. *et al.* Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. *United Eur. Gastroenterol. J.* **7**, 297–306 (2019).

244. Schlemper, R. J. *et al.* The Vienna classification of gastrointestinal epithelial neoplasia. *Gut* **47**, 251–255 (2000).
245. Garcia-Peraza-Herrera, L. *et al.* Interpretable Fully Convolutional Classification of Intrapapillary Capillary Loops for Real-Time Detection of Early Squamous Neoplasia. *Arxiv* (2018).
246. Shimizu, Y. *et al.* Long-term outcome after endoscopic mucosal resection in patients with esophageal squamous cell carcinoma invading the muscularis mucosae or deeper. *Gastrointest. Endosc.* **56**, 387–390 (2018).
247. Wang, W.-L. *et al.* A training program of a new simplified classification of magnified narrow band imaging for superficial esophageal squamous cell carcinoma. *J. Gastroenterol. Hepatol.* n/a-n/a doi:10.1111/jgh.14071.
248. Bollschweiler, E. *et al.* High rate of lymph-node metastasis in submucosal esophageal squamous-cell carcinomas and adenocarcinomas. *Endoscopy* **38**, 149–156 (2006).
249. García-Peraza-Herrera, L. C. *et al.* Intrapapillary capillary loop classification in magnification endoscopy: open dataset and baseline methodology. *Int. J. Comput. Assist. Radiol. Surg.* **15**, 651–659 (2020).
250. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. & Torralba, A. Learning deep features for discriminative localization. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 2921–2929 (2016).
251. Guo, L. *et al.* Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest. Endosc.* (2019).

252. Mizumoto, T. *et al.* Diagnosis of superficial esophageal squamous cell carcinoma invasion depth before endoscopic submucosal dissection. *Dis. Esophagus* **31**, dox142–dox142 (2018).
253. Nakagawa, K. *et al.* Classification for invasion depth of esophageal squamous cell carcinoma using a deep neural network compared with experienced endoscopists. *Gastrointest. Endosc.* **90**, 407–414 (2019).

Appendix 1 – Data collection proforma used for the MCM5 study

CASE REPORT FORM
STUDY TITLE: FEASIBILITY STUDY OF THE USE OF ARQUER'S MCM5 ELISA TEST TO AID IN THE DIAGNOSIS OF OESOPHAGEAL CANCER USING GASTRIC ASPIRATES

STUDY SITE: University College London Hospitals
PRINCIPAL INVESTIGATOR: Dr. <u>Rehan Haidry</u>

I am confident that the information supplied in this case record form is complete and accurate data. I confirm that the study was conducted in accordance with the protocol and any protocol amendments and that written informed consent was obtained prior to the study.	
Investigator's Signature:	
Investigator's name:	
Date of signature:	

WITHDRAWAL: COMPLETE IF THE SUBJECT HAS WITHDRAWN FROM THE STUDY	
Date of withdrawal	
Reason	

Inclusion Criteria

Yes No*

1 Patients who have been referred for endoscopy and fall into one of four subgroups defined as: Control (Acid Reflux), Barrett's Oesophagus, Barrett's Dysplasia or Oesophageal Cancer

☐ ☐

2 Is the subject 18 years of age or older?

☐ ☐

3 Is the subject suitable for standard investigations as part of normal clinical practice?

☐ ☐

4 Is the subject able to understand the purpose of the study, give consent, and have a gastric aspirate taken?

☐ ☐

5 Has the subject willingly given written informed consent?

☐ ☐

* If any inclusion criteria are ticked 'No' then the patient is not eligible for the study.

Exclusion Criteria

Yes
* No

1 Patients who are undergoing chemotherapy for a previously diagnosed malignancy

☐ ☐

2 Has the subject had any instrumentation to the upper GI tract within the last 2 weeks?

☐ ☐

3 Has the subject already received treatment (chemotherapy or radiotherapy)?

☐ ☐

* If any exclusion criteria are ticked 'Yes' then the patient is not eligible for the study.

Year of birth:

CLINICAL INFORMATION		
Acid Reflux (Y/N)		
Non dysplastic Barrett's Oesophagus (Y/N)		
Dysplastic Barrett's Oesophagus (Y/N)		
Oesophageal Cancer (Y/N)		
<i>If Applicable:</i> Grade and stage of diagnosed cancer		
Result of any other relevant investigations	Cytology	
	Bloods	
	Other	

SAMPLE COLLECTION											
Sample Identification No:					-					-	
Clinical description of visit											
Date sample collected											
Time sample collected											
Total volume of gastric aspirate collected											
Time sample process and centrifuged											

Appendix 2 – Data collection proforma used for the OE study

ZOOM ENHANCEMENT ENDOSCOPY IN BARRETT'S OESOPHAGUS with OPTICAL ENHANCEMENT (ZEEBO 2 STUDY) RECORDING SHEET

PATIENT DETAILS

Name: _____
 Gender Male Female (please circle)
 Date of birth: / /
 Hospital No/Identifier: _____

ENDOSCOPY DETAILS

Endoscopy date: / /
 Endoscopist: _____
 Scope number: _____
 Recording Sequence No. _____

PRE-RECORDING CHECKLIST

N-Acetylcysteine applied? Y ☐ N ☐
 Simeticone (Infacol) applied? Y ☐ N ☐

Barrett's measurement: GOJ _____ cm C _____ cm M _____ cm

RECORDING CHECKLIST (please spend a minimum of 10 seconds during each enhancement mode)

PULL THROUGH RECORDING (Please tick)

HD-WLE recording: ☐

i-Scan OE recording: ☐

**Magnification view of
macroscopically normal
area in HD-WLE and OE
AND matched biopsy** ☐

WERE ANY ABNORMALITIES DETECTED? Y ☐ N ☐

If yes, please record each using **zoom function** and complete the below table for each abnormality.

Level	Location (please circle)	Lesion recorded in all enhancement modes?	Matched biopsy /EMR obtained?	M/V Score (please circle)	Comments (i.e. why lesion chosen?)
____cm	P L A R	Y <input type="checkbox"/> N <input type="checkbox"/>	Y <input type="checkbox"/> N <input type="checkbox"/>	M 1 2 V 1 2	
____cm	P L A R	Y <input type="checkbox"/> N <input type="checkbox"/>	Y <input type="checkbox"/> N <input type="checkbox"/>	M 1 2 V 1 2	
____cm	P L A R	Y <input type="checkbox"/> N <input type="checkbox"/>	Y <input type="checkbox"/> N <input type="checkbox"/>	M 1 2 V 1 2	
____cm	P L A R	Y <input type="checkbox"/> N <input type="checkbox"/>	Y <input type="checkbox"/> N <input type="checkbox"/>	M 1 2 V 1 2	
____cm	P L A R	Y <input type="checkbox"/> N <input type="checkbox"/>	Y <input type="checkbox"/> N <input type="checkbox"/>	M 1 2 V 1 2	

Appendix 3 – Protocol used for image capture and histology recording in OE study

ZOOM ENHANCEMENT ENDOSCOPY WITH I-SCAN and I-Scan OE TO DETECT NEOPLASIA IN BARRETT'S OESOPHAGUS (ZEEBO II)

Short Title: ZOOM ENHANCEMENT ENDOSCOPY with I-Scan OE IN BARRETT'S OESOPHAGUS (ZEEBO STUDY) - ZEEBO II

Protocol: Version 1.5 (19/10/2016)

M. Everson, R. Haidry

Introduction

Barrett's oesophagus (BE) is a change to the cells lining the oesophagus and occurs in about 10% of people with acid reflux disease. This condition is of great importance as it represents a pre-invasive precursor lesion to oesophageal adenocarcinoma. Adenocarcinoma of the oesophagus is a deadly disease with a 5 year survival of less than 15%. This cancer is increasing rapidly in the West with a 350% increase in incidence in the USA over the last 20 years, with a similar increase in the U.K. It is thought that 64-86% of all adenocarcinomas arise in BE, with the incidence of oesophageal cancer increased 30-100-fold above that for the general population. As such a major goal in recent years is the early detection of dysplastic changes in high risk patients with BE, to enable early therapeutic intervention and so prevention of the development or prognosis of adenocarcinoma.

High grade dysplasia (HGD) is at present the most robust routinely used clinical marker of cancer progression in BE; its presence confers a 16-59% risk of developing cancer within 5 years of the diagnosis of HGD. HGD is characterized by changes in the architecture and cellular components of the oesophageal mucosa, which may be detected microscopically and often macroscopically. There are now several minimally invasive treatment options that can be also delivered through the endoscope. Many patients with Barrett's oesophagus therefore undergo surveillance endoscopies at intervals to detect HGD in line with international guidelines. Conventional white light endoscopy is not sensitive enough to detect HGD alone and the standard approach is therefore to take one tissue biopsy in every quadrant of

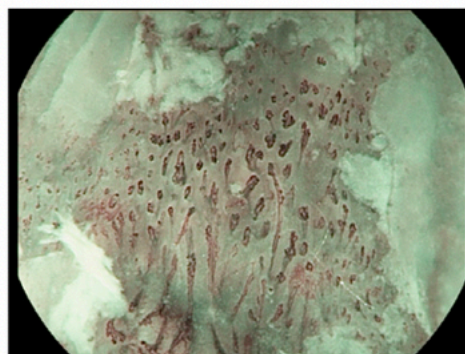
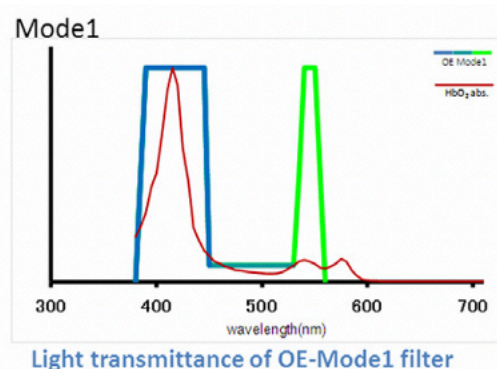
the Barrett's segment every 1-2 cm and send for histopathological review (Seattle Protocol). Analysis of biopsies is time consuming and costly, as such more accurate and cheaper methods to detect HGD are urgently needed.

A promising new technology developed by HOYA Pentax (Tokyo, Japan), optical enhancement imaging (OE) utilises novel post processing technology (i-scan) alongside newly developed optical filters, which emit wavelengths of light corresponding with the main absorption peaks of haemoglobin. This limited spectral transmission has been shown in other technologies such as narrow band imaging (NBI) to allow improved visualisation of microvascular structures within the gastrointestinal mucosa and confers greater colour distinction abnormalities within the mucosa. In addition this new technology delivers a continuous baseline transmission of light between three peaks in the red, blue and green wavelengths. This increased transmission allows the endoscope to provide higher intensity illumination of the target area, currently a limitation of NBI, thereby providing better quality imaging.

OE

Predominantly transmits blue and green wavelengths corresponding to the main absorption spectra of haemoglobin, but additionally raises baseline transmission of wavelengths across the entire visible light spectrum.

ESCN lesions typically demonstrate high levels of neovascularisation and as a result higher levels of circulating haemoglobin. We have previously demonstrated that zoom endoscopy in conjunction with i-scan post processing technology can detect dysplasia in BE based on the mucosal and vascular patterns. Now that this technology has developed further, we hypothesise that using OEI technology, in combination with acetic acid chromoendoscopy and enhanced optical zoom imaging, will allow even greater quality enhancement and distinction of microvascular and mucosal structures within BE tissue. This greater quality enhancement of mucosal structures, alongside higher intensity illumination, should lead to greater sensitivity in the recognition of neoplasia in BE tissue.



Although the benefits of OE in the detection and characterisation of have not yet been demonstrated in large human trials, this technology could offer an effective, time efficient screening tool for the detection of high grade dysplasia in patients undergoing assessment of BE neoplasia

Aims

The aims of this study are to:

1. Establish a simple endoscopic classification system for Barrett's Oesophagus (BE) related neoplasia (London BE classification)
2. To validate this classification system using a 'Pentax HiLine' series magnification endoscope using the new i-scan OE optical imaging platform.

Inclusion Criteria

1. Patients undergoing endoscopic assessment for BE-related neoplasia across 3 separate sites – University College Hospital London (UCLH), Nottingham University Hospital, University Hospitals Leuven.
2. **Patients with or without known low grade dysplasia (LGD), high-grade dysplasia (HGD) or early cancer undergoing endoscopy.**
3. All procedures must be performed using a 'Pentax HiLine' series Magniview endoscope with i-scan OE
4. Patient able to consent.
5. Minimum BE segment of C0M2 or C1M1 as per the Prague Classification.

Exclusion Criteria

1. Patients in whom endoscopy and biopsy is contraindicated
2. Patients who are unable to give informed consent.
3. Pregnant women.
4. People under the age of 21 years.

Classification System

The previously validated classification system will utilise a simple tool by which mucosal and vascular patterns are analysed such that:

Mucosal

M1: regular circular, oval or villous/gyrus-like ridged pits

M2: distorted, irregular pits or featureless mucosa

Vascular

V1: regular and uniform vessels

V2: irregular or dilated 'corkscrew-like' vessels

Dependant upon individual's interpretation of the mucosal and vascular patterns, the following combinations and likely diagnosis are possible:

Table 1 – possible outcomes using classification system and subsequent likely diagnosis

MV Classification	Likely diagnosis
M1V1	ND-BE
M1V2	D-BE
M2 V1	D-BE
M2 V2	D-BE

* ND-BE – non-dysplastic Barrett's oesophagus

* D-BE – dysplastic Barrett's

Endoscopy and lesion recording protocol

All endoscopy procedures will be performed by one of six consultant gastroenterologists:

- Dr RJ Haidry, UCLH (UK)
- Dr R Sweis, UCLH (UK)
- Dr K Ragunath, Nottingham University Hospital (UK)
- Dr J Ortiz Fernando Sordo, Nottingham University Hospital (UK)
- Prof R Bisschops, University Hospitals Leuven (Belgium)
- **Dr J Esteban (Madrid)**

Pre-recording steps

1. Please record patient, device (i.e. scope model number) and endoscopist details on the **recording sheet** provided.
2. Please ensure you have the necessary space on your hard drive recorder before starting the procedure.
3. Prior to recording and capturing lesions of interest, please ensure you have washed the entire BE segment with N-Acetylcysteine (2%) and Simeticone (Infacol).
4. Ensure you have accurately measured the Barrett's segment using the Prague Classification.

Recording steps

1. Please manoeuvre your endoscope so that it is at the level of the gastro-oesophageal junction (GOJ).
2. Make a note of the level of the GOJ on the recording sheet (see below).
3. Ensure your endoscope is set to plain white-light endoscopy (WLE).
4. **Switch on** your recording device

5. Now perform the '**pull through**' procedure – beginning at the GOJ, retract your endoscope through the entire segment of Barrett's mucosa. You should aim to spend approximately 10 seconds for every 1cm of Barrett's present (i.e. it should take 30 seconds to retract through a 3cm segment of Barrett's segment).
6. Repeat the pull through (starting at the GOJ) using **i-Scan OE**
7. If you identify any lesions of interest please make a note of their level and orientation (i.e. 32L – 32cm, left oesophageal wall)

Magnification recording

Once you have performed and recorded the pull through procedure using all enhancement modes, you are now ready to begin recording with magnification endoscopy on any lesions of interest you have already highlighted.

1. Return to any lesions you have highlighted during your pull through procedure. For cases in which you plan to record more than one lesion using the zoom scope, it is important you highlight the order in which you have recorded these lesions on the recording sheet - **it needs to clear to the person reviewing the videos which lesion corresponds to which recording.**
2. Once you have arrived at your lesion of interest please **switch on** your recording device.
3. **Switch on the zoom function** on your endoscope and take a detailed recording of the abnormality in question (aim to spend at least 30 seconds doing this) **using HD-HD WLE followed by i-Scan OE mode**
4. Switch off the zoom function and take a biopsy from the area you have just recorded.
5. Please ensure you send this biopsy in its own, separate formalin pot, which has been correctly labelled with the location from which it has been taken.
6. Once complete, **switch off** your recording device.
7. Please repeat steps 2-6 for every lesion you zoom in on and record.

Obtaining non dysplastic control images with matched histology

1. Identify an area of the BE that appears macroscopically non-dysplastic
2. Use magnification view to obtain still images of macroscopically non dysplastic area in iScan OE modes
3. Please take one matched biopsy from non dysplastic area, ensure this sample is clearly labelled

Post Magnification recording

Once you have completed your pull through and zoom scope recordings (with matched biopsies), please perform any further biopsies (or treatment) as per routine clinical care. These do **not** need to be recorded for the purposes of this study.

Post endoscopy and recording

1. Once the procedure is complete please review your recording to ensure all information has been captured correctly.
2. Please ensure you have fully completed the recording details not forgetting to make a note of the recording sequence reference.
3. Please file away the recording sheet safely in a secure area as it contains confidential patient information.

Editing and uploading videos to 'DropBox' anonymously

Once videos have been recorded, they will need to be edited so that they are preferably no longer than **90 seconds**. This can be done using the video editing software of your choice (i.e. for Windows use 'Windows Movie Maker' and Apple 'iMovie'). Please ensure you save the video in a universal format that will be able to be viewed by all parties. The most common viewing format is **.mp4**.

Once videos have been edited and converted to the appropriate format, please upload all videos to the shared DropBox folder. Please ensure all videos are anonymised and contain no patient identifiable information. It is important you organise videos in a format that allows you to go back and establish the eventual histopathological diagnosis.

Appendix 4 – Full dataset for MCM5 study

Histology	AR	NDBE	HGD	CANCER
MCM5 expression level	7.3551	22.9557	230.1155	300
	9.1615	47.3419	5.5487	148.9923
	18.0291	2.4286	7.4371	3.3635
	9.4077	262.8768	20.4925	500
	13.5132	98.4955	8.78325	9.76625
	10.0646	23.6947	67.04486	6.197667
	3.4139	3.0854	250	428.5714
	15.54283	4.5316	300	166.6667
	2.082375	26.81663	2.760889	31.92667
	8.660917	0.371	94.8099	110.277
	81.00919	10.39283	1.354	12.4865
	20.99688	3.1135	189.4128	86.0635
	2.729	187.5	45.03	187.5
	0	3.011214	300	3.829333
			0.007167	194.5525
			300	
			198.1853	
			0	
Median	9.2846	16.67427	56.03743	110.277
Mean	14.42616	49.75824	112.2768	146.0129

Appendix 5 – Peer reviewed publications arising from this thesis